CHAPTER 6

THE EQUILIBRIUM ANALYSIS OF ALTERNATIVE HOUSING ALLOWANCE PAYMENTS*

Other chapters have described quasi-static analyses of housing market behavior and policy issues that may be confronted by a national housing allowance program. However, any such program is likely to be of sufficient size to alter significantly both demand and supply conditions. Hence a long-run analysis is needed to determine the market implication of the study's findings regarding demand and supply behavior. For example, the partial-static analyses have found income elasticities of demand for housing substantially less than unity, and price discrimination by landlords towards various types of households. Clearly, these findings have profound implications for housing allowances—lower demand elasticities suggest less interest on the part of allowance recipients to spend additional income on housing. However, quantitative estimates of how these findings influence the price, quality and earmarking effects of specific allowance payment formulae requires simultaneous consideration of both supply and demand behavior. Toward this end, part of the study involved the simulation of various housing allowance payment formulae using a ten-year model of urban housing markets.

*This chapter represents the joint research of both Joseph Ferreira, Jr., and Dennis W. Carlton. We gratefully acknowledge the able assistance of Dennis Fromholzer who performed many of the computations needed to carry out this study.
1.1 Introduction

Unfortunately, no housing market simulation model was sufficiently developed by 1973 to be easily adopted for studying the long-run implications of the Joint Center's static analyses. Developing and calibrating such models is an enormous (and recent) undertaking, particularly when many non-classical market assumptions are needed.

After undertaking a review of existing housing market simulation models, it was decided, as described in the January 1973 interim report, that the Urban Institute Housing Model was the most suitable for our simulation analysis. Though simple in theory, the model allowed for neighborhood externalities, household characteristics, price and quantity changes and demand elasticities.

Though the theory underlying the Urban Institute model is rigorous and well suited to our needs, the simulation model was still in the developmental stages. Certain technical and mathematical problems with the model's solution algorithm hampered its use and complicated efforts to modify and disaggregate the model (see Technical Appendix *). Faced with time and cost limitations and the need to develop and calibrate parts of the model, we focused on studying the effects of alternative characterizations of behavior. Emphasis was placed on redefining the Urban Institute model's utility function to accommodate income elasticities less than unity and on modifying the ad hoc method used in calibrating the permanent incomes originally used in the model. The modified model was recalibrated to approximate the Pittsburgh housing market and used to analyze the impact of several income-gap and percent-of-rent

* The Technical Appendix will be drafted for the final report.
housing allowance payment formulae. The assistance and cooperation of the Urban Institute Housing Study Group was particularly helpful in enabling us to rapidly modify, calibrate and use a version of their model on the MIT-Harvard computing facilities.

Time permitting, we would have liked to have further disaggregated the zonal and neighborhood externalities of the model in order to examine the housing allowance implications of market segmentation and "crisis ghetto" theories. A discussion of how this might be done is contained in the appendices. As we shall see, such an effort requires improvement in the model solution algorithm. Suggestions along these lines are also in the Technical Appendix.

The next section briefly describes the Urban Institute Housing Model. Section 1.3 outlines the remainder of the report.

1.2 The Urban Institute Housing Model

The Urban Institute Housing Model which is modified for use in this chapter is described in detail in "The Distribution of Housing Services," Urban Institute Working Paper #208-6, by Frank de Leeuw. This section summarizes the theory, solution algorithm, and calibration procedure of the model. It provides background information that will be cited in the remaining parts of this chapter.

The essence of the UI model is a micro-economic picture of households and landlords contracting for housing at qualities and prices determined in several sub-markets. The model considers a hypothetical city in which housing is located in any of several residential zones. These zones differ insofar as the model is concerned
in their housing stock in the characteristics of their residents and in their exogenously
specified values for $t(z)$, the average travel time to work for residents of zone $z$.

On the supply side, a housing unit is characterized by the "quantity of service"
it provides (an assumed one-dimensional combination of size and quality) and the
price per unit of service charged as rent. For all housing, the operating cost per unit
of service ($P_o$) and capital cost per unit of service ($P_c$) are assumed to be constant.
Several model housing units are distinguished corresponding to alternative types of
units in the actual market. Landlords profit-maximize in selecting the quantity of
service $Q(j)$. (The model is oriented toward a rental market though some provision is
made for homeowners.) The actual supply curve assumed for model housing unit $j$ is
given by

$$Q (j) = \left[ \beta_1 + \frac{2}{3} \beta_2 \left( \frac{P(j)}{P_o} - 1 \right) \right] Q_o (j). \quad (1.1)$$

where $Q(j)$ and $P(j)$ are market levels of housing services and price per unit of service,
respectively, $Q_o (j)$ is the initial level of housing service, and $0 < \beta_1, \beta_2 < 1, P(j) > P_o$.
Note that $(1-\beta_1)$ is a depreciation rate -- how much $Q$ would drop during one
time period if $P(j) = P_o$, the minimum operating cost per unit of service. The other
parameter $\beta_2$ affects the slope of the supply curve. Figure 1.1 illustrates the case
where $\beta_1 = 0.4, \beta_2 = 0.90, P_o = 0.44, and P_c = 0.77$. Four housing units are
shown with $Q_o (j) = 20, 30, 125 and 150$ units of service initially. Note that all
supply curves intersect the $Q(j) = 0$ axis at $P(j) = P_o - \left( \frac{3}{2} P_c \beta_1 / \beta_2 \right)$ and hence,
supply curves can never be parallel.
Figure 1.1

TYPICAL SUPPLY CURVES

\[ Q(j) = [\beta_1 + \frac{2}{3} \beta_2 \left( \frac{P(j) - P_o}{P_c} \right)] Q_o(j), \]

where \( \beta_1 = 0.4, \beta_2 = 0.9, P_o = 0.44, P_c = 0.77. \)
An unlimited amount of new construction is available in one homogeneous
model zone at the price $P_n = P_o + P_c$ per unit of service.* Thus, the supply curve
for new construction is a perfectly elastic horizontal line at price $P_n$.

On the demand side, model households are characterized by any one of four
different preference schedules (reflecting age, family status, and race) and by
current income. These preference schedules are utility functions which compare
(using a multiplicative form) various amounts of housing expenditure, disposable
income, leisure time, neighborhood income, and racial composition. The specific
form of the utility function $U_{ij}$ specifying the utility of household $i$ for model unit
$j$ is given by

$$
U_{ij} = \left( Q_j - \gamma_i \alpha_i \frac{\gamma_i}{P_n} \right)^{\alpha_i} \cdot \left( \gamma_i - P_j Q_j - \gamma_i (1 - \alpha_i) \gamma_i \right)^{1-\alpha_i}
\cdot \left( 200 - t(z_j) \right)^{0.5} \cdot \left( \alpha_i - \alpha_j \right) \cdot l_j \cdot R_{ij},
$$

where

$$
l_j = \left[ \frac{Q_j (z_j) (P (z_j) - P_o)}{Q (P - P_o)} \right]^{.01} \gamma_2
$$

$$
R_{ij} = \left[ W (z_j) + \frac{1000}{100 \gamma_3 + 1} \right] \gamma_4
$$

*Current UI research on the model is disaggregating the new construction zone
and further distinguishing demand behavior according to tenure.
and previously undefined variables are given by,

\[ Y_i = \text{income of household } i, \]

\[ Q(j) = \text{average } Q \text{ of houses in zone of unit } j, \]

\[ P(j) = \text{average } P \text{ of houses in zone of unit } j, \]

\[ W(j) = \text{the fraction of (white) households in the zone of unit } j \text{ when house } i \text{ is (white),} \]

\[ Q \cdot (P_j - P_o) = \text{average net rent (less operating costs) of all units.} \]

The parameters \( \alpha_1, \gamma_1, \gamma_2, \text{ and } \gamma_3 \) indicate relative preferences for housing, other commodities, leisure time, neighborhood quality and racial composition. Only \( \alpha_1 \) depends upon household type and \( 0 < \alpha_1 < 1 \). Other constraints are \( 0 \leq \gamma_1 \leq 1 \), and \( \gamma_2, \gamma_3 \geq 0 \).

In comparing this model with classical equilibrium models, it is important to note that (1) the price, \( P \), charged per unit of housing service, \( Q \), is not assumed to be constant for existing model dwellings, and (2) the new construction option which allows a household to purchase any amount of \( Q \) (above a minimum) at a fixed price \( P_0 + P_o \) per unit of \( Q \) enables the model to avoid a situation where the rents of existing dwellings would be undetermined because the number of dwellings equals the number of households, and all households are forced to select exactly one dwelling.

To date, the Urban Institute Housing Study Group has calibrated the model for five test cities to simulate ten-year changes in each city's housing market. The
parameters in (2) relate to rent/income ratios and are estimated from regional Census data relating rent paid to family characteristics and income. To account for homeowners, a rent/value ratio is used to convert values into rents. This ratio is the same for all zones and is based on a hedonic price regression for the state containing the test city.

The operating and capital costs per unit of service, \( P_0 \) and \( P_c \), are estimated from FHA and BLS data. The quantity of service initially provided by model dwellings is estimated from Census tract data on average rents after adjusting for locational factors and the relative cost per unit of service for housing compared to that of other goods. Census tract data is also used to generate model households distinguished by type and income. Zonal boundaries are chosen to include a contiguous group of Census tracts that is as homogeneous as possible. The number of zones is generally predetermined to be four or five, to provide as much spatial distinction as possible without sacrificing computational ease.

The remaining household behavior parameters \( \gamma_1, \gamma_2, \) and \( \gamma_3 \), and the remaining supply parameters, \( \alpha_1 \) and \( \alpha_2 \), are estimated by fitting model runs for 1960 and 1970 to actual data. For the 1960 run, \( \alpha_1 \) and \( \alpha_2 \) are set equal to 1.0 and 0.0 to reflect inelastic short-term conditions. The model is initialized for 1960 prices and housing stock and the gammas which "best" match households to 1960 assignments are estimated. (Here, "best" means the lowest normalized "root mean square error" averaged over several criteria.) To estimate \( \alpha_1 \) and \( \alpha_2 \), the model is run with 1960 starting conditions of the fitted gammas. Those \( \alpha_1 \) and \( \alpha_2 \) values
which yield the "best model results for 1970 are selected. This procedure assumes
that 1960 and 1970 market conditions are close to "equilibrium."

A typical run of the Urban Institute model involves about 40 households (combinations
of the four basic types with different incomes) and a similar number of
housing units. Let us refer to this discrete set of households and dwellings as "model"
households and "model" housing units. Initially, the model assumes a housing stock
of certain "initial" characteristics and location, and a resident population with speci-
fied "current" characteristics. Housing allowance programs are translated into altered
incomes and relative prices (and hence utilities) for recipient households and for
other households whose taxes are increased to provide the subsidy.

Solving the model requires adjustment of the prices and the quantities of service
provided by the housing market until a satisfactory housing assignment is found for
each "model" household. An assignment is defined to be satisfactory if (1) no un-
occupied, existing house is priced above \( P_o \), (2) no two households are assigned the
same unit, and (3) households prefer their assigned house to any other existing house
(at its offered price and quantity) and to any new house (unless they are assigned
an identical new house.) During each iteration of the model's solution algorithm,
a specific set of prices for the "model" housing units is checked to determine whether
the household choices they induce constitute a satisfactory assignment.

Iterations of the Urban Institute model are not intended to correspond to elapsed
time in the actual housing market. If, during an iteration, no household (or more than
one) chooses a particular unit, then the price per unit of service of that housing unit
Is lowered (or raised) for the next iteration. Although this procedure appears to represent a bidding process, the time unit corresponding to one iteration is not specified, and the solution procedure begins with high prices not necessarily related to what would be observed when a housing allowance program begins and transaction costs are not considered.

A two-step solution procedure is presently used. First, each housing unit is "offered" at its current price beginning with the unit with highest initial $Q_o$. Unassigned households (and certain tentatively assigned households) consider each offering and tentatively select a particular unit if it provides a $U_{ij} > U_{ik}$ for any other unit $k$ at that unit's price, $P_k$. The utilities used for this comparison reflect neighborhood characteristics that existed prior to the current series of offerings. Assignments made in this way are "frozen" until all units have been tentatively assigned.* This procedure prevents altered utility functions due to changing neighborhood characteristics from adjusting these tentative assignments before all units have been offered. One such offering of all existing units constitutes one "major iteration."

At the end of each major iteration, zonal wealth and transportation characteristics affecting the utility function are updated and the tentative assignment procedure begins again with the unit with the largest $Q_o$ again offered first. To facilitate convergence and enable various solutions with different racial composition to be identified, each zone's racial characteristics are preset and unchanged throughout the

*For reasons discussed more extensively in the Technical Appendix, a few households are left unfrozen during minor iterations.
solution process. Solutions with racial compositions much different from the present values must be discarded and rerun after adjusting the predetermined percentages.

If four successive major iterations meet the solution criteria, the "offering" process is continued without "freezing." The procedure stops when four successive major iterations that do not allow freezing continue to meet the solution criteria.

In calibrating the model for a particular city, one must select a particular level of aggregation by choosing the number of zones, the number of "model" households and dwellings and the number of household types that are distinguished. The time required to identify equilibrium solutions increases exponentially with the number of "model" households. To make the Pittsburgh model manageable, the Urban Institute Housing study group characterized the Pittsburgh housing market in 1970 in terms of 41 "model" households and 39 "model" dwellings that existed prior to 1960. The model city was assumed to encompass the area included in Allegheny and Westmoreland Counties. Four zones were distinguished: a "poverty" zone labeled #1; a somewhat dispersed zone "2" including the "better" part of the central city; and two suburban zones 3 and 4. A finer disaggregation of zones complicates the model's solution since the distinctions among zones would decrease. Also, the number of "model" dwellings in each zone would drop thereby reducing the detail with which each zone's housing is characterized.

The model's solution algorithm prohibits fractional assignments of "model" households to "model" dwellings. Hence, each "model" household and dwelling must correspond to an equal number of "actual" households and dwellings. For Pittsburgh the ratio is about one to 15,000. This integer assignment requirement implies that a swap of assignments between two "model" households is the smallest change that can be distinguished. In a zone such as #2 with only four "model" dwellings, such a change can markedly affect the zonal characteristics of the solution if, for example, the households involved in the swap are of different race. Hereafter, we shall refer to this limitation as the "discreteness problem."*

The four household types in the Pittsburgh calibration distinguish white and black headed households broken down by multiple-person families with heads under age 65 and single-person households or families with heads over age 65. Household types are assumed to differ only in the alpha-parameter values of their utility functions. Hence, distinguishing additional types would not complicate the solution algorithm. However, household types that are uncommon in Pittsburgh cannot usefully be distinguished since each "model" household represents about 15,000 actual households and several households of each type are needed before their solution assignments can be usefully interpreted.

In modifying and calibrating the Urban Institute model, we retained the same level of aggregation, zonal boundaries and household types. The next section identifies

*Several such problems and possibilities related to the model's solution algorithm are discussed in the Technical Appendix to this chapter.
the changes that were made and outlines their discussion in the remainder of this chapter.

1.3 Outline of the Report

In the Urban Institute Housing Model, household preferences, as described by the Cobb-Douglas type utility function of (1.2) assume a unit elasticity of demand for housing with respect to changes in permanent income. The permanent income of "model" households is set equal to the product of the household's current income raised to the 0.6 power and the average income of the household type raised to the 0.4 power — a choice of exponents designed to match observed and predicted housing expenditures as well as estimate permanent income.

An alternative procedure for estimating the permanent income of model households is described in Part 2. Using it enables permanent income adjustments and income elasticity assumptions to be separated. The econometric technique is of significant theoretical interest since it appeared to work quite well and yet avoids the problem of requiring time-series income data for a particular population in order to estimate their distribution of permanent income.

The theory underlying modifications of the household utility function is described in Part 3. The assumptions about household expectations are altered to allow income elasticities less than one. Elasticities of approximately 0.5 were chosen based on independent estimates described previously in Chapter 4. Pittsburgh Census data is then used to estimate the \( \alpha \) parameters of the modified utility function. Finally, the permanent income distributions for the modified model are compared with those of the Urban Institute's calibration for Pittsburgh.
Parts 4 and 5 describe the calibration of the five gamma and beta parameters of the model. The gamma estimates were similar to those of the Urban Institute, but the beta parameters had very different supply-side interpretations.

Part 6 describes the results of using this calibration of the modified model to simulate several income-gap and percent-of-rent housing allowance formulations. For purposes of comparison additional simulations were run which in effect assumed that recipients react to housing allowances as if their income elasticities were unity. The 0.5 income elasticity significantly limited the impact of the allowances and altered the comparison of income-gap and percent-of-rent forms.

Conclusions concerning the appropriateness of the modified model are discussed in Part 7. Conclusions and recommendations regarding housing allowance payment formulas are organized in Part 8 into two sections. The first concerns the effect of housing allowances in Pittsburgh. The second concerns recommendations regarding the appropriate form and level of allowances in various market situations.
A Method of Determining the Distribution of Permanent Income Based on Cross-Sectional Observation of Current Income

In using the Urban Institute Housing Model for Pittsburgh, it is necessary to obtain the distribution of permanent income for various demographic groups in that SMSA for the years 1960 and 1970. As is quite common, cross sectional data on the distribution of current, not permanent, income is available for the particular year in which we are interested. This paper describes an approach to infer, for any particular group, the distribution of their permanent income from the distribution of their current income, for any year in which data on their current income distribution is available.

Aside from its obvious importance in the application of any housing model, the problem of estimating the distribution of permanent income from that of current income is an important one which arises quite frequently in empirical work in all branches of economics. Therefore, the method proposed in this paper to solve this problem should be of quite general interest.

This paper is organized as follows. Section 1 describes why methods which are frequently used in the study of permanent income and which are based only on the one year of cross sectional data, are inadequate to deal with the particular problem of estimating the distribution of permanent income. Section 2 describes an alternative

[I] Data on current income is available from the Census at 10 year intervals, for the largest SMSA's.

* This part of the study is the research of Dennis W. Carlton. Joseph Ferreira Jr. provided valuable comments, and Dennis Fromholtzer greatly assisted in the computations.
approach which makes use of an additional independent data source, namely the Michigan Panel study. The details of this new approach, together with the variables and data used, are then presented in Sections 13 and 24. Experiments to investigate the robustness of the results are described. Finally, in Sections 25 and 26, the results and their implications are discussed. Appendix 1 shows how the method of Section 24 can be extended to obtain a statistically more sophisticated (but much more complicated) treatment of the problem. It is shown that the problem reduces to the quite interesting methodological problem of estimating the characteristic function of a distribution. Appendixes 2, 3 and 4 appearing on pages 6-83 through 6-87 deal with some technical points discussed in the text of Part 2.
Section 2.1 - Inadequacy of One Typical Approach

In this section there is a discussion of why one typical approach of dealing with permanent income is not well suited for the problem of estimating the distribution of permanent income. The purpose of this section is not to criticize any particular study, but rather, through a discussion of a typical approach, to bring into sharper focus the main issues and pitfalls involved in the estimation of the distribution of permanent income from that of current income.

Let \( Y_p, Y_t, Y_c \) stand for permanent, transitory, and current income respectively at some particular point in time. As described above, the problem is to determine the distribution of \( Y_p \) from that of \( Y_c \), for various demographic categories, which are described by a vector of attributes \( x \). By definition

\[
(1) \quad Y^i_c = Y^i_p + Y^i_t,
\]

for each individual \( i \).

For the moment, assume that \( Y_p \) and \( Y_t \) are independent and that \( E(Y_t) = 0 \). Furthermore, assume that for any individual \( i \) belonging to group \( j \), with characteristics \( x_j \), his permanent income equals a systematic component which depends on \( x_j \), namely \( Y^{i,j}_p x_j \) plus an individual component, namely \( Y^{i,\text{ind}}_p \), which is independent of \( x_j \), and which depends on certain, unobservable, characteristics of individual \( i \), such as his productivity, ambition, etc. Hence, we can write

\[
(2) \quad Y^i_p = Y^{i,j}_p x_j + Y^{i,\text{ind}}_p,
\]
By absorbing a constant into \( Y_{p,x_j} \), we can assume that \( E(Y^i_{p,ind}) = 0 \) (expectation is over \( i \)). To say that permanent income has a frequency distribution for a certain demographic group \( x_j \), is equivalent to saying that at any particular time, \( Y^i_p \), or more precisely, \( Y^i_{p,ind} \), has some probability distribution across individuals \( i \) in group \( j \).

If the systematic component, \( Y_{p,x_j} \), could be estimated, then the \( E(Y^i_p) \) could be predicted for a randomly chosen individual belonging to the demographic group \( j \).

However, knowledge of \( Y_{p,x_j} \) does not provide any information about the distribution of the random variable \( Y^i_{p,ind} \). In other words, the ability to summarize the entire income distribution of group \( j \) by its expected value, namely \( E(Y^i_p) = Y_{p,x_j} \), does not enable one to obtain any idea about the distribution of \( Y^i_p \) for individuals \( i \) within group \( j \). It is important to investigate the distribution of \( Y^i_p \), because not only the mean, but also the variability, or, more precisely, the particular shape of (permanent) purchasing power will have a crucial effect on the type of behavior that occurs in the market place.

In order to determine the distribution of \( Y^i_p \), not only the deterministic variable, \( Y_{p,x_j} \), but also the stochastic variable \( Y^i_{p,ind} \) must be investigated. But, as can readily be seen from eqs (1) and (2), it is impossible to distinguish the random variable \( Y^i_t \) from the random variable \( Y^i_{p,ind} \), based only on one year of cross-sectional data for group \( j \). Therefore, it is hopeless to attempt to infer the distribution of

---

1. As discussed above, random variables are random over \( i \).
$Y_p^i$, and hence $Y_p^i$, using only cross sectional data on group $i$ for one year. Similarly, any attempt to predict $Y_p^i$ based on data from a single year, cannot account for the effect of variables which are constant over that year. For example, abnormal economic conditions in one particular year will certainly affect $E(Y_p^i)$, but need not affect $E(Y_p^i)$, since $Y_p^i$ is viewed as a long run average of income over several years. Once we allow that $E(Y_p^i) 
eq 0$ for a particular year, then from eq (1), one can see that $Y_p,i$ and $Y_p,x_i$ can not even be identified from cross-sectional data.

In summary, only if $Y_p^i$, and not just $Y_p,x_i$, can be observed or estimated, can the relation between the distribution of current and permanent income be investigated. It is only through the use of a cross-sectional time series that one will be able to estimate $Y_p^i$, and then be able to discern the relationship among the distributions of $Y_c^i$ and $Y_p^i$ for the various $X$ values, which characterize the demographic groups that one is interested in. The next section will explain how these relationships can be determined from longitudinal data and how they can be used to estimate the distribution of $Y_p^i$ from available information on the distribution of $Y_c^i$, for the particular time, place, and group in which this study is interested.

---

1. Simply because the approach characterized by eqs (1) and (2) is not fruitful for estimating the distribution of $Y_p$, does not mean that the approach is never useful. In fact, the reader probably recognizes the close connection to the instrumental variables method to estimate coefficients of independent variables, subject to measurement error.
Using the notation of the previous section, let us again consider

\[ Y_c^i = Y_p^i + Y_t^i. \]

This system holds for all individuals \( i \), and it is assumed that \( Y_t^i \) is a random variable, independent of \( Y_p^i \), and independent and identically distributed for individuals in the same group.

Suppose that data is available on the empirical distribution of \( Y_c^i \) for each of the \( J \) groups, characterized by the vector \( x_i, i=1,...,J \). The distribution of \( Y_c \) for each of these \( J \) groups can be estimated from this data. Also, suppose that an estimate is available of the distribution of \( Y_t^i \) across the members of a particular demographic group, group \( J_0 \). Since we now have the estimated distributions of \( Y_t \) and \( Y_c \) for this group, it should be possible to derive the distribution of \( Y_p \) for this class. This can be done because relation (3) expresses \( Y_c \) as a convolution of \( Y_p \) and \( Y_t \). Letting \( F_{yt} \), \( F_{yc} \) stand for the estimated characteristic functions of \( Y_t \) and \( Y_c \), respectively for group \( J_0 \), we can use (2) to write:

\[ F_{yp} = [F_{yt}]^{-1} F_{yc}, \]

where \( F_{yp} \) is the characteristic function of \( Y_p \) in the group \( J_0 \). Relation (3) uniquely determines the distribution of \( Y_p \). The density function of \( Y_p \) can be found by inverting the Fourier Transform (3). Notice that this approach uses the

---

1. Such data is available from the Census, for the largest SMSA's, every 10 years.
distribution of $Y_c$ across a class to determine the distribution of $Y_p$ across that same class. There is no attempt to predict $Y_p^i$ for each individual $i$.

More generally, suppose that it is possible to estimate the dependence of the characteristic function, $F_{yt}$, of group $j$, on the vector of characteristics, $x_i$, which describe that group. Also, we may wish to allow economic conditions to affect $F_{yt}$, so that additional variables, $e$, may also influence $F_{yt}$. Then, since an estimate of $F_{yc}$ is available for each group, the density function of $Y_p$ for group $j$, can be estimated as

$$f_{yp} (x_{1i}, e_i) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-iuY_p} \frac{F_{yc}(iu, x_{1i}, e_i)}{F_{yt}(u, x_{1i}, e_i)} du, \text{ where } i = \sqrt{-1},$$

and where $F_{yc}$ refers to the estimated characteristic function for current income for group $i$.

Thus, unless additional structure is imposed on the problem, there will be two methodological issues to face. First, $F_{yc}$ has to be estimated from observed data. The second and more difficult problem is the estimation of the dependence of $F_{yt}$ on the vector of characteristics $(x_i, e)$. The rather complicated solution to this interesting problem is discussed more fully in Appendix I.

An alternative, and quite attractive, way of proceeding is to recognize that the difficult problem of estimating distribution and/or characteristic functions can be bypassed if some additional structure is imposed onto the relation expressed in eq (2).

More specifically, if the random variables $Y_p^i$ and $Y_t^i$ are constrained to all belong to the Pareto-Levy class of distribution\(^1\), then there is no need to estimate entire

1. This class of distributions has the property that it is closed under convolution.
distribution functions. Instead, there will be a relation between the finite number of parameters determining the distribution of $Y_p$, and the parameters of the $Y_t$ and $Y_c$ distributions. The parameters of the $Y_p$ distribution will uniquely determine its distribution. Such a procedure not only has the obvious computational advantage of greatly simplifying the problem, but also will allow us to obtain, without too much work, statistical confidence regions for the parameters of the $Y_p$ distribution. The Pareto-Levy family of distributions has been used quite frequently in economic studies, especially of the distribution of income.

Although useful for exposition purposes, the assumption associated with eq (2), that $Y_t^i$ is independent of $Y_p^i$, is not a reasonable way to describe the problem. It is more reasonable to assume that the variability of $Y_t^i$ depends positively on the magnitude of $Y_p^i$, where rich persons experiencing larger absolute deviations in $Y_t^i$ than poor persons. A quite natural way of formulating such a relationship would be

$$ Y_c^i = Y_p^i (1 + v_i) $$

where $Y_t^i = v_i Y_p^i$, and $v^i$, not $Y_t^i$, is assumed to be independent of $Y_p^i$. Taking logs of both sides of eq (5) yields

$$ \ln Y^i = \ln Y_p^i + \ln (1 + v_i) $$

or since $1 + v_i = (Y_c/Y_p)^i$,

$$ \ln Y^i_c = \ln Y^i_p + \ln (Y_c/Y_p)^i. $$

Eq (6') brings us right back to the case of a linear equation which expresses one random variable as the convolution of two independent random variables. The "transitory" component of income is now interpreted to be $Y_c/Y_p$. Suppose that $Y_c^i$ and $(Y_c/Y_p)^i$ are both assumed to follow log-normal distributions across the
members of any class 1. Once log normal distributions for \( Y_c \) and \((Yc/Yp)^i\) have been estimated a simple relation between the parameters of the distributions of \( Y_c^i \) and \( Y_p^i \) can be used to derive the distribution of \( Y_p \). Specifically, the two defining parameters of the \( Y_p \) distribution, can be estimated as

\[
(7) \quad u \ln Y_p = u \ln Y_c - u \ln (Yc/Yp), \quad \text{and}
\]

\[
(8) \quad \sigma^2_{\ln Y_p} = \sigma^2_{\ln Y_c} - \sigma^2_{\ln (Yc/Yp)}, \quad \text{where}
\]

\( u \) and \( \sigma^2 \) stand for the mean and variance of the relevant normal distributions. 2 Thus, through the assumption of log-normal distributions, the quite difficult problem of estimating the functional dependence of a characteristic function on \( \gamma_i \), the vector of characteristics of group \( i \), and then inverting a Fourier transform, is replaced with the somewhat simpler problem of estimating the parameters of two log-normal distributions and determining the dependence of one set of these, i.e., those associated with the distribution of \((Yc/Yp)^i\) on \( \gamma \). Let us now turn to a discussion of a log-normal distribution.

Aside from its obvious convenience, the choice of log-normal distributions for income seem well justified for our problem, in light of previous empirical 3 studies which use the log-normal distribution to characterize the income distribution. Aitchinson and Brown conclude that regarding “the introduction of the distribution of incomes into econometric model...the log normal hypothesis seems to have considerable advantages over most other candidates. In our case, we will be fitting log normal income distributions not to

1. In relation to previous discussions, \( \ln Y_c^i \) and \( \ln (Yc/Yp)^i \) each follow a normal distribution, which is the most well known member of the Pareto-Levy class.

2. The two parameters \((u, \sigma^2)\) completely characterize a log-normal distribution. Relations (7) and (8) follow immediately from eq(6) and the assumptions associated with that equation.

3. For a theoretical derivation of a log-normal distribution of income from a stochastic model, see Aitchinson and Brown, The Log-Normal Distribution, Cambridge University Press, 1963, Ch. 11 "The Distribution of Inces."

the entire heterogeneous population, but rather to several stratified (and hence more homogeneous) demographic groups. This fact strengthens the case for the log-normal assumption for the distribution of $Y_c$ since 'the evidence...suggests that the more homogeneous the group of income recipients is, the more likely is the log-normal curve to yield a good description of the income distribution.'\textsuperscript{1} There is the sometimes-raised objection with regard to the use of the log normal distribution, that although it may fit well throughout most of the income range, it does not fit especially well at the high income range - its right hand tail often doesn't have enough weight. 'The finding of Quenel [D] that the log-normal curve is the better approximation in the lower range of income, whilst the Pareto curve is better in the higher range.'\textsuperscript{2}

Although underestimating the number of very rich is an undesirable feature of the log-normal assumption, it could very well happen that, depending on the use to which the estimates of permanent income will be put, this undesirable feature of the log-normal distribution will not present serious difficulties. For example, in the application of this research, which is to study the effects of housing allowances, there are at least two reasons why this objection to the log normal distribution is of small consequence. First, the purpose of this study, is to predict the changes in housing behavior, that occur from increasing the permanent purchasing power of the poor through a housing allowance. It is most important for us to correctly estimate $Y_p$ for those who might be eligible for a housing allowance - a category that definately excludes high income people. Second, in any housing market, (and in the Urban Institute's representation of it), the richest members of the population tend to purchase new homes. Now, the size of

\textsuperscript{1} Aitchinson and Brown, op. cit. p. 118.

\textsuperscript{2} Aitchinson and Brown, op. cit. p. 108.
that new house is of little consequence in predicting what will happen to the existing housing stock. Indeed, in the Urban institute model, the size of the new home a person purchases has zero effect in predicting the outcome of market behavior in the rest of the housing market, which is dominated by the existing stock of housing. The will tend to underestimate the income, for say, even the richest 10% of the population fact that the log-normal distribution will not matter, since these people will wind up in new housing anyway — even with their underestimated (though still high) incomes.

The equilibrium solution of the U.I. model would not be changed if these people had higher incomes and bought larger new homes.

Once it has been decided that the assumption of the log-normal distribution is adequate to describe the distribution of $Y_c$, there seems to be no reason to doubt, for much the same reasons, that it would also provide an adequate description of the distribution of $Y_p$. Hence, the assumption of log-normality for the relevant distributions in eq(6) seems reasonably well justified, and will, through eq(7) and (8), provide a quite useful vehicle for determining the distribution of permanent income.

The procedure that will be followed, then, is to estimate the parameters of a log-normal distribution for current income for the particular demographic groups in Pittsburgh in which this study is interested for the years 1960 and 1970. Then, the Michigan Survey will be used to investigate the parameters of the log-normal distribution of "transitory" income, namely $Y_c/Y_p$, for these same demographic groups. Relationships will be established between the parameters of the distribution of $\ln(Y_c/Y_p)$ and various economic and demographic characteristics. These relations will then be used to

---

1. A log-normal distribution for $Y_a$ and $Y_p$ implies a log-normal distribution for $Y_c/Y_p$ under the assumptions on p. 6-22.

2. Actually, the years are 1960 and 1970. Data for these years for Pittsburgh appears in the 1960 and 1970 Census. To keep the same terminology as the Urban institute, I will use "1960" and "1970" to refer to this Census data from Pittsburgh.
predict the parameters of the distribution of $\ln Y_c/Y_p$ for the demographic groups in Pittsburgh for the years 1960 and 1970. Once the parameters of the distribution of $Y_c/Y_p$ have been predicted for the demographic groups in Pittsburgh, eqs (8) and (7) will be used to combine these estimates with those of the parameters of the distribution of $Y_c$ to obtain the parameters of the distribution of permanent income for these demographic groups in Pittsburgh for the years 1960 and 1970. The next two sections of this paper will describe the details of the estimation of the parameters of the distribution of current income, $Y_c$, in Pittsburgh for the demographic groups in 1960 and 1970, and the estimation from the Michigan data of the functional dependence of the parameters of the parameters of the distribution of $\ln Y_c/Y_p$ on various demographic and economic variables.
2.3 The Log-Normal Distribution of Current Income, \( Y_c \), in Pittsburgh for Selected Demographic Groups in 1960 and 1970.

Theoretical Specification of an Equation to Estimate the Parameters of a Log-Normal Distribution

As mentioned in the previous section, it is necessary to estimate a separate log-normal distribution of current income \( Y_c \) for each of the selected demographic groups in Pittsburgh, for both 1960 and 1970. Data is available on the cumulative frequency points of the income distribution for each of these selected demographic groups. A decision must be made as to which estimation method should be chosen to estimate these log-normal distributions. Two of the most popular methods of estimating log-normals are 1) method of fractiles and 2) maximum likelihood estimation. In the method of fractiles, two fractiles of the cumulative frequency curve are observed, and then \( (\mu, \sigma^2) \) of the log-normal distribution are found so as to exactly generate these 2 fractiles.\(^1\) In maximum likelihood, it is assumed that the random variable \( Y_c \) is exactly log-normal, and \( (\mu, \sigma^2) \) are chosen to maximize the constructed likelihood function.

Both of these approaches have serious drawbacks. The method of fractiles is obviously inefficient (in the statistical sense), since it uses information on only two fractiles, and ignores other useful data that is available. The method of maximum likelihood is also deficient, because it totally ignores the fact that the income distribution is not exactly log-normal, but only approximately so. The failure to specify what

\[ \text{1. A log-normal distribution of a random variable } Y \text{ is completely characterized by two parameters } (\mu, \sigma^2), \text{ where } \mu = \text{E}(\ln Y), \text{ and } \sigma^2 = \text{var}(\ln Y). \]
"approximately" log-normal means is a drawback of the maximum likelihood procedure. 1

The estimation procedure, employed in this paper, will make use of an equation which relates \( \ln Y_c \) to fractiles of the relevant population. If \( Y_c \) has a log-normal distribution with parameters \((u, \sigma^2)\), then the following relation 2 holds between observed income \( Y_c \) and the fractile \( q \) of the population with income below \( Y_c \):

\[
(9) \quad \ln Y_c = u + \sigma \cdot v_q \quad \text{where} \quad v_q = \text{that point on a standard normal curve which corresponds to a fractile } q. \]

Stated in a slightly different way, eq(9) says that when a log-normal frequency distribution is graphed on "log-normal" paper, the result is a straight line ("log-normal" paper has \( \ln Y \) as ordinate, and \( v_q \) as abscissa). The slope of this straight line will be the standard deviation \( \sigma \), of \( \ln Y_c \), while the constant will equal the mean of \( \ln Y_c \).

To say that a distribution is approximately log-normal will be interpreted to mean that eq(9) approximately holds. In other words, the relevant equation to be estimated to determine \((u, \sigma)\) is:

\[
(10) \quad \ln Y_c = u + \sigma \cdot v_q + e \quad \text{where} \quad e \text{ is a stochastic error term.}
\]

---

1. This method is not the familiar "maximum likelihood" procedure with which econometricians are acquainted. Usually, the likelihood function is constructed from the distribution of the stochastic term (i.e., the "approximately" part of the relation). Here, there is no specification of the "approximately" part of the relation. It will soon be seen that the meaning of "approximately" must be specified beforehand, if one wishes to choose a consistent (in the statistical sense) estimation method.

2. Aitchinson and Brown, Ch. 3.

3. e.g. \( v = 0 \) means 50% of the distribution lies to the left, while 1.96 means 97.5% of the distribution lies to the left.
However, the analysis does not end here. It is still unknown whether \( v_q \) or \( \ln Y_c \) should be regarded as the dependent variable in eq(10). Interestingly enough, the answer to this question depends crucially on what is meant by "approximately" log-normal.

Consider the following two interpretations of what "approximately" log-normal means. If \( Y \) is approximately log-normal with parameters \((u, \sigma)\) then:

1. One can predict the mean \( q \ln Y \). The following question can be considered: What is the average income of the bottom quartile of the population?

2. One can predict the "normal-frequency percentage" that corresponds to \( \ln Y \). The following question can be considered: What is the fractile of the population with income below \( Y_c \)?

If interpretation 1 is appropriate, then \( \ln Y \) should be the dependent variable, while if interpretation 2 is appropriate, then \( v_q \) should be the dependent variable. It is quite difficult to decide which interpretation is the "correct" one. Both seem eminently reasonable. However, interpretation 1 seems to be a slightly more natural way of viewing the problem. (Note that eventually the estimated income distribution will be used to estimate average purchasing power for various fractiles, which is precisely the question that interpretation 1 seeks to answer.) Such reasoning suggests that eq(10) should be used to find the desired parameters, with \( \ln Y_c \) regarded as the dependent variable. Regarding \( \ln Y_c \) as the dependent variable in eq(10) will yield, under the usual normality assumption on \( e \) as maximum likelihood estimates of the parameters \((u, \sigma)\).

---

1. This time, the maximum likelihood procedure is the one econometricians know. Interestingly enough, the previous "maximum likelihood" procedure is not even consistent, unless a) interpretation 2 is adopted and most importantly, as the number of observations approach infinity, the mean of \( v_q \) goes to zero. This last condition is quite strong and emphasizes the importance of examining the stochastic assumptions. In other words, the "approximate" part of the relation before an estimation procedure is chosen.
The choice of which variables to regard as independent will affect the estimates of the coefficients \((u, \sigma)\). However, if \(R^2\) is high, then the coefficient estimates will be insensitive to whether interpretation 1 or 2 is adopted. Fortunately, as will be seen below, \(R^2\) is always very high for all estimated equations of the form of eq. (10).

Hence, although it is an important theoretical concern to decide whether interpretation 1 or 2 is more justified, practically, it is of little consequence in the estimation of the coefficients \((u, \sigma)\) when eq(10) is used.

DATA

The data that will be used to estimate the log-normal distributions of current income are cumulative frequency tables of the income distribution of the various demographic groups in Pittsburgh in 1960 and 1970. These tables are published in Metropolitan Housing Characteristics, HC-2 for 1960 and 1970. There are 9 and 11 demographic groups for 1960 and 1970, respectively, with separate tables for black/white, and owner/renter categories for both the city and SMSA. A demographic group is characterized by such traits as sex of head, age of head, marital status, and single/multiple person households. In 1970 certain categories (the first three of "male") were combined to make the demographic groupings identical to those of 1960.

Although the data tables usually provide 9 points on the income frequency curve for each group, the income tables for renters in 1970 is quite aggregated and provide only three data points. To remedy this unsatisfactory situation, an additional data source, the Population Census, table 198, was used to estimate a more refined

1. This is the same source that the Urban Institute uses to generate their income distributions.
breakdown of the distribution of income for renters.\(^1\) Unfortunately, several of the
demographic categories in the Population Census were not quite identical to those of
the Metropolitan Housing Characteristics, and, hence, certain assumptions and
corrections were made to make the two tables comparable. The details of these
corrections are more fully described in Appendix 2. One final point to note about
the data used, is that in the Metropolitan Housing Characteristics, the category
"other male head" appeared to be rather arbitrarily defined. Furthermore, the number
of people in this category was usually a small percentage of the total population
(3-5\%). Because of this small number, the category does not even have enough members
to constitute "one" household in the Urban Institute model. For these reasons, it
was felt best to simply distribute the "other male head" category among the category
"male head"\(^2\). Table 1 illustrates the 7 demographic groups x 4 categories which were
used in our estimating procedure. As discussed in \(\text{Section 2.2}\) the assumption that is being
made is that the distribution of current income for each of these groups in 1960
and 1970 can be adequately approximated by a log-normal distribution.

---

1. Of course, the census tape -6th count-for-Pittsburgh could have been used to
obtain the necessary data if the extra effort to obtain even more precision were warranted.

2. The details of how the category was distributed among the "male head" category
are described in Appendix 2.
Results of Estimation of Distribution of Income in Pittsburgh

Following the procedure described in the previous pages, log normal distributions were fitted to each of the seven basic demographic groups of Table 1 for owner/renter, and black/white categories for 1960 and 1970. The results of these 56 regressions are displayed in Table 2. The fit of these regressions to the actual data is quite good.

A few graphs of the income distribution of selected demographic groups are displayed on pg. 6-34. These graphs are plots of the group's income distribution on "log-normal" graph paper. The plotted points for these two graphs, as well as all most of the other 56 graphs, seem quite clearly to lie close to a straight line. As explained in the previous section, such a result lends support to the use of the log normal distribution as an approximation to the distribution of current income. However, a rather consistent feature of the 56 graphs (of which only 2 are displayed) is that the first point tends to fall below the fitted line. This means that the log-normal distribution tends to underestimate the number of people with extremely low incomes. This finding is consistent with that of Aitchinson and Brown, p. 116, who observed the same phenomenon. This fact undoubtedly reflects that the process of income formation, at the very low end of the income scale (below $2000 annually), differs markedly from that of the rest of the income range. However, this underestimation at the very low income end is not a serious problem, since for the most populous of the 56 groups (namely subgroups 1 and 2 of Table 1), the bottom point usually represents at most 6% of the entire population of that group. Furthermore, as the high $R^2$ would indicate, all other points lie quite close to the fitted line for each of the 56 cases.
Table 1

Demographic Groups

For black/white x owner/renter, the 7 demographic groups (hereafter referred to as subgroups) are:

1. male head, family size above 1, age = 0-45
2. male head, family size above 1, age = 45-64
3. male head, family size above 1, age = over 65
4. female head, family size above 1, age = under 65
5. female head, family size above 1, age = over 65
6. single person, age = under 65
7. single person, age = over 65
Graph of the distribution of income for black renters in 1970, aged 45-64.

V9 scale - measured in terms of corresponding cumulative frequency percentages.

Graph of the distribution of income for white owners in 1960, aged 25-45.

V9 scale - measured in terms of corresponding cumulative frequency percentages.
The Parameters of the Log Normal Distributions of Current Income for the Demographic Groups in Pittsburgh in 1960 and 1970

<table>
<thead>
<tr>
<th>Group</th>
<th>( \mu )</th>
<th>( \sigma )</th>
<th>( \hat{\sigma} )</th>
<th>( \hat{\mu} )</th>
<th>( R^2 )</th>
<th>Number of Observations</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-1-1</td>
<td>8.80</td>
<td>0.036</td>
<td>0.477</td>
<td>0.029</td>
<td>0.975</td>
<td>9</td>
</tr>
<tr>
<td>1-1-2</td>
<td>8.89</td>
<td>0.037</td>
<td>0.586</td>
<td>0.034</td>
<td>0.976</td>
<td>9</td>
</tr>
<tr>
<td>1-1-3</td>
<td>8.34</td>
<td>0.015</td>
<td>0.810</td>
<td>0.019</td>
<td>0.996</td>
<td>9</td>
</tr>
<tr>
<td>1-1-4</td>
<td>8.33</td>
<td>0.033</td>
<td>0.704</td>
<td>0.035</td>
<td>0.983</td>
<td>9</td>
</tr>
<tr>
<td>1-1-5</td>
<td>8.32</td>
<td>0.044</td>
<td>0.804</td>
<td>0.053</td>
<td>0.970</td>
<td>9</td>
</tr>
<tr>
<td>1-1-6</td>
<td>7.90</td>
<td>0.059</td>
<td>0.718</td>
<td>0.045</td>
<td>0.973</td>
<td>9</td>
</tr>
<tr>
<td>1-1-7</td>
<td>6.74</td>
<td>0.086</td>
<td>1.174</td>
<td>0.051</td>
<td>0.987</td>
<td>9</td>
</tr>
<tr>
<td>1-2-1</td>
<td>8.53</td>
<td>0.039</td>
<td>0.494</td>
<td>0.039</td>
<td>0.970</td>
<td>7</td>
</tr>
<tr>
<td>1-2-2</td>
<td>8.50</td>
<td>0.035</td>
<td>0.570</td>
<td>0.040</td>
<td>0.976</td>
<td>7</td>
</tr>
<tr>
<td>1-2-3</td>
<td>8.09</td>
<td>0.029</td>
<td>0.739</td>
<td>0.034</td>
<td>0.989</td>
<td>7</td>
</tr>
<tr>
<td>1-2-4</td>
<td>7.90</td>
<td>0.038</td>
<td>0.781</td>
<td>0.036</td>
<td>0.990</td>
<td>7</td>
</tr>
<tr>
<td>1-2-5</td>
<td>7.91</td>
<td>0.058</td>
<td>0.715</td>
<td>0.056</td>
<td>0.970</td>
<td>7</td>
</tr>
<tr>
<td>1-2-6</td>
<td>7.73</td>
<td>0.082</td>
<td>0.625</td>
<td>0.059</td>
<td>0.958</td>
<td>7</td>
</tr>
<tr>
<td>1-2-7</td>
<td>6.92</td>
<td>0.235</td>
<td>0.783</td>
<td>0.114</td>
<td>0.904</td>
<td>7</td>
</tr>
<tr>
<td>2-1-1</td>
<td>8.50</td>
<td>0.027</td>
<td>0.495</td>
<td>0.022</td>
<td>0.986</td>
<td>9</td>
</tr>
<tr>
<td>2-1-2</td>
<td>8.65</td>
<td>0.030</td>
<td>0.576</td>
<td>0.030</td>
<td>0.981</td>
<td>9</td>
</tr>
<tr>
<td>2-1-3</td>
<td>8.14</td>
<td>0.013</td>
<td>0.803</td>
<td>0.014</td>
<td>0.998</td>
<td>9</td>
</tr>
<tr>
<td>2-1-4</td>
<td>7.98</td>
<td>0.038</td>
<td>0.693</td>
<td>0.030</td>
<td>0.987</td>
<td>9</td>
</tr>
<tr>
<td>2-1-5</td>
<td>8.06</td>
<td>0.046</td>
<td>0.832</td>
<td>0.047</td>
<td>0.978</td>
<td>9</td>
</tr>
<tr>
<td>2-1-6</td>
<td>7.97</td>
<td>0.044</td>
<td>0.692</td>
<td>0.034</td>
<td>0.983</td>
<td>9</td>
</tr>
<tr>
<td>2-1-7</td>
<td>6.93</td>
<td>0.058</td>
<td>1.070</td>
<td>0.035</td>
<td>0.993</td>
<td>9</td>
</tr>
<tr>
<td>2-2-1</td>
<td>8.21</td>
<td>0.029</td>
<td>0.513</td>
<td>0.027</td>
<td>0.983</td>
<td>7</td>
</tr>
<tr>
<td>2-2-2</td>
<td>8.24</td>
<td>0.030</td>
<td>0.551</td>
<td>0.030</td>
<td>0.985</td>
<td>7</td>
</tr>
<tr>
<td>2-2-3</td>
<td>7.77</td>
<td>0.032</td>
<td>0.671</td>
<td>0.025</td>
<td>0.993</td>
<td>7</td>
</tr>
<tr>
<td>2-2-4</td>
<td>7.35</td>
<td>0.037</td>
<td>0.768</td>
<td>0.023</td>
<td>0.995</td>
<td>7</td>
</tr>
<tr>
<td>2-2-5</td>
<td>7.23</td>
<td>0.098</td>
<td>1.137</td>
<td>0.084</td>
<td>0.973</td>
<td>7</td>
</tr>
<tr>
<td>2-2-6</td>
<td>7.59</td>
<td>0.089</td>
<td>0.584</td>
<td>0.053</td>
<td>0.961</td>
<td>7</td>
</tr>
<tr>
<td>2-2-7</td>
<td>6.66</td>
<td>0.279</td>
<td>0.860</td>
<td>0.129</td>
<td>0.898</td>
<td>7</td>
</tr>
</tbody>
</table>

1 For notational simplicity, every group is described by three numbers 1-i-K. The last number refers to the subgroups of table 1. The second number indicates race with the value 1 used for whites, and the value 2 used for blacks. The meaning of the first number is given below:

- 1 = owners 1960
- 2 = renters 1960
- 3 = owners 1970
- 4 = renters 1970

So, for example, group 3=1-2 refers to owners 1970, white, subgroup 2.

2 \( \hat{\sigma} \) refers to the estimated standard error of the coefficient.
### Table 2 (cont'd)

<table>
<thead>
<tr>
<th></th>
<th>$u$</th>
<th>$s_u$</th>
<th>$\sigma$</th>
<th>$s_{\sigma}$</th>
<th>$R^2$</th>
<th>Number of Observation</th>
</tr>
</thead>
<tbody>
<tr>
<td>3-1-1</td>
<td>9.29</td>
<td>0.098</td>
<td>0.520</td>
<td>0.058</td>
<td>0.921</td>
<td>9</td>
</tr>
<tr>
<td>3-1-2</td>
<td>9.36</td>
<td>0.087</td>
<td>0.607</td>
<td>0.057</td>
<td>0.942</td>
<td>9</td>
</tr>
<tr>
<td>3-1-3</td>
<td>8.66</td>
<td>0.010</td>
<td>0.795</td>
<td>0.011</td>
<td>0.999</td>
<td>9</td>
</tr>
<tr>
<td>3-1-4</td>
<td>8.72</td>
<td>0.045</td>
<td>0.693</td>
<td>0.042</td>
<td>0.974</td>
<td>9</td>
</tr>
<tr>
<td>3-1-5</td>
<td>8.68</td>
<td>0.046</td>
<td>0.794</td>
<td>0.049</td>
<td>0.974</td>
<td>9</td>
</tr>
<tr>
<td>3-1-6</td>
<td>8.31</td>
<td>0.052</td>
<td>0.768</td>
<td>0.046</td>
<td>0.976</td>
<td>9</td>
</tr>
<tr>
<td>3-1-7</td>
<td>7.32</td>
<td>0.061</td>
<td>1.082</td>
<td>0.042</td>
<td>0.990</td>
<td>9</td>
</tr>
<tr>
<td>3-2-1</td>
<td>9.09</td>
<td>0.079</td>
<td>0.548</td>
<td>0.055</td>
<td>0.935</td>
<td>9</td>
</tr>
<tr>
<td>3-2-2</td>
<td>9.09</td>
<td>0.062</td>
<td>0.595</td>
<td>0.047</td>
<td>0.959</td>
<td>9</td>
</tr>
<tr>
<td>3-2-3</td>
<td>8.34</td>
<td>0.005</td>
<td>0.748</td>
<td>0.004</td>
<td>0.999</td>
<td>9</td>
</tr>
<tr>
<td>3-2-4</td>
<td>8.36</td>
<td>0.032</td>
<td>0.757</td>
<td>0.029</td>
<td>0.990</td>
<td>9</td>
</tr>
<tr>
<td>3-2-5</td>
<td>8.22</td>
<td>0.058</td>
<td>0.796</td>
<td>0.050</td>
<td>0.973</td>
<td>9</td>
</tr>
<tr>
<td>3-2-6</td>
<td>8.11</td>
<td>0.085</td>
<td>0.740</td>
<td>0.064</td>
<td>0.951</td>
<td>9</td>
</tr>
<tr>
<td>3-2-7</td>
<td>7.14</td>
<td>0.228</td>
<td>0.943</td>
<td>0.121</td>
<td>0.896</td>
<td>9</td>
</tr>
<tr>
<td>4-1-1</td>
<td>8.99</td>
<td>0.059</td>
<td>0.546</td>
<td>0.042</td>
<td>0.960</td>
<td>9</td>
</tr>
<tr>
<td>4-1-2</td>
<td>9.10</td>
<td>0.060</td>
<td>0.644</td>
<td>0.048</td>
<td>0.962</td>
<td>9</td>
</tr>
<tr>
<td>4-1-3</td>
<td>8.54</td>
<td>0.016</td>
<td>0.761</td>
<td>0.016</td>
<td>0.997</td>
<td>9</td>
</tr>
<tr>
<td>4-1-4</td>
<td>8.38</td>
<td>0.030</td>
<td>0.709</td>
<td>0.025</td>
<td>0.991</td>
<td>9</td>
</tr>
<tr>
<td>4-1-5</td>
<td>8.48</td>
<td>0.021</td>
<td>0.784</td>
<td>0.021</td>
<td>0.995</td>
<td>9</td>
</tr>
<tr>
<td>4-1-6</td>
<td>8.30</td>
<td>0.077</td>
<td>0.797</td>
<td>0.071</td>
<td>0.947</td>
<td>9</td>
</tr>
<tr>
<td>4-1-7</td>
<td>7.31</td>
<td>0.029</td>
<td>1.127</td>
<td>0.020</td>
<td>0.998</td>
<td>9</td>
</tr>
<tr>
<td>4-2-1</td>
<td>8.81</td>
<td>0.046</td>
<td>0.615</td>
<td>0.038</td>
<td>0.974</td>
<td>9</td>
</tr>
<tr>
<td>4-2-2</td>
<td>8.79</td>
<td>0.046</td>
<td>0.660</td>
<td>0.041</td>
<td>0.973</td>
<td>9</td>
</tr>
<tr>
<td>4-2-3</td>
<td>8.12</td>
<td>0.022</td>
<td>0.749</td>
<td>0.017</td>
<td>0.997</td>
<td>9</td>
</tr>
<tr>
<td>4-2-4</td>
<td>7.92</td>
<td>0.076</td>
<td>0.732</td>
<td>0.050</td>
<td>0.968</td>
<td>9</td>
</tr>
<tr>
<td>4-2-5</td>
<td>8.14</td>
<td>0.027</td>
<td>0.766</td>
<td>0.022</td>
<td>0.994</td>
<td>9</td>
</tr>
<tr>
<td>4-2-6</td>
<td>7.88</td>
<td>0.082</td>
<td>0.798</td>
<td>0.057</td>
<td>0.965</td>
<td>9</td>
</tr>
<tr>
<td>4-2-7</td>
<td>6.72</td>
<td>0.335</td>
<td>1.069</td>
<td>0.166</td>
<td>0.856</td>
<td>9</td>
</tr>
</tbody>
</table>
If $Y$ is distributed log-normal, viz. $N$ is distributed normal), in 1960, then if the fundamental force affecting income distribution over the 60-70 decade was the rate of growth of per-capita income for various groups, then

$$\ln Y_{70} = \ln g \cdot Y_{60}$$

for each person, where $g$ = growth rate of income per capita for some particular group. The implication of eq(11) is that $\ln Y_{1970}$ has a log-normal distribution with the same variance as in 1962, but with a mean that exceeds the 1960 mean of $\ln Y_{60}$ by $\ln g$. The hypothesis that the $\sigma$ of various groups remained unchanged between 60 and 70 is strongly supported by a visual comparison between the $\sigma$ estimates for a group in 1960 and in 1970 in Table 2.

To test the hypothesis about identical $\sigma$'s for various groups for the years 1960 and 1970, constrained regressions were run and $F$ tests were performed. In almost all cases, the hypothesis of equal $\sigma$'s between 1960 and 1970 for each group was strongly accepted at the 95% level. The only cases, where the hypothesis of equal $\sigma$'s for 1960 and 1970 was rejected, were black renters - subgroups 3, 5, and 6 (male head over 65, female over 65, single under 65 respectively). For these groups, forces other than per capita income growth must have affected their income distribution during the 60-70 decade.

In the estimation of these log-normal equations, in certain cases where it appeared

---

1. The constrained regressions which use both 1960 and 1970 data, are of the form

$$\ln Y_{70} = u_{1970} + D_{1970} \cdot u_{1970} + D_{1960} = \sigma \cdot \ln Y_{60},$$

where the notation is the same as in eq (10), but where $u_{1960} = 1960$ constant, $u_{1970} = 1970$ constant, $D_{1960} =$ dummy indicator for 1960, and $D_{1970}$ is dummy indicator for 1970.
obvious that the bottom point was not explained by a log-normal hypothesis, the
bottom point was omitted in the regression equation. As previously mentioned, the
bottom point usually represents a small percent (below 5%) of the population of
that group. In all cases where a bottom point was omitted, the coefficient

\[ (\mu, \sigma) \] of eq. (10) or (10+) remained virtually unchanged; while the estimate

of the standard error of regression usually fell considerably, and the Durbin-Watson

usually improved.

For a few stubborn equations, (mostly those of the elderly), the Durbin-Watson

of the single equations (of Table 2) were so low as to indicate serious autocorrelation,

and hence too low to justify using an F test to test the hypothesis of equal \[ \sigma \] 's in

1960 and 1970. In those few cases, separate autocorrelation corrections were run.

In cases where corrections for autocorrelation were made in either the 1960 or 1970

equation, there was no attempt to impose the constraint that the \[ \sigma \] in 1960 be

identical to the \[ \sigma \] in 1970. The reason for this is that 1) to impose this constraint

would undoubtedly have introduced heteroskedasticity between the 1960 variance of

the stochastic error in eq. (10) (footnote 1, p. 637) and its variance in 1970, and 2)

with only 13 observations, combining the 1960 & 1970 equations into one equation,

and correcting for the resulting heteroskedasticity would have produced negligible,

if any, gains in efficiency. The final results are presented in Table 3. In that

table, a \( \times \) indicates that the estimate of \[ \sigma \] for a particular group is the same as

that reported in the 1950 equation for the group.

---

1. This is especially true for the lower segments of the population, namely subgroups

1 and 2. For elderly subgroups, the bottom point usually represents a sizable proportion of the population.
**TABLE 3**

The Final Estimated Parameters of the Log-Normal Distributions of Current Income for the Demographic Groups in Pittsburgh for 1960-1970 (1)

<table>
<thead>
<tr>
<th>Group</th>
<th>Constant</th>
<th>Stnd Error of Constant</th>
<th>Slope</th>
<th>Stnd Err of Slope</th>
<th>R²</th>
<th>D.W.</th>
<th>No. of Observations</th>
<th>Value of Autocorrelation Coefficient (If Applicable)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-1-1</td>
<td>8.797</td>
<td>.0311</td>
<td>.4584</td>
<td>.0179</td>
<td>0.982</td>
<td>1.205</td>
<td>16</td>
<td></td>
</tr>
<tr>
<td>1-1-2</td>
<td>8.902</td>
<td>.0191</td>
<td>.5277</td>
<td>.0133</td>
<td>0.993</td>
<td>1.720</td>
<td>15</td>
<td></td>
</tr>
<tr>
<td>1-1-3</td>
<td>8.348</td>
<td>.0136</td>
<td>.8002</td>
<td>.0118</td>
<td>0.997</td>
<td>1.039</td>
<td>16</td>
<td></td>
</tr>
<tr>
<td>1-1-4</td>
<td>8.394</td>
<td>.0193</td>
<td>.6305</td>
<td>.0146</td>
<td>0.993</td>
<td>1.024</td>
<td>15</td>
<td></td>
</tr>
<tr>
<td>1-1-5</td>
<td>8.380</td>
<td>.0260</td>
<td>.730</td>
<td>.0310</td>
<td>0.989</td>
<td>1.088</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>1-1-6</td>
<td>7.894</td>
<td>.0427</td>
<td>.7193</td>
<td>.0272</td>
<td>0.981</td>
<td>1.386</td>
<td>17</td>
<td></td>
</tr>
<tr>
<td>1-1-7</td>
<td>6.740</td>
<td>.0860</td>
<td>1.174</td>
<td>.0510</td>
<td>0.987</td>
<td>1.443</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>1-2-1</td>
<td>8.531</td>
<td>.0304</td>
<td>.4786</td>
<td>.0187</td>
<td>0.986</td>
<td>1.544</td>
<td>14</td>
<td></td>
</tr>
<tr>
<td>1-2-2</td>
<td>8.495</td>
<td>.0279</td>
<td>.5380</td>
<td>.0194</td>
<td>0.988</td>
<td>1.347</td>
<td>14</td>
<td></td>
</tr>
<tr>
<td>1-2-3</td>
<td>8.088</td>
<td>.0156</td>
<td>.7447</td>
<td>.0111</td>
<td>0.987</td>
<td>1.202</td>
<td>16</td>
<td></td>
</tr>
<tr>
<td>1-2-4</td>
<td>7.839</td>
<td>.0255</td>
<td>.7392</td>
<td>.0178</td>
<td>0.994</td>
<td>1.331</td>
<td>15</td>
<td></td>
</tr>
<tr>
<td>1-2-5</td>
<td>7.892</td>
<td>.0436</td>
<td>.7342</td>
<td>.0313</td>
<td>0.982</td>
<td>1.319</td>
<td>15</td>
<td></td>
</tr>
<tr>
<td>1-2-6</td>
<td>7.643</td>
<td>.0820</td>
<td>.7014</td>
<td>.0459</td>
<td>0.950</td>
<td>1.541</td>
<td>16</td>
<td></td>
</tr>
<tr>
<td>1-2-7</td>
<td>6.996</td>
<td>.0869</td>
<td>.7153</td>
<td>.0440</td>
<td>0.969</td>
<td>2.606</td>
<td>12</td>
<td></td>
</tr>
</tbody>
</table>

Table continues
## TABLE 3

<table>
<thead>
<tr>
<th>Group</th>
<th>Constant</th>
<th>Std Error of Constant</th>
<th>Slope $a$</th>
<th>Std Error of Slope $a$</th>
<th>$R^2$</th>
<th>D.W.</th>
<th>No. of Observation</th>
<th>Value of Autocorrelation Coefficient (If Applicable)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2-1-1</td>
<td>8.506</td>
<td>.0229</td>
<td>.4868</td>
<td>.0140</td>
<td>0.990</td>
<td>1.317</td>
<td>16</td>
<td></td>
</tr>
<tr>
<td>2-1-2</td>
<td>8.655</td>
<td>.0259</td>
<td>.5709</td>
<td>.0185</td>
<td>0.987</td>
<td>1.097</td>
<td>16</td>
<td></td>
</tr>
<tr>
<td>2-1-3</td>
<td>8.150</td>
<td>.0131</td>
<td>.7924</td>
<td>.0105</td>
<td>0.998</td>
<td>1.436</td>
<td>17</td>
<td></td>
</tr>
<tr>
<td>2-1-4</td>
<td>7.952</td>
<td>.0608</td>
<td>.6962</td>
<td>.0398</td>
<td>0.995</td>
<td>1.337</td>
<td>9</td>
<td>$\rho = .5908$</td>
</tr>
<tr>
<td>2-1-5</td>
<td>8.020</td>
<td>.0744</td>
<td>.8335</td>
<td>.0611</td>
<td>0.991</td>
<td>1.356</td>
<td>9</td>
<td>$\rho = .5883$</td>
</tr>
<tr>
<td>2-1-6</td>
<td>7.953</td>
<td>.0430</td>
<td>.7118</td>
<td>.0281</td>
<td>0.980</td>
<td>1.552</td>
<td>17</td>
<td></td>
</tr>
<tr>
<td>2-1-7</td>
<td>6.876</td>
<td>.0359</td>
<td>1.1025</td>
<td>.0198</td>
<td>0.995</td>
<td>0.987</td>
<td>18</td>
<td></td>
</tr>
<tr>
<td>2-2-1</td>
<td>8.190</td>
<td>.0365</td>
<td>.5518</td>
<td>.0225</td>
<td>0.983</td>
<td>1.760</td>
<td>15</td>
<td></td>
</tr>
<tr>
<td>2-2-2</td>
<td>8.225</td>
<td>.0308</td>
<td>.5900</td>
<td>.0205</td>
<td>0.988</td>
<td>1.512</td>
<td>15</td>
<td></td>
</tr>
<tr>
<td>2-2-3</td>
<td>7.77</td>
<td>.0320</td>
<td>.671</td>
<td>.025</td>
<td>0.993</td>
<td>2.123</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>2-2-4</td>
<td>7.386</td>
<td>.0633</td>
<td>.7412</td>
<td>.0325</td>
<td>0.977</td>
<td>1.303</td>
<td>16</td>
<td></td>
</tr>
<tr>
<td>2-2-5</td>
<td>7.23</td>
<td>.0980</td>
<td>1.137</td>
<td>.084</td>
<td>0.973</td>
<td>2.354</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>2-2-6</td>
<td>7.59</td>
<td>.0890</td>
<td>.584</td>
<td>.053</td>
<td>0.961</td>
<td>1.651</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>2-2-7</td>
<td>6.676</td>
<td>.1183</td>
<td>.8539</td>
<td>.0518</td>
<td>0.958</td>
<td>1.323</td>
<td>15</td>
<td></td>
</tr>
</tbody>
</table>

Table continues
### TABLE 3

The Final Estimated Parameters of the Log-Normal Distributions of Current Income for the Demographic Groups in Pittsburgh for 1960-1970(1)

<table>
<thead>
<tr>
<th>Group</th>
<th>Constant</th>
<th>Std Error of Constant</th>
<th>Std Error of Slope</th>
<th>$R^2$</th>
<th>D.W.</th>
<th>No. of Observations</th>
<th>Value of Autocorrelation Coefficient (If Applicable)</th>
</tr>
</thead>
<tbody>
<tr>
<td>3-1-1</td>
<td>9.333</td>
<td>0.0339</td>
<td>*</td>
<td>*</td>
<td>0.982</td>
<td>1.205</td>
<td>16</td>
</tr>
<tr>
<td>3-1-2</td>
<td>9.380</td>
<td>0.0202</td>
<td>*</td>
<td>*</td>
<td>0.993</td>
<td>1.720</td>
<td>15</td>
</tr>
<tr>
<td>3-1-3</td>
<td>8.657</td>
<td>0.0126</td>
<td>*</td>
<td>*</td>
<td>0.997</td>
<td>1.039</td>
<td>16</td>
</tr>
<tr>
<td>3-1-4</td>
<td>8.796</td>
<td>0.0179</td>
<td>*</td>
<td>*</td>
<td>0.993</td>
<td>1.024</td>
<td>15</td>
</tr>
<tr>
<td>3-1-5</td>
<td>8.72</td>
<td>0.0390</td>
<td>0.746</td>
<td>0.043</td>
<td>0.981</td>
<td>0.586</td>
<td>8</td>
</tr>
<tr>
<td>3-1-6</td>
<td>8.376</td>
<td>0.0395</td>
<td>*</td>
<td>*</td>
<td>0.981</td>
<td>1.386</td>
<td>17</td>
</tr>
<tr>
<td>3-1-7</td>
<td>7.363</td>
<td>0.1020</td>
<td>1.083</td>
<td>0.0583</td>
<td>0.995</td>
<td>0.958</td>
<td>9</td>
</tr>
<tr>
<td>3-2-1</td>
<td>9.153</td>
<td>0.0286</td>
<td>*</td>
<td>*</td>
<td>0.986</td>
<td>1.544</td>
<td>14</td>
</tr>
<tr>
<td>3-2-2</td>
<td>9.136</td>
<td>0.0263</td>
<td>*</td>
<td>*</td>
<td>0.988</td>
<td>1.347</td>
<td>14</td>
</tr>
<tr>
<td>3-2-3</td>
<td>8.342</td>
<td>0.0144</td>
<td>*</td>
<td>*</td>
<td>0.997</td>
<td>1.202</td>
<td>16</td>
</tr>
<tr>
<td>3-2-4</td>
<td>8.393</td>
<td>0.0231</td>
<td>*</td>
<td>*</td>
<td>0.994</td>
<td>1.331</td>
<td>15</td>
</tr>
<tr>
<td>3-2-5</td>
<td>8.297</td>
<td>0.0431</td>
<td>*</td>
<td>*</td>
<td>0.982</td>
<td>1.319</td>
<td>15</td>
</tr>
<tr>
<td>3-2-6</td>
<td>8.140</td>
<td>0.0688</td>
<td>*</td>
<td>*</td>
<td>0.950</td>
<td>1.541</td>
<td>16</td>
</tr>
<tr>
<td>3-2-7</td>
<td>7.391</td>
<td>0.0748</td>
<td>*</td>
<td>*</td>
<td>0.969</td>
<td>2.606</td>
<td>12</td>
</tr>
</tbody>
</table>

$\rho = 0.7230$

Table continues
TABLE 3

The Final Estimated Parameters of the Log-Normal Distributions of Current Income for the Demographic Groups in Pittsburgh for 1960-1970

<table>
<thead>
<tr>
<th>Group</th>
<th>Constant</th>
<th>Std Error of Constant</th>
<th>Slope</th>
<th>Std Err of Slope</th>
<th>R²</th>
<th>D.W.</th>
<th>No. of Observations</th>
<th>Value of Autocorrelation Coefficient (If Applicable)</th>
</tr>
</thead>
<tbody>
<tr>
<td>4-1-1</td>
<td>9.043</td>
<td>.0256</td>
<td>*</td>
<td>*</td>
<td>0.990</td>
<td>1.317</td>
<td>16</td>
<td></td>
</tr>
<tr>
<td>4-1-2</td>
<td>9.140</td>
<td>.0295</td>
<td>*</td>
<td>*</td>
<td>0.987</td>
<td>1.097</td>
<td>16</td>
<td></td>
</tr>
<tr>
<td>4-1-3</td>
<td>8.513</td>
<td>.0131</td>
<td>*</td>
<td>*</td>
<td>0.998</td>
<td>1.436</td>
<td>17</td>
<td></td>
</tr>
<tr>
<td>4-1-4</td>
<td>8.41</td>
<td>.025</td>
<td>.680</td>
<td>.020</td>
<td>0.995</td>
<td>1.009</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>4-1-5</td>
<td>8.453</td>
<td>.0763</td>
<td>.7611</td>
<td>.028</td>
<td>0.998</td>
<td>0.944</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>4-1-6</td>
<td>8.407</td>
<td>.0406</td>
<td>*</td>
<td>*</td>
<td>0.980</td>
<td>1.552</td>
<td>17</td>
<td></td>
</tr>
<tr>
<td>4-1-7</td>
<td>7.336</td>
<td>.0308</td>
<td>*</td>
<td>*</td>
<td>0.995</td>
<td>0.987</td>
<td>18</td>
<td></td>
</tr>
<tr>
<td>4-2-1</td>
<td>8.848</td>
<td>.0327</td>
<td>*</td>
<td>*</td>
<td>0.983</td>
<td>1.760</td>
<td>15</td>
<td></td>
</tr>
<tr>
<td>4-2-2</td>
<td>8.829</td>
<td>.0280</td>
<td>*</td>
<td>*</td>
<td>0.988</td>
<td>1.512</td>
<td>15</td>
<td></td>
</tr>
<tr>
<td>4-2-3</td>
<td>8.12</td>
<td>.022</td>
<td>.749</td>
<td>.017</td>
<td>0.997</td>
<td>1.764</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>4-2-4</td>
<td>7.912</td>
<td>.0532</td>
<td>*</td>
<td>*</td>
<td>0.977</td>
<td>1.303</td>
<td>16</td>
<td></td>
</tr>
<tr>
<td>4-2-5</td>
<td>8.14</td>
<td>.027</td>
<td>.766</td>
<td>.022</td>
<td>0.994</td>
<td>2.284</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>4-2-6</td>
<td>7.88</td>
<td>.082</td>
<td>.798</td>
<td>.057</td>
<td>0.965</td>
<td>1.514</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>4-2-7</td>
<td>6.987</td>
<td>.1079</td>
<td>*</td>
<td>*</td>
<td>0.958</td>
<td>1.323</td>
<td>15</td>
<td></td>
</tr>
</tbody>
</table>

(1) For an explanation of the notation used, see footnotes 1 and 2 of Table 2.

A "*" in a group I-J-K indicates that the relevant values are the same as those for the group in 1960 (i.e. group (1-2)-J-K)
Using the specific values of Table 3, the validity of the argument regarding the growth of income can be examined more closely. It was argued above, that the difference in the constant term in Table 3 for any group between 1960 and 1970 should reflect the growth of per capita nominal income between 1960 and 1970. Table 4 presents the average difference between the 1960 and 1970 estimated constant term for the four main groups (i.e., owner/renter x black/white). The second column of Table 4 presents similar results, except now a weighted average is used to compute the average difference, with the weights being proportional to the population of each subgroup. Each table indicates that the income growth of blacks, especially black owners was substantially higher than that of whites. Examination of individual subgroups shows that the highest rates of growth of income were achieved by subgroups 1 and 2, with the blacks in these subgroups having considerably higher growth rates of income than their white counterparts.

The actual growth in per capita nominal income between 1960 and 1970 was \(1905^{3130}\), or about 1.64. The ln 64 \(\approx .495\), which accords almost identically with the implications of table 4. Such a result lends strong support to the usefulness of the log-normal distribution as an approximation to the distribution of income.

In addition to their good fit, the equations of Table 3 perform quite well in another, though related, sense. Each equation is able to predict the median income extremely well, usually within an error of 0-3%. This property is quite desirable.


2. If \(Y\) is ln normally distributed with parameters \((\mu, \sigma)\) then \(e^\mu = \text{median of } Y\). Hence a 0-5\% error in predicting the median, means approximately a 0-.05 error in predicting the mean of \(\ln Y\).

3. This finding is in contradiction to Metcalf's finding that median income is underpredicted when a log-normal distribution is used to approximate the income distribution. Metcalf uses the method of fractiles for his estimation, a method whose disadvantages were discussed on p. 13. See Metcalf, Charles, An Econometric Model of the Income Distribution, Markham Publishing Co., Chicago; p. 116.
### Table 4

<table>
<thead>
<tr>
<th></th>
<th>whites</th>
<th></th>
<th>weighted average of individual differences</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>owners</td>
<td>.453</td>
<td>.478</td>
</tr>
<tr>
<td></td>
<td>renters</td>
<td>.456</td>
<td>.487</td>
</tr>
<tr>
<td>blacks</td>
<td>owners</td>
<td>.481</td>
<td>.544</td>
</tr>
<tr>
<td></td>
<td>renters</td>
<td>.521</td>
<td>.517</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>.478</td>
<td>.484</td>
</tr>
</tbody>
</table>
for purposes of our research, since we will be summarizing the income distribution
by dividing it up into a small number of fractiles. Hence, it is quite nice to know
that we are able, at least, to divide the distribution of income in half quite accurately.

Before discussing the results in Table 3 in more detail, it will be useful to digress
briefly to explain the interpretation of the parameters \( \mu \) and \( \sigma \). If \( Y \) is log
normally distributed with parameters \((\mu, \sigma)\), then \( E(\ln Y) = \mu \), and \( \text{var}(\ln Y) = \sigma^2 \).
The parameter \( \mu \) is directly related to the median of \( Y \), by the relation \( e^\mu = \text{Ymedian} \).
The parameter \( \sigma^2 \) essentially provides a measure of the variability in the percentage
deviations of \( Y \) across the population. The variance of \( Y \) (not \( \ln Y \)), is related
to both \( \mu \) and \( \sigma^2 \), by the formula \( \text{var} Y = e^{2\mu} - e^{\sigma^2} (e^{\sigma^2} - 1) \). \(^1\) Let us now
consider the results of Table 3 in more detail.

Basically, the results of Table 2 and Tables 3 and 4 agree quite well with
expectations about relative incomes of various groups. White owners have the
highest median income, followed by white renters, black owners, and black renters.
The difference in median income between owners and renters is much greater in
percentage terms for blacks, than for whites. As previously mentioned, the rate
of income growth during the decade for blacks was higher than for whites. This
higher growth rate, no doubt, reflects the influence of the civil rights movements
during the 60's.

\(^1\) Aitchinson and Brown - op. cit., Metcalf - op. cit. Another interpretation of \( \sigma \),
not discussed in this paper, is that \( \sigma \) is monotonically related to Lorenz's measure of
inequality (see Aitchinson and Brown, p. 113).
As regards the estimates of $\sigma$ the patterns are less clear cut than the estimates of $u$. One pattern that did emerge was that the $\sigma$ estimate seemed to increase from subgroups 1 and 3 for both blacks and whites. This would indicate that for male headed households, the variability in the percentage deviations of $Y$ (i.e., the variability of $\ln Y$) increases with age.

There seems to be little, if any, systematic differences between the "$\sigma$--" estimates for blacks and whites, except in the case of subgroups 4 and 7 (female headed households under 65, and elderly single people). For these subgroups, the "$\sigma$--" estimates for whites considerably exceed those for blacks; indicating that the variability in percentage deviations of $Y$ for whites exceeds that for blacks for those particular subgroups. As mentioned above, the variance of $Y_c$ for any group, depends on both $u$ and $\sigma^2$, and is given by the formula $\text{var } Y = e^{2u} \cdot e^{\sigma^2 - 1}$. Since, for each subgroup, $u$ for whites tends to be higher than $u$ for blacks, while $\sigma$ for whites and blacks show no systematic difference, one can conclude from the results of Table 3, that the variance of income is considerably higher for white subgroups than for the corresponding black subgroups; while the variance of the percentage changes in income appears similar for corresponding white and black subgroups. Once again, these results accord well with intuition.

A comparison of the "$\sigma$--" between owners and renters, for corresponding subgroups also reveals little systematic differences between the estimates except for black subgroups 5, 6, and 7 (female headed, over 65, single under 65, and single over 65). Indeed, one might well be led to believe that the $\sigma$ for owners and renters and for blacks and whites were equal, for corresponding subgroups (though, of course, the
variance of income is higher for owners and whites than for renters and blacks.

In this study, the constraint of equal $\sigma^2$'s for owners and renters, and for blacks and whites in corresponding subgroups, was not imposed for two reasons. First, the justification for imposing such constraints seemed weak, especially in comparison to the rather strong theoretical justification of p. 6-27 that argued for the constraint of identical $\sigma^2$'s for each subgroup in 1960 and 1970. Secondly, the data set was not considered sufficiently rich to allow a statistical test of these hypotheses on $\sigma^2$, that would be powerful enough to overcome this relative lack of solid theoretical justification for imposing such constraints. However, the hypothesis of equal $\sigma^2$'s for owners and renters, and for blacks and whites in corresponding subgroups, does seem interesting, and deserving of future empirical investigation with other bodies of data.

In summary, the log-normal distribution appears to provide an excellent approximation to the distribution of actual income. One seems well justified in using it in the derivation of the distribution of permanent income. The estimation of the parameters of the log-normal income distribution for each group in Table 3 is the first half of the solution to the problem of estimating the parameters of the log-normal distribution of permanent income for these same groups. Let us now turn to the second half of the solution, namely, the estimation of the log-normal distribution of $\ln Y_e/Y_p$ for these same groups.
2.4 The Log-Normal Distribution of "Transitory" Income, $Y_c/Y_p$:
Dependence on Demographic and Economic Variables

Proceeding in the effort to estimate the distribution of permanent income, the
next step is to closely examine the distribution of the "transitory" component of income,
namely $Y_c/Y_p$. Once the relationship between permanent income and past years'
Income has been specified, a cross-sectional time series of data can be used to obtain
a cross-sectional data series on $(Y_c/Y_p)$ for each individual $i$ in the cross section.
Using this series on $(Y_c/Y_p)^i$, the relationship between the assumed log-normal
distribution of this variable and certain demographic and economic characteristics of
the population can then be examined. If a representative sample of individuals is used
for this cross-sectional study, then it is reasonable to assume that the relations that
are present in this cross-sectional sample will carry over to Pittsburgh and can be used
to "predict" the log-normal distribution of $Y_c/Y_p$ for Pittsburgh in any year, for any
particular demographic group.

To say that the log-normal distribution of $Y_c/Y_p$ for group $j$ will depend on the
demographic $(x_j)$ and economics $(e_j)$ characteristics of that group simply means that
$(u_j, \sigma_j^2)$ of the normal distribution of $\ln Y_c/Y_p$ will depend on $x_j$ and $e_j$ where
$u_j = E \ln (Y_c/Y_p)$, and $\sigma_j^2 = \text{var} (\ln Y_c/Y_p)$. Since, the demographic characteristics
with which this study is most concerned are the ones listed in Table I of the previous
section (there are 28 groups in all), the logical characteristics to include in $x$ are these.
However, it is not necessarily appropriate (and this will be verified later by the results
to be reported) to adopt the assumption of an approximate log-normal distribution of
$Y_c/Y_p$ for each of the groups listed in Table 1. In particular, it may be wise to use
the finer age breakdown of 0-45, 45-65 for subgroups 4 and 6 (viz., female headed
under 65, single under 65). By estimating an appropriate log normal distribution for
this finer age breakdown, and then combining these results to obtain an approximate
log-normal distribution for subgroups 4 and 6, it will be possible to detect important
forces that affect different age segments of subgroups 4 and 6, but which could not have
been discovered, if the more aggregate 0-65 age category had been used.

There are at least 2 ways of proceeding to estimate the relation of $\nu$ and $\sigma^{-2}$
on $x$ and $e$, depending on how one wishes to parametrize the problem. First,
suppose that the population is stratified into the 28 demographic groups of Table 1.
Then for each group $j$, we postulate that

\begin{equation}
\mu_t^{(j)} = f_t(e_t), \quad \sigma_t^{(j)^2} = h_t(e_t),
\end{equation}

where the subscript $t$ refers to time. Alternatively, we could impose the same
parametrization on the distributions for each group $j$, and then write that

\begin{equation}
\mu_t^{(j)} = f(x_t, e_t), \quad \sigma_t^{(j)^2} = h(x_t, e_t)
\end{equation}

The first method could run into the problem that the number of people in
certain cells will be very small. Also, if the parameterization (13) is approximately
correct, then estimating a model based on the specification in eq. (13) will
sharpen our accuracy when we predict the desired parameters, as functions of $x$ and $e$,
for the distribution of $Y_c/Y_p$ in Pittsburgh, for various groups. For these reasons,
eq. (13) will be used as the basis of the estimation procedure.
The parameters of $\beta$ and $h$ in eq. (13) can be estimated by running the regressions

$$\ln \left( \frac{Y_c}{Y_p} \right)^i = f(x_i, e_i) + n_1, \text{ and}$$

$$\ln \left( \frac{Y_c}{Y_p} \right)^i = E \ln \left( \frac{Y_c}{Y_p} \right)^i + h(x_i, e_i) + n_2 \text{ where } n_1 \text{ and } n_2 \text{ are error terms and where individual } i \text{ belongs to group } j.$$  

Note that $E\ln(Y_c/Y_p)^i = u_i = f(x_i, e_i)$, and $E(\ln(Y_c/Y_p)^i - E\ln(Y_c/Y_p)^i)^2 = \sigma^2_i = h(x_i, e_i)$, for individual $i$ in group $j$.

One expects that economic conditions could exert significant influence on the mean of $\ln(Y_c/Y_p)^i$. On the other hand, it is difficult to see why demographic characteristics would ever influence $u_i$, except through some interaction with economic conditions. In fact, the assumption in $\Lambda$, is that in "normal" economic conditions, $E(Y_c/Y_p) = 1$ for all groups. Since $\ln \left( \frac{Y_c}{Y_p} \right) \propto Y_c/Y_p - 1$, the range of $Y_c/Y_p$ that will most commonly be observed, it follows that in "normal" economic conditions, $E(\ln(Y_c/Y_p)^i) = 0$. Based on the foregoing reasoning, eq. (14a) can be rewritten as

$$\ln \left( \frac{Y_c}{Y_p} \right)^i = f(x_i, e_i) + e_i,$$

where $e_i$ is an economic-demographic interaction term and $f(\cdot) = 0$ in "normal" economic conditions. On the other hand, regarding the variance equation, eq. (14b) it is quite reasonable to expect that both $x_i$ and $e_i$ influence the variance.

---

(1) Eq. (14b) implies that there is heteroskedasticity in the $e$ of eq. (14a). Once eq. (14b) is estimated, it can be used in eq. (14a) to correct for the heteroskedasticity.
Eqs. (14b) and (15) will be the basic equations for estimating \( f \) and \( h \).

**Data and Variables Used in Cross-Section Study**

The cross sectional data set that is used to estimate the distribution of \( \ln Y_c/Y_p \) is the five year panel study, known as the Michigan Data Survey. There are 5060 families on this data set, with each family providing information for five years (1968-1972). However, not all families are suitable for our purposes. First, "split-offs", i.e., members of families previously interviewed, who now form their own family (e.g., a son marries), are included in these 5060 families, with the data for the years when he was with the original family being identical to that of the original family. Obviously, the income of a son's father should not appear in the son's calculation of his permanent income, once the son marries. For this reason, only families which remained unchanged, or had a change in family composition other than head and wife were included.

Another major category of exclusion was rural and small city families. Because our study is concerned solely with the distribution of transitory and permanent income in Pittsburgh, and because it is likely that the pattern of income variability for rural and small city areas is much different from that of large urban areas, families which lived in areas where the largest city had less than 100,000 people were excluded from our data set. Also, families who lived in foreign nations were omitted.

---

(1) Income data is provided for 1967-1971, while economic data is provided for 1968-1972.
A final minor category of exclusion was families for whom data was available for certain questions. After excluding families which did not meet the above requirements, there remained 1399 observations on individual families for each of the five years of the data set.

As mentioned above, the type of demographic variables to include in $x$ are those of Table 1. Dummy variables are used to indicate 1) age category, 2) sex of head, 3) race, 4) owner-renter status and 5) size of household (over 1 or not). The race variable is measurable as black/white. Other minorities (i.e. Spanish) were included in blacks (this was an insignificant percentage of the non-white population). The economic variable that is used to indicate the level of economic activity in the region, is the unemployment rate of the county. 1 Actually, the deviation from the previous three-year unemployment rate is used. The assumption is that it is deviation from the "normal" economic conditions that influences the distribution of $Y_c/Y_p$ for any particular demographic group and time period $t$. Recall that in "normal" economic conditions the mean of $\ln Y_c/Y_p$ approximately equals 0.

A simple linear specification for eqs. (14b) and (15) in the variables $x$ and $e$ misses what are probably important interaction effects between the demographic and economic variables. The idea is that a high level of unemployment may affect the young more than the middle aged, and blacks more than whites. For this reason, several economic interaction variables were used. Interactions between race, age, sex, size of family, and economic conditions were all attempted in estimating both eqs. (14b) and (15).

1 The unemployment rate is reported as falling into one of the following five categories, 0-2%, 2-4%, 4-6%, 6-10%, and over 10%. The following values were used to represent these categories: 2, 3, 5, 7.5, and 11%. Practically, no observations fell in the first and last categories.
The year that was chosen to estimate the equations was 1970. It appeared that the year 1971 (the last year for which income data is available on the Michigan tapes) was not a good year to use because of the wage-price freeze that began in August 1971. It was felt that the association between economic conditions and \( Y_o/Y_p \) could be distorted because of the freeze. (1) By choosing 1970, the data allows us to form a four year (1967-1970) weighted average of past income for an estimate of \( Y_p \), and a three year (1968-1970) average of unemployment as an estimate of "normal" economic conditions.

**Relation of Permanent Income to Past Years' Income**

The permanent income of each individual will be estimated by a weighted average of his previous four years' income:

\[
(16) \quad Y_p^t = \sum_{z=0}^{3} C^{-z} Y^t_{t-z}.
\]

A declining weighting scheme seems most reasonable, while an equal weighting scheme represents a straight arithmetic average.

Before a weighted average of previous year's income can be formed, each yearly income figure should be corrected for general inflation and for regional

---

(1) Preliminary experiments using 1971 data did in fact show rather perverse patterns of signs on coefficients, with unemployment being positively correlated with positive transitory income. Whether the influence of the wage-price freeze was sufficient to explain this relation was not investigated.
price differences. The annual consumer price index was used to correct for the yearly effects of inflation. Unfortunately, regional price indices for the county of residence are available only for 1970. If we assume that 1) people don't move, and 2) regional price differences remain constant at their 1970 relative values, then we can use the 1970 regional price index to correct yearly income for regional price differences. However, since we are always concerned with a ratio of two incomes, namely $Y_c/Y_p$, then under the two aforementioned assumptions, the regional price correction cancels out.

The two assumptions mentioned above which justify ignoring regional price variation are much stronger than required. If $\ln Y_c/Y_p$ is used in eq. (15), with incomes uncorrected for regional price variation, then to avoid possible bias in our estimates of $f(\cdot)$, because of regional price differentials, it is required that 1) people move more randomly among counties - i.e. there is no tendency for people to move from say low to high-price index counties, and 2) unemployment levels and regional price inflation are uncorrelated across regions at one point in time. Both conditions seem quite reasonable. People do not look at the general price level of a region before they migrate, but rather at the real wage that they can expect to earn for several years in that region. There is no reason to expect that changes in the long run real wage of a region are strongly associated with the rate of price inflation of that region. The second condition is also reasonable. There seems to be little correlation between regional price inflation and unemployment.

(1) This implies that inflation is constant across all regions.
A graph, in Appendix 3, of these 2 series for the 25 largest cities, illustrates this point quite clearly for the year 1970. The conclusion is that that inability to correct for regional differences in the rate of inflation will not cause serious bias in our estimation of f(•) and h(•) of eqs. (15) and (14b).

Using the inflation corrections described in the previous paragraph, two different estimates for $Y^i_p$ were used. The first was

$$Y^i_p = 0.412 Y^i(0) + 0.275 Y^i(-1) + 0.190 Y^i(-2) + 0.123 Y^i(-3),$$

while the second was a straight four year average. The weights for the first estimate are, except for normalization, the first four weights used by Friedman in "A Theory of the Consumption Function". The first estimate of $Y^i_p$ with declining weights seems more reasonable than the second. Basically, the second estimate of $Y^i_p$ is used solely to investigate the sensitivity of the results to different weighting schemes.

Implication of the Sum of the Weights Used to Estimate $Y^i_p$ in Eq. (16)

In the specification for $Y^i_p$ in eq. (16), not only the relative weights, but also the sum of the weights are quite important. It is not necessarily true that the $c$'s should add up to unity. For example, if the income of the members of some group is growing quite rapidly, then it is perfectly reasonable to expect the sum of the weights to exceed unity. In particular, different rates of income growth are likely for blacks, whites, and for the young and old.

(1) Laumas, Journal of Political Economy 1971 also makes this point that the sum of the weights has implications for the growth of income for a particular group.
A priori, one may have only vague quantitative ideas about how the sum of the weights should vary by group, and may be hesitant to use these ideas in the empirical study. Fortunately, there is no need to specify beforehand what the sum of the c's should be for each group; the data can tell us. To understand this point, consider once again eq. (15),

\[ \ln(Y_c/Y_p)^\dagger = f(e_i, t) + n_1 \]

Now, suppose \( \hat{Y}_p^\dagger \) is estimated from eq. (16) as \( \check{Y}_p^\dagger \). In forming this estimate, \( \hat{Y}_p^\dagger \), the correct relative weights are used, but the sum of the weights equals one.

Suppose that this individual belongs to a group whose income growth is so rapid that the weights should sum to

\( g(>1) \). Then \( \hat{Y}_p^\dagger \propto g \cdot \hat{Y}_p^\dagger \), and it immediately follows that

\[ \ln(Y_c/Y_p)^\dagger = \ln(Y_c/Y_p)^\dagger - \ln(g) \]

If \( \hat{Y}_p^\dagger \) is used in eq. (15) instead of \( \check{Y}_p^\dagger \), then the equation that is actually being estimated is

\[ \ln(Y_c/Y_p)^\dagger = \ln g + f(e_i, t) + n_1 \]

But eq. (17) implies that \( \ln g \) will appear as the value of a dummy variable for group \( i \). Hence, if eq. (15) is estimated with dummy variables for each demographic group, then it is possible to capture the differential growth rates of income among the various groups. Such reasoning suggests that the equation that should be used to estimate the dependence of the mean \( u \) on socio-economic characteristics is

\[ \ln(Y_c/Y_p)^\dagger = a^\dagger \cdot D + f(e_i, t) + n_1 \]

where individual \( i \) belongs to group \( j \), \( n_1 \) is the stochastic error, and \( D \) is a 0-1
dummy vector for demographic tracts. The relative magnitudes of the coefficients of \( b^i \) will provide information on the relative growth rates of income for the various demographic groups.

**Empirical Results**

Before turning to a discussion of the specific results, it will be useful to briefly consider the type of results that can be expected. Since, by definition, "transitory" income is a random variable whose randomness arises from the interaction of innumerable unobservable forces, it is not expected that \( \ln Y_c/Y_p \) can be explained very well for each individual with, basically, one economic indicator. However, it is crucial to recognize that the objective is not necessarily to explain \( \ln Y_c/Y_p \) very closely. Instead, the objective is to accurately determine the mean (i.e. the expected value, or more precisely the \( f(\cdot) \) function of eq. (15)) of this random variable.

The variability of \( \ln Y_c/Y_p \) around \( f(\cdot) \) has absolutely no importance. In other words, the statistic \( R^2 \) will be completely irrelevant in judging the results of the equations. The statistics that will be meaningful are the magnitude and statistical significance of the estimated coefficients of \( f(\cdot) \). The \( F \) statistic\(^{(1)} \) statistic will reveal if the equation is explaining anything at all. An identical argument applies to the estimates of \( h(\cdot) \) in eq. (14b).

The following highly simplified model will provide a very rough idea about the expected size of the coefficients of the economic variables in the equation used to

\(\text{(1) The } F\text{-statistic indicates whether the equation is explaining a significant amount more than a pure constant term would.}\)
explain $E(\ln Y_c/Y_p)$. Suppose, that unemployment is the sole source of income fluctuation, let $u$ represent the deviation of unemployment from its normal level. Then, $Y_c = Y_p(1 - u)$, or $\ln Y_c/Y_p = \ln(1 - u) \approx -u$. In this simple example, a 1% change in unemployment will cause a .01 change in $\ln Y_c/Y_p$. Based on this simple model, we expect rather small coefficients (.0-.10) on the economic variables for most demographic groups in eq. (18). In other words, in estimating eq. (18), we hope to discern, through a complicated net of unobservable forces, a rather small, but positive, association between the mean of $\ln (Y_c/Y_p)$ and economic conditions for certain demographic groups. Let us now consider the specific results of estimating the mean eq. (18).

**Results From Estimating the "Mean" Equation (18)**

After experimenting with several interaction terms, the form of eq. (18) that was chosen to explain $\ln(Y_c/Y_p)$ was

\[
\begin{align*}
\ln Y_c/Y_p & = 0.0589 A + 0.0324 RA6 + 0.0369 ERA6 + \\
& \quad + 0.02473 ENA4 + 0.0515 ENA6 - 0.0301, \quad (1) \\
& \quad + 1.982 (2.629) (3.928) \\
R^2 & = 0.04 \tag{2} \\
F & = 11.53 \\
\text{Degrees of Freedom} & = 1393, \text{ where}
\end{align*}
\]

(1) $t$-ratios are in parenthesis.

(2) As mentioned on previous page, $R^2$ is meaningless for purposes of judging eq. (19). It is reported only for completeness.
A = 1 if the head of household is 0-45 years, but not a white female, 0 otherwise,
RA6 = 1 if the head of household is black, 0-45 years, 0 otherwise.
ERA6 = RA6 * EC,
ENA4 = NA4 * EC,
ENA6 = NA6 * EC,
EC = \frac{UN68 + UN69 + UN70}{3} = UN70

UNt = unemployment rate in year 1960 + t,
NA4 = single, 45-65, and
NA6 = single, 0-45.

There are several interesting features to note about eq. (19). First, economic conditions seem to matter for only three demographic groups. (Blacks aged 0-45, single aged 0-45, singles aged 45-65). Attempts to include economic variables for other demographic groups always resulted in usually positive but relatively small and highly insignificant (t ratios below 1) coefficients for those variables. In light of the discussion of the previous page, these results are not terribly surprising.

The coefficients of the economic variables that do matter in eq. (19) appear to be in the right ballpark, based on previous arguments. It is quite reasonable to believe that the "transitory" income (as measured by \( Y_c/Y_p \)) of young blacks and single people is sensitive to economic conditions. As might have been expected the coefficient of ENA6 exceeds that of ENA4, reflecting the fact that younger single people are more sensitive to economic conditions than the older members of their group. One possible reason for the importance of economic conditions for single people is the greater mobility and hour flexibility that single people have over married people.
Based on the discussion relating the coefficients of the non-economic variables (i.e., the coefficients in eq. (18)) to the relative growth rates of permanent income of different groups, the interpretation of the non-economic variables also accords well with intuition. Eq. (19) indicates that young (0-45) people have a higher rate of growth of permanent income than older persons. Young white women have a lower rate of growth of permanent income than their male counterparts. Young blacks have higher rate of growth of permanent income than their white peers. This is not an unreasonable finding, especially in view of the civil rights movement during the 1960's. In fact, in the previous section of this paper, it was found that blacks had a higher rate of current income growth over the 60-70 decade, than did whites.

As a glance at the variable definitions of eq. (19) will show, no distinction is made between black men and women, aged 0-45, while for whites such a distinction is made. F tests were able to accept the hypothesis that black men and women ages 0-45 have the same coefficients. The two stratified regressions for the age category 0-45 for blacks, indicate how close the relationships were for the males and females. These two regressions were

\[
\ln \frac{Y_c}{Y_p} = 0.060 + 0.038 \text{ EC} \quad \text{for black men 0-45, and}
\]

\[
\ln \frac{Y_c}{Y_p} = 0.065 + 0.038 \text{EC} \quad \text{for black women 0-45.}
\]

As an experiment to examine the sensitivity of the results of eq. (19) to the particular weighting scheme used for \( Y_f \), a straight four year arithmetic average was also used in the estimation of eq. (19). The result of that estimation is given below:
\[ (20) \quad \ln \frac{Y_c}{Y_p} = 0.078A + 0.0535 \cdot RA6 + 0.049 \cdot RA6 + 0.034 \cdot ENA4 + 0.0557 \cdot ENA6 + 0.025 \]
\[ (4.6) \quad (2.26) \quad (3.64) \]
\[ (2.158) \quad (2.25) \quad (2.63) \]

\[ R^2 = 0.04 \]
\[ F = 13.0 \]

Degree of Freedom = 1393

As a comparison with eq. (19) shows, the coefficients of economic variables seem quite robust to a change in the definition of \( Y_p \). (Shortly, it will be seen that these are the only coefficients of eq. (19) which will play an important part in the estimation of the distribution of permanent income.)

As might have been expected, the coefficients of the non-economic variables are less robust. However, the relative magnitude of the non-economic variables, as well as the economic variables, are unchanged between eq. (19) and eq. (20).

The performance of eq. (19) seems quite satisfactory. The F statistic is highly significant. (Recall that this is the statistic that will indicate if the equation is explaining anything at all). The signs and relative magnitudes of the coefficients are what one would theoretically expect. One coefficient is significant at the 90% level, while all others are significant at the 95% level. Furthermore, the results of eq. (19) seem reasonably robust to a different definition of \( Y_p \). Let us now turn to the results of the estimation of the variance equation (14b).
Results of Estimating the Variance Equation (14b)

The final form of eq. (14b) that was chosen to explain \( \ln(\frac{Y_c}{Y_p})^2 = \bar{E}( \ln(\frac{Y_c}{Y_p})^2) \) was:

\[
(21) \quad (\ln(\frac{Y_c}{Y_p}))^2 = 0.02495 RA4 + 0.0430 SA4 + 0.0236 NA4 + 0.0476 NA6 + 0.0282 \\
(2.88) \quad (2.72) \quad (1.7) \quad (2.8) \quad (7.816)
\]

\[R^2 = 0.024\]
\[F = 8.641\]

Degrees of Freedom = 1394, where

RA4 = 1 if the head of household is black, aged 45-65,
0 otherwise, and

SA4 = 1 if the head of household is a white female, aged 45-65,
0 otherwise,

and all other variables were defined beneath eq. (19). The left hand side of eq. (21)
was estimated by the square of the residuals for eq. (19).

It is interesting that no economic variables appear in eq. (21). In experimenting
with various specifications, the coefficients of economic variables\(^{(1)}\) were always very
small and very insignificant. Such a result seems quite possible, though.

Eq. (21) indicates that single persons below 64, and blacks and white women aged
45-65, have greater dispersion in \( \ln(\frac{Y_c}{Y_p}) \) than does the rest of the population.

\(^{(1)}\) Aside from using interaction variables, based on \( EC \), defined beneath eq. (19),
variables depending on \( |EC| \) and \( EC^2 \) were also tried.
This means that the percentage deviations in "transitory," income (i.e. $Y_c/Y_p$) are much larger for these groups than for the rest of the population.

As with eq. (19), the F statistic is highly significant, (so the equation does explain a significant amount more than a constant term alone would). One coefficient is significant at the 90% level, while the remaining four are significant at the 95% level.

Eq. (21) essentially explains the variance of the disturbance term $\eta_1$ in the equation (18) for $\ln Y_c/Y_p$. As explained in the footnote to eq. (21) should be used to correct for heteroskedasticity in eq. (18). Using eq. (21) to correct eq. (18) for heteroskedasticity one obtains the almost identical mean equation:

\[
(22) \quad \ln Y_c/Y_p = 0.0588 A + 0.03365 RA6 + 0.03730 ERA6 + 0.02442 ENA4 + 0.05167 ENA6 - 0.0304 \\
(4.641) \quad (1.94) \quad (3.778) \quad (1.66) \quad (2.27) \quad (2.98)
\]

$R^2 = .04$

Degree of Freedom = 1393

$F = 11.61$

**Sensitivity of Results to City Size**

Eqs. (19) through (22) were all run for individuals who resided in areas with cities of over 100,000 people. For a variety of reasons, one might believe that the coefficients of eq. (21) and (22) will be sensitive to different city sizes. To investigate this sensitivity, eqs. (21) and (22) were run on the subset of the
sample which resided in areas with cities of greater than 500 thousand people.

A Chow test was performed to determine whether the coefficients of the "variance" equation (21), differed by city size. The test accepted, at the 95% level, the hypothesis of equal coefficients for cities with populations of 100-500 thousand, and for cities with populations greater than 500 thousand. Once this hypothesis was accepted, a "quasi" F test was performed to determine if the coefficients of the mean equation (22) varied according to city size. Once again, the hypothesis of equal coefficients for cities of 100-500 thousand, and for cities of over 500 thousand was accepted at the 95% level.

In summary, eqs. (22) and (21) provide estimates of $f(\cdot)$ and $h(\cdot)$ of eqs. (14b) and (15). The estimated relations are highly significant statistically; most coefficients are significant at the 95% level, and all at the 90% level. The estimated coefficients have the correct signs, and relative magnitudes that are theoretically quite reasonable.

The coefficients were fairly robust when estimated under different conditions. Eqs. (21) and (22) are the final versions of the equations which will be used to represent the dependence of the mean and variance of $\ln(Y_c/Y_p)$ on socioeconomic variables.

(1) "quasi" - because eq. (21) was used to correct for heteroscedasticity. Asymptotically, the "quasi" F statistic becomes a true F statistic.
2.5 Predicting the Parameters of the Distribution of \( \ln(Y_c/Y_p) \) for Pittsburgh in 1960 and 1970

Since eqs. (21) and (22) represent the functional dependence of the parameters of the distribution of \( \ln Y_c/Y_p \) on socioeconomic characteristics, they can be used to predict the parameters of the distribution of \( \ln Y_c/Y_p \) for the demographic groups in Pittsburgh for 1960 and 1970\(^{(1)}\). If \( \hat{a} \) represents the estimated coefficients in eq. (22), and \( x_g \) represents the vector of socioeconomic characteristics describing group \( g \), then \( E(\ln Y_c/Y_p) \) can be estimated as \( \hat{a} \cdot x_g \), and the variance of this estimate will be \( x_g \cdot V \cdot x_g \), where \( V \) is the variance-covariance matrix of \( \hat{a} \). An analogous procedure will provide an estimate of \( \text{var}(\ln Y_c/Y_p) \) and the variance of this estimator.

There is a rather subtle point that arises in using eq. (22), though not eq. (21), to predict the desired parameters. In using eq. (22) for prediction, the non-economic terms in that equation should not be used. In terms of the discussion preceding eq. (17), what needs to be predicted is not \( E(\ln(Y_c/Y_p)^1) \), but rather \( E(\ln(Y_c/Y_p)^1 \cdot g^\lor Y_p = Y_p \). This latter expression will be predicted only if the non-economic coefficients in eq. (22) are ignored.

Using the fact that, for Pittsburgh, \( EC(1960) = .667 \), and \( EC(1970) = .3 \), eqs. (21) and (22) were used to generate predicted values of \( u = E(\ln Y_c/Y_p) \) and \( \sigma^2 = \text{var}(\ln Y_c/Y_p) \) for the demographic groups in Pittsburgh for the years 1960 and 1970.

\(^{(1)}\) Recall from \textbf{Section 2.2}, that when dealing with the Pittsburgh data the years which "1960" and "1970" refer to are usually 1959 and 1969.
1970. (1) Table 5 presents these predicted parameters \((u, \sigma^2)\) for the various demographic groups in Pittsburgh. Notice that in Table 5, the estimates of \(\sigma^2\) do not differ between 1960 and 1970 because no variable relating to economic conditions appears in eq. (21). Also, no distinction between owner and renter appears in Table 5 because none appears either in eq. (21) or eq. (22). (2)

Beside each prediction in Table 5 is the estimated standard error of this predicted value. As Table 5 indicates, the statistical significance of these predicted values is extremely good, with all predicted values being significant at the 95% level.

There is one problem, though. Predictions of the relevant parameters need to be made for the demographic groups of Table 1. However, as remarked in Sec. 2.4, for certain groups, finer age or sex breakdowns than those in Table 1, were used so as to better capture certain forces. For example, subgroup 4 (females under 65), is composed of females aged 0-45, and females aged 45-65. The assumption is that for each of these finer age categories \(\ln Yc/Yp\) has an approximate log-normal distribution. When these two finer age breakdowns are combined to obtain subgroup 4 of Table 1 (i.e., females under 65) an approximate log-normal distribution for \(\ln Yc/Yp\) is obtained for this combined category. (If two distributions are exactly log-normal, a weighted average of the distributions is definitely not log-normal. However, if two distributions are "approximately" log-normal, it is quite possible for the weighted average to be "approximately" log-normal.) The question then arises as to how the parameters of the combined aggregate age category should be related to those of the finer age categories.

---

(1) Actually, EC(1959) = -.667, and EC (1969) = .3. See Footnote 1 of previous page for an explanation of which years are used.

(2) An owner-renter variable was always small and insignificant in each equation.
TABLE 5a
Estimates of \( u \) for \( \ln Yc/Yp \)

<table>
<thead>
<tr>
<th>whites - 1969</th>
<th>Estimated Mean ( u )</th>
<th>Standard Error of ( Yc ) Estimate of ( u )</th>
</tr>
</thead>
<tbody>
<tr>
<td>subgroup (1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0-45</td>
<td>.0155</td>
<td>.00678</td>
</tr>
<tr>
<td>45-65</td>
<td>.0073</td>
<td>.00442</td>
</tr>
<tr>
<td>whites - 1959</td>
<td></td>
<td></td>
</tr>
<tr>
<td>subgroup</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0-45</td>
<td>-.0344</td>
<td>.0151</td>
</tr>
<tr>
<td>45-65</td>
<td>-.0163</td>
<td>.0098</td>
</tr>
<tr>
<td>blacks - 1969</td>
<td></td>
<td></td>
</tr>
<tr>
<td>subgroup</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>.0112</td>
<td>.00296</td>
</tr>
<tr>
<td>4</td>
<td>.0112</td>
<td>.00296</td>
</tr>
<tr>
<td>0-45</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>45-65</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>6</td>
<td>.0267</td>
<td>.00693</td>
</tr>
<tr>
<td>0-45</td>
<td>.0073</td>
<td>.00442</td>
</tr>
<tr>
<td>45-65</td>
<td></td>
<td></td>
</tr>
<tr>
<td>blacks - 1959</td>
<td></td>
<td></td>
</tr>
<tr>
<td>subgroup</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>-.0249</td>
<td>.00659</td>
</tr>
<tr>
<td>0-45</td>
<td>-.0249</td>
<td>.00659</td>
</tr>
<tr>
<td>45-65</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>6</td>
<td>.0593</td>
<td>.00154</td>
</tr>
<tr>
<td>0-45</td>
<td>-.0163</td>
<td>.0098</td>
</tr>
<tr>
<td>45-65</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

For any group not shown in this Table, \( u = 0 \).

(1) The subgroups refer to those of Table 1.
**TABLE 5-B**

Estimates of $\sigma^2$ for $\ln(Y_c/Y_p)$

<table>
<thead>
<tr>
<th>Whites</th>
<th>Estimated Variance $\sigma^2$</th>
<th>Standard Error of Estimate of $\sigma^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Subgroup</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>0.0282</td>
<td>0.0036</td>
</tr>
<tr>
<td>2</td>
<td>0.0282</td>
<td>0.0036</td>
</tr>
<tr>
<td>3</td>
<td>0.0282</td>
<td>0.0036</td>
</tr>
<tr>
<td>4</td>
<td>0.0282</td>
<td>0.0360</td>
</tr>
<tr>
<td>5</td>
<td>0.0712</td>
<td>0.0155</td>
</tr>
<tr>
<td>6</td>
<td>0.0282</td>
<td>0.0036</td>
</tr>
<tr>
<td>7</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| Blacks          |                               |                                        |
| **Subgroup**    |                               |                                        |
| 1               | 0.0282                        | 0.0036                                 |
| 2               | 0.0532                        | 0.0079                                 |
| 3               | 0.0282                        | 0.0036                                 |
| 4               | 0.0282                        | 0.0360                                 |
| 5               | 0.0712                        | 0.0155                                 |
| 6               | 0.0282                        | 0.0036                                 |
| 7               |                               |                                        |
Let \( c_1 \) = percent of the population in category 1
\( c_2 = \text{percent of the population in category 2} \)
\( c_2 = 1 - c_1 \)
\( f_1 = \text{distribution function of category 1, and} \)
\( f_2 = \text{distribution function of category 2.} \)

Consider \( f(x) = c_1 f_1(x) + c_2 f_2(x) \). Then, it can be shown\(^{(1)}\) that

\[
(23) \quad u_o = c_1 u_1 + c_2 u_2, \text{ and}
\]

and

\[
(24) \quad \sigma_o^2 = c_1 \sigma_1^2 + c_2 \sigma_2^2 + c_1 c_2 (u_1 - u_2)^2, \text{ where}
\]

\( u \) and \( \sigma^2 \) refer to the mean and variance of the relevant distributions.

After obtaining the particular values of \( c \) for the age and sex breakdowns
that are used in Table 5, eqs. (23) and (24), can be used with the results of Table 5,
to obtain the estimates of the parameters of the distribution of \( \ln Yc/Yp \) for the more
aggregated subgroup 4 (females under 65) and subgroup 6 (all singles under 65).
Table 6 presents the predicted values of \( \mu = E(\ln (Yc/Yp)) \) and \( \sigma^2 = Var(\ln Yc/Yp) \) for
the demographic groups of Table 1, for Pittsburgh in 1960 and 1970. For subgroups 4 and 6,
the 1960 variance estimate differs from that of 1970, because the relative composition
of each subgroup (i.e., the "c's" used in eqs. (23) and (24)) shifted between these two
years.

\( \quad \)

\( (1) \) See Appendix 4.
TABLE 6

Estimates of $u$ and $\sigma^2$ for $\ln \frac{Yc}{Yp}$ for the Demographic Groups of Table 1

### Estimates of $u$

<table>
<thead>
<tr>
<th>Subgroups (1)</th>
<th>1960</th>
<th>1970</th>
</tr>
</thead>
<tbody>
<tr>
<td>whites 6</td>
<td>-0.0245</td>
<td>0.0114</td>
</tr>
</tbody>
</table>

### Estimates of $\sigma^2$

<table>
<thead>
<tr>
<th>Subgroups</th>
<th>1960</th>
<th>1969 if Different From 1959</th>
</tr>
</thead>
<tbody>
<tr>
<td>whites</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>0.0282</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>0.0282</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>0.0282</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>0.0574</td>
<td>0.0539</td>
</tr>
<tr>
<td>5</td>
<td>0.0282</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>0.0775</td>
<td>0.0781</td>
</tr>
<tr>
<td>7</td>
<td>0.0282</td>
<td></td>
</tr>
<tr>
<td>blacks</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>0.0282</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>0.0532</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>0.0282</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>0.0367</td>
<td>0.0367</td>
</tr>
<tr>
<td>5</td>
<td>0.0282</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>0.0771</td>
<td>0.0766</td>
</tr>
<tr>
<td>7</td>
<td>0.0282</td>
<td></td>
</tr>
</tbody>
</table>

(1) The subgroups refer to those of Table 1.
2.6 Obtaining the Distribution of Permanent Income for Demographic Groups in Pittsburgh for 1960 and 1970

Combining the results of Sections 2.3, 2.4, and 4.5 all the ingredients are available to obtain the distribution of permanent income for the various demographic groups in Pittsburgh. Since \( \ln Yc = \ln Yc/Yp \), then, following the reasoning underlying eqs. (7) and (8), the following relations can be used to estimate the parameters of the distribution of \( \ln Yc \):

\[
(7) \quad u_{yp} = u_{yc} - u_{yc/yp}, \quad \text{and} \\
(8) \quad \sigma^2_{yp} = \sigma^2_{yc} - \sigma^2_{yc/yp}
\]

Table 3 can provide estimates of \( u_{yc} \) \( \sigma^2_{yc} \), while Table 6 can provide estimates of \( u_{yc/yp} \) and \( \sigma^2_{yc/yp} \).

Although it would be useful to have statistical confidence regions on our estimates derived from eqs. (7) and (8), unfortunately, these estimates do not have any tabulated distribution. However, some rough bounds can be obtained as follows.

Suppose \( c = a + b \), where \( a, b, c \) are all random variables. Let \( c^* = a^* + b^* \), and

\[
Pr(a < c^*) = 1 = Pr(b < b^*). \quad \text{Then} \quad Pr(c < c^*) \geq Pr(a < a^*) \cdot Pr(b < b^*) = (1 - \alpha)^2.
\]

Choosing \( \alpha = .999 \), we find that \( Pr(c < c^*) \geq 98.1 \%. \) Similarly, a lower limit, \( c^* \), can be found, such that \( Pr(c^* < c < c^*) \geq 96.2 \%. \) Essentially, a confidence region for \( a \), and \( b \) can be formed, and by adding together their upper and their lower limits, a confidence region for \( c \) can be obtained. These confidence intervals
will of course be much larger than exact 96.2% confidence intervals, but they will
give some idea about confidence bounds.\(^{(1)}\)

Table 7 presents the final estimates of the parameters of the distribution of \(\ln Y_p\)
for the demographic groups of Table 1 in Pittsburgh in 1960 and 1970, together with
their confidence intervals which were described above.

The general relation between the parameters of the permanent income
distributions of the different demographic groups are basically the same as those of
the current income distribution of Table 3. Hence, the general comments of Section 2.3
on comparisons between the income distributions between groups also apply to the
permanent income distribution of these groups with little change. Using the same
reasoning as in Section 2.3, where the rate of current income growth was
examined by looking at the shift in \(u\) between 60 and 70, it can be seen that in regard
to permanent income growth, the conclusions reached in Section 2.3 about the growth
of current income would have to be slightly altered to allow for a faster rate of
growth of permanent income for singles under 65, black males 0-45, and black females
under 65. A greater differential rate of permanent income growth for blacks (by .01)
than for whites is indicated, than was the case for current income growth. This
last fact corroborates the independent finding from the Michigan data, that the
growth rate of \(Y_p\) for young blacks exceeded that for their white peers. Similarly,
the results of Table 7 also corroborate the quite reasonable finding of eq. (21) that the

\(^{(1)}\) The following example will illustrate just how conservative a confidence interval,
derived in this manner, actually is. Suppose \(a, b\) are independent, and distributed
normally - \(N(0,1)\). Then an exact 98% confidence region is \([-3.29, 3.29]\). The exact
96.2% confidence region is, of course, smaller. However, the inexact 96.2% confidence
region, used above, would yield the region \([-4.66, 4.66]\).
### Table 7

#### Estimated Parameters of the Distribution of Permanent Income

<table>
<thead>
<tr>
<th>Group</th>
<th>mean, $u$</th>
<th>a conservative 96.2% confidence region</th>
<th>variance $\sigma^2$</th>
<th>a conservative 96.2% confidence region</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-1-1</td>
<td>8.797</td>
<td>[8.715, 8.879] *</td>
<td>.1819</td>
<td>[.1323, .2361]</td>
</tr>
<tr>
<td>1-1-3</td>
<td>8.348</td>
<td>[8.312, 8.384] *</td>
<td>.6121</td>
<td>[.5547, .6715]</td>
</tr>
<tr>
<td>1-1-4</td>
<td>8.394</td>
<td>[8.342, 8.446] *</td>
<td>.3401</td>
<td>[.2699, .4194]</td>
</tr>
<tr>
<td>1-1-5</td>
<td>8.38</td>
<td>[8.298, 8.462] *</td>
<td>.5047</td>
<td>[.3636, .6647]</td>
</tr>
<tr>
<td>1-1-6</td>
<td>7.919</td>
<td>[7.778, 8.059] *</td>
<td>.44</td>
<td>[.3068, .5840]</td>
</tr>
<tr>
<td>1-1-7</td>
<td>6.74</td>
<td>[6.47, 7.01] *</td>
<td>1.35</td>
<td>[.9910, 1.76]</td>
</tr>
<tr>
<td>1-2-1</td>
<td>8.556</td>
<td>[8.458, 8.65] *</td>
<td>.2008</td>
<td>[.1464, .2535]</td>
</tr>
<tr>
<td>1-2-2</td>
<td>8.495</td>
<td>[8.419, 8.571] *</td>
<td>.2362</td>
<td>[.1639, .3142]</td>
</tr>
<tr>
<td>1-2-3</td>
<td>8.088</td>
<td>[8.047, 8.129] *</td>
<td>.5264</td>
<td>[.4780, .5795]</td>
</tr>
<tr>
<td>1-2-4</td>
<td>7.856</td>
<td>[7.778, 7.934] *</td>
<td>.5097</td>
<td>[.4301, .5956]</td>
</tr>
<tr>
<td>1-2-5</td>
<td>7.892</td>
<td>[7.775, 8.009] *</td>
<td>.5116</td>
<td>[.3863, .6495]</td>
</tr>
<tr>
<td>1-2-6</td>
<td>7.679</td>
<td>[7.433, 7.925] *</td>
<td>.4143</td>
<td>[.2293, .6290]</td>
</tr>
<tr>
<td>1-2-7</td>
<td>6.996</td>
<td>[6.751, 7.241] *</td>
<td>.4835</td>
<td>[.3129, .6848]</td>
</tr>
<tr>
<td>2-1-1</td>
<td>8.506</td>
<td>[8.445, 8.567] *</td>
<td>.2088</td>
<td>[.1659, .2547]</td>
</tr>
<tr>
<td>2-1-2</td>
<td>8.655</td>
<td>[8.586, 8.724] *</td>
<td>.2978</td>
<td>[.2346, .3658]</td>
</tr>
<tr>
<td>2-1-3</td>
<td>8.151</td>
<td>[8.116, 8.184] *</td>
<td>.5997</td>
<td>[.5482, .6524]</td>
</tr>
<tr>
<td>2-1-4</td>
<td>7.952</td>
<td>[7.778, 8.134] *</td>
<td>.4273</td>
<td>[.2530, .6361]</td>
</tr>
<tr>
<td>2-1-5</td>
<td>6.020</td>
<td>[7.797, 8.243] *</td>
<td>.6665</td>
<td>[.3893, 1.014]</td>
</tr>
<tr>
<td>2-1-6</td>
<td>7.978</td>
<td>[7.836, 8.119] *</td>
<td>.4291</td>
<td>[.2941, .5795]</td>
</tr>
<tr>
<td>2-1-7</td>
<td>6.876</td>
<td>[6.783, 6.969] *</td>
<td>1.192</td>
<td>[.1.141, 1.312]</td>
</tr>
<tr>
<td>2-2-1</td>
<td>8.215</td>
<td>[8.102, 8.328] *</td>
<td>.2763</td>
<td>[.2050, .3549]</td>
</tr>
<tr>
<td>2-2-2</td>
<td>8.225</td>
<td>[8.142, 8.308] *</td>
<td>.2949</td>
<td>[.2147, .3812]</td>
</tr>
<tr>
<td>2-2-3</td>
<td>7.77</td>
<td>[7.662, 7.878] *</td>
<td>.4207</td>
<td>[.3078, .5336]</td>
</tr>
<tr>
<td>2-2-4</td>
<td>7.403</td>
<td>[7.224, 7.58] *</td>
<td>.5127</td>
<td>[.3811, .6609]</td>
</tr>
<tr>
<td>2-2-5</td>
<td>7.23</td>
<td>[6.900, 7.560] *</td>
<td>1.265</td>
<td>[.7101, 1.9788]</td>
</tr>
<tr>
<td>2-2-6</td>
<td>7.626</td>
<td>[7.298, 7.954] *</td>
<td>.2639</td>
<td>[.05, .5084]</td>
</tr>
<tr>
<td>2-2-7</td>
<td>6.676</td>
<td>[6.359, 6.993] *</td>
<td>.7009</td>
<td>[.4747, .9658]</td>
</tr>
<tr>
<td>3-1-3</td>
<td>8.657</td>
<td>[8.624, 8.690] *</td>
<td>.3436</td>
<td>[.2770, .4194]</td>
</tr>
<tr>
<td>3-1-4</td>
<td>8.796</td>
<td>[8.748, 8.844] *</td>
<td>.3436</td>
<td>[.2770, .4194]</td>
</tr>
<tr>
<td>3-1-5</td>
<td>8.72</td>
<td>[8.597, 8.843] *</td>
<td>.5283</td>
<td>[.3395, .7566]</td>
</tr>
<tr>
<td>3-1-6</td>
<td>8.365</td>
<td>[8.248, 8.482] *</td>
<td>.4393</td>
<td>[.3058, .5833]</td>
</tr>
<tr>
<td>3-1-7</td>
<td>7.363</td>
<td>[7.057, 7.669] *</td>
<td>1.145</td>
<td>[.7883, 1.562]</td>
</tr>
</tbody>
</table>
Table 7 (cont'd)

<table>
<thead>
<tr>
<th>Group</th>
<th>mean, u</th>
<th>α-conservative 96.2% confidence region</th>
<th>variance $\sigma^2$</th>
<th>α-conservative 96.2% confidence region</th>
</tr>
</thead>
<tbody>
<tr>
<td>3-2-1</td>
<td>9.142</td>
<td>[9.057, 9.227]</td>
<td>.5107</td>
<td>[0.4311, 0.5949]</td>
</tr>
<tr>
<td>3-2-3</td>
<td>8.342</td>
<td>[8.304, 8.380]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3-2-4</td>
<td>8.385</td>
<td>[8.318, 8.452]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3-2-5</td>
<td>8.297</td>
<td>[8.181, 8.413]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3-2-6</td>
<td>8.124</td>
<td>[7.929, 8.319]</td>
<td>.4148</td>
<td>[0.2293, 0.6290]</td>
</tr>
<tr>
<td>3-2-7</td>
<td>7.391</td>
<td>[7.180, 7.602]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4-1-1</td>
<td>9.043</td>
<td>[8.975, 9.111]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4-1-2</td>
<td>9.140</td>
<td>[9.062, 9.218]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4-1-3</td>
<td>8.513</td>
<td>[8.479, 8.547]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4-1-4</td>
<td>8.41</td>
<td>[8.331, 8.489]</td>
<td>.4085</td>
<td>[0.3022, 0.5229]</td>
</tr>
<tr>
<td>4-1-5</td>
<td>8.453</td>
<td>[8.224, 8.682]</td>
<td>.5511</td>
<td>[0.4183, 0.6989]</td>
</tr>
<tr>
<td>4-1-6</td>
<td>8.396</td>
<td>[8.276, 8.516]</td>
<td>.4285</td>
<td>[0.2931, 0.5752]</td>
</tr>
<tr>
<td>4-1-7</td>
<td>7.336</td>
<td>[7.256, 7.416]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4-2-1</td>
<td>8.337</td>
<td>[8.74, 8.932]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4-2-2</td>
<td>8.829</td>
<td>[8.754, 8.904]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4-2-3</td>
<td>8.12</td>
<td>[8.054, 8.186]</td>
<td>.5328</td>
<td>[0.4507, 0.6201]</td>
</tr>
<tr>
<td>4-2-4</td>
<td>7.904</td>
<td>[7.758, 8.050]</td>
<td>.5137</td>
<td>[0.3821, 0.6602]</td>
</tr>
<tr>
<td>4-2-5</td>
<td>8.14</td>
<td>[8.059, 8.221]</td>
<td>.559</td>
<td>[0.4335, 0.6724]</td>
</tr>
<tr>
<td>4-2-6</td>
<td>7.86</td>
<td>[7.605, 8.123]</td>
<td>.5602</td>
<td>[0.2833, 0.8959]</td>
</tr>
<tr>
<td>4-2-7</td>
<td>6.987</td>
<td>[6.698, 7.276]</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Footnotes:

1. For explanation of notation, see footnote 1 of Table 2.

2. A "**" indicates an exact 98% confidence region.

Because of the way estimates for subgroups 4 and 6 were constructed, the confidence regions reported above are conservative 94% regions.
young tend to have higher rates of permanent income growth than do the older members of the population.

The differences in the estimates of the mean \( \mu \), between Tables 3 which deals with current income and \( \gamma \), which deals with permanent income, are quite small, and are always less than 1/4 of 1%. On the other hand, the difference in the estimates of \( \sigma^2 \) between Tables 3 and 7 are much more substantial, ranging from a low of 2 1/2% for groups 2-1-7 (renters, white, single over 65 in 1960 and 1970), to a high of 20% for group 2-2-6 (renters, black, singles under 65, 1960). To facilitate comparison, Table 8 presents the estimates of \( \sigma^2_{\ln Y_c} \) derived from Table 3, and the \( \sigma^2_{\ln Y_p} \) of Table 7. The overwhelming majority of cases involve reductions in \( \sigma^2 \) which range from 5 to 14%. This last result shows that as expected, \( \ln Y_c \) has greater variance than \( \ln Y_p \).

The main purpose of this paper was to derive the distribution of \( Y_p \), because it was felt that it differed significantly from the distribution of \( Y_c \). It was felt that the distribution of \( Y_c \) was much more dispersed than that of \( Y_p \). To obtain a general idea of the reduction in the variance in \( Y_c \), not \( Y_p \), that occurs by using permanent instead of current income, consider the following typical example. Let \( (\mu_p, \sigma^2_p) \) and \( (\mu_c, \sigma^2_c) \) stand for the mean and variance of the logarithm of permanent and current income respectively. Then, the reduction, \( R \), in the variance that occurs by using the permanent, not current, income distribution can be computed from the following relation

\[
1 - R = e^{2\mu_p} \cdot e^{\sigma^2_p} (e^{\sigma^2_p} - 1) / e^{2\mu_p} \cdot e^{\sigma^2_c} (e^{\sigma^2_c} - 1)
\]

(1)

(1) If \( Y \) is distributed log-normally with parameters \( (\mu, \sigma) \), then var \( Y = e^{2\mu} + \sigma^2(e^{\sigma^2} - 1) \).
Comparison of Estimates of $\sigma^2$ for $\ln Y_c$ and $\ln Y_p$

<table>
<thead>
<tr>
<th>group</th>
<th>$\sigma^2$ for $\ln Y_c$</th>
<th>$\sigma^2$ for $\ln Y_p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-1-1</td>
<td>.2101</td>
<td>.1819</td>
</tr>
<tr>
<td>1-1-2</td>
<td>.2784</td>
<td>.2502</td>
</tr>
<tr>
<td>1-1-3</td>
<td>.6403</td>
<td>.6121</td>
</tr>
<tr>
<td>1-1-4</td>
<td>.3975</td>
<td>.3401</td>
</tr>
<tr>
<td>1-1-5</td>
<td>.5329</td>
<td>.5047</td>
</tr>
<tr>
<td>1-1-6</td>
<td>.5174</td>
<td>.4400</td>
</tr>
<tr>
<td>1-1-7</td>
<td>1.378</td>
<td>1.330</td>
</tr>
<tr>
<td>1-2-1</td>
<td>.2290</td>
<td>.2008</td>
</tr>
<tr>
<td>1-2-2</td>
<td>.2894</td>
<td>.2362</td>
</tr>
<tr>
<td>1-2-3</td>
<td>.5546</td>
<td>.5264</td>
</tr>
<tr>
<td>1-2-4</td>
<td>.5464</td>
<td>.5097</td>
</tr>
<tr>
<td>1-2-5</td>
<td>.5398</td>
<td>.5116</td>
</tr>
<tr>
<td>1-2-6</td>
<td>.4914</td>
<td>.4143</td>
</tr>
<tr>
<td>1-2-7</td>
<td>.5117</td>
<td>.4835</td>
</tr>
<tr>
<td>2-1-1</td>
<td>.2370</td>
<td>.2098</td>
</tr>
<tr>
<td>2-1-2</td>
<td>.3260</td>
<td>.2978</td>
</tr>
<tr>
<td>2-1-3</td>
<td>.6279</td>
<td>.5997</td>
</tr>
<tr>
<td>2-1-4</td>
<td>.4847</td>
<td>.4273</td>
</tr>
<tr>
<td>2-1-5</td>
<td>.6947</td>
<td>.6665</td>
</tr>
<tr>
<td>2-1-6</td>
<td>.5066</td>
<td>.4291</td>
</tr>
<tr>
<td>2-1-7</td>
<td>1.22</td>
<td>1.198</td>
</tr>
<tr>
<td>2-2-1</td>
<td>.3045</td>
<td>.2763</td>
</tr>
<tr>
<td>2-2-2</td>
<td>.3481</td>
<td>.2949</td>
</tr>
<tr>
<td>2-2-3</td>
<td>.4489</td>
<td>.4207</td>
</tr>
<tr>
<td>2-2-4</td>
<td>.5494</td>
<td>.5127</td>
</tr>
<tr>
<td>2-2-5</td>
<td>1.293</td>
<td>1.265</td>
</tr>
<tr>
<td>2-2-6</td>
<td>.341</td>
<td>.2639</td>
</tr>
<tr>
<td>2-2-7</td>
<td>.7291</td>
<td>.7009</td>
</tr>
<tr>
<td>3-1-1</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>3-1-2</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>3-1-3</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>3-1-4</td>
<td>*(.3975)</td>
<td>.3436</td>
</tr>
<tr>
<td>3-1-5</td>
<td>.5565</td>
<td>.5283</td>
</tr>
<tr>
<td>3-1-6</td>
<td>*(.5174)</td>
<td>.4393</td>
</tr>
<tr>
<td>3-1-7</td>
<td>1.173</td>
<td>1.145</td>
</tr>
</tbody>
</table>
### Table 8 (cont'd)

<table>
<thead>
<tr>
<th>Group</th>
<th>$\sigma^2$ for $\ln Y_e$</th>
<th>$\sigma^2$ for $\ln Y_p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>3-2-1</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>3-2-2</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>3-2-3</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>3-2-4</td>
<td>*(.5464)</td>
<td>.5107</td>
</tr>
<tr>
<td>3-2-5</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>3-2-6</td>
<td>*(.4914)</td>
<td>.4148</td>
</tr>
<tr>
<td>3-2-7</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>4-1-1</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>4-1-2</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>4-1-3</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>4-1-4</td>
<td>.4624</td>
<td>.4085</td>
</tr>
<tr>
<td>4-1-5</td>
<td>.5793</td>
<td>.5511</td>
</tr>
<tr>
<td>4-1-6</td>
<td>*(.5066)</td>
<td>.4285</td>
</tr>
<tr>
<td>4-1-7</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>4-2-1</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>4-2-2</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>4-2-3</td>
<td>.561</td>
<td>.5328</td>
</tr>
<tr>
<td>4-2-4</td>
<td>.5494</td>
<td>.5137</td>
</tr>
<tr>
<td>4-2-5</td>
<td>.5867</td>
<td>.559</td>
</tr>
<tr>
<td>4-2-6</td>
<td>.6268</td>
<td>.5602</td>
</tr>
<tr>
<td>4-2-7</td>
<td>*</td>
<td>*</td>
</tr>
</tbody>
</table>

1. See footnote 1, Table 2 for an explanation of notation.

A "*" in a subgroup I-J-K means that the value is identical to the one listed for subgroup (I-2) - J-K. (i.e., the same group in 1960).
Typical values for the parameters are $u_p = u_c$, and $\sigma^2_p = \sigma^2_c = .03$. (Actually, $\sigma^2_c$ is the smallest correction made, but it is made quite frequently. The reduction in variance is even greater when $\sigma^2_c - \sigma^2_p > .03$.) For $\sigma^2 = .4$, $R = .87$, for a 13% reduction in variance, while for $\sigma^2 = .6$, $R = .90$, and for $\sigma^2 = .20$, $R = .81$, for a 19% reduction in the variance. In few cases where $u_p$ exceeds $u_c$ (only for certain groups in 1960) the reduction in variance is slightly less (by at most 4%) than the typical case, while in those cases where $u_p$ is less than $u_c$ (only in 1970) the reduction in the variance is slightly greater than the typical example would indicate. Noting from Table 8, that $\sigma^2_c$ is quite low for the two largest categories, subgroups 1 and 2 (male headed 0-45, and 45-65), it is seen that quite sizable reductions (about 17%) in the variance of the income distribution have been achieved for a very large segment of the population.

In summary, Table 7 represents the estimates of the parameters of the distribution of permanent income for the demographic groups in which this study is interested. The results of Table 7 corroborate the independent findings from the Michigan data about the relative growth rates of permanent income among black and white demographic groups. Compared to the current income distributions of Table 3, the results of Table 7 imply that the dispersion of permanent income $Y_p$, is quite substantially below that of current income, $Y_c$, for major sectors of the population. The sizable differences between the distribution of $Y_p$ and $Y_c$ emphasizes the need to carefully estimate the distribution of $Y_p$, when it is believed that market behavior depends on permanent and not current income.

(1) For example, in 1970, subgroups 1 and 2 represented roughly 70% of the total population in Pittsburgh.
Conclusions

This paper has developed a new approach for estimating the distribution of permanent income for a group of individuals for whom only current income data is available. The assumption of log-normality of current income seems to work surprisingly well, and affords great simplification. (Appendix 1 discusses a more sophisticated and much more complicated approach which does not impose the log-normality assumption). The equations which estimate the dependence of $E(\ln Y_c/Y_p)$ and $\text{var}(\ln Y_c/Y_p)$ on socioeconomic variables produced quite reasonable results. Interpretations based on the results of Table 7 and eq. (21), regarding the relative growth rate of permanent income among various demographic groups, are in close agreement with each other.

The distributions of permanent and current income were seen to differ significantly from each other. The derived permanent income distributions, as one would hope, have a considerably smaller dispersion than those of current income for a large segment of the population. It would seem that only by using an accurate estimate of the true purchasing power, on which consumers base their decisions, can one expect to properly model the impact of forces in the housing market, and be able to predict the effects of the housing allowance.

The ability to derive the distribution of permanent income from only a one-period cross-sectional distribution of current income appears to be quite valuable and should have general applicability in several areas of applied economic research.
APPENDIX 1

An Alternativc Solution Method which Allows Very General Parametrizations of the Distribution Functions of the Underlying Random Variables

In Section 22 of this paper, it was shown that the problem we face can be stated as follows:

There are three random variables $x, y, z$, $z = x + y$ and $x$ and $y$ are independent. It is desired to estimate the distribution of $z$ for some group, for whom data only on $z$ is available. Somehow, the distribution of $y$ across each group, and then this estimated distribution must be combined with the estimate of the distribution of $z$, to obtain the distribution of $x$.  

One approach to this problem is to parametrize the distributions of $x$, $y$ and $z$ in terms of a few estimable unknowns. This was the approach adopted in the paper. It was seen that a parametrization in terms of log normal distributions was able to fit the data quite well and was able to generate quite useful and reasonable results. In this appendix, more general methods will be outlined which avoid assumption of a particular form of the distribution functions of $x$, $y$ and $z$.

From eq. (4), it is seen that the Fourier transform of both $z$ and $y$ must be estimated for each group. Since data on $z$ is available for each group, the Fourier transform of $z$ can be estimated as follows:

Divide the $z$-range into intervals. For each interval $[z_i, z_{i+1}]$, estimate

$$P(z \in [z_i, z_{i+1}]) = P_i. \text{ Then form } E(e^{itz}) = \mathcal{P}_z(t) = \sum e^{iz_k} + \frac{1}{z_{i+1} - z_i}.$$ 

(1) Professor R. Hall first suggested this alternative type of approach to me.

(2) $\mathcal{P}$ stands for a characteristic function.
as an approximation to the characteristic function of z.

The more difficult problem is estimating \( \Phi_y(t) \) for a particular group. Since data on y only is not available for any group, an independent data source must be used to estimate \( \Phi_y(t) \) as a function of socioeconomic variables \( s \). Once \( \Phi_y(t/s) \) has been estimated as a function of \( s \), the \( s^* \) of each particular group can be used to predict \( \Phi_y(t/s^*) \). To estimate \( \Phi_y(t/s) \) as a function of \( s \), from our independent data source, one can proceed as follows:

1. Fix \( k \). Then regress the imaginary value \( e^{i \cdot k \cdot s} \) on the vector of observed socioeconomic variables (including possible interaction terms) \( s \).

Then, \( \text{E}(e^{i \cdot k \cdot s}) = \overset{\sim}{\alpha}_k \cdot s \).

Notice that the parametrization of the distribution of y is achieved by the vector \( \overset{\sim}{\alpha}_k \), for all values of \( k \). We may want to constraint our estimated \( \Phi_y(t/s) \) further, by requiring certain coefficients of \( \overset{\sim}{\alpha}_k \) to be equal for all \( k \).

Once \( \Phi_z(t) \), and \( \Phi_y(t/s^*) \) have been estimated, the following equation\(^{(1)}\)

\[
(A-1) \quad f(x/s^*) = \sum_{k \in K} e^{-ikx} \frac{\Phi_z(k)}{\Phi_y(k/s^*)}.
\]

Of course, careful thought must be given as to which values of \( k \) to sum over in eq. (A-1) to insure consistency of the estimator \( f(x/s) \).

---

\( (1) \) This is the same equation as eq. (3), except for notation.
In light of the fact that the log normal assumption seems to be reasonably consistent with the data, and in light of the favorable results that were obtained with the log-normality assumptions, the more general and more complicated method just outlined, is not desirable for the purposes of this particular study. However, this alternative approach does have the potential of being quite useful in the empirical analysis of distribution functions.
APPENDIX 2
Data Problems

Renters in 1970

For the "Renter" category in 1970, only three observations were available for the cumulative income distribution in "Metropolitan Housing Characteristics - HC-2" (MHC). To remedy this situation, an alternative data source, "The Census Population Characteristics (CPP), was used. Unfortunately, because of slightly different coverage and definitions, the numbers were not identical to those in the MHC. To obtain the points on the cumulative income frequency curve for renters, the same relative distribution of income, as indicated in the CPP, was used to distribute the people in the three large income categories of MHC into smaller income classes.

"Other Males"

In MCH, the categories "other male" are defined to include male head, wife present, nonrelative present, and male head, wife not present.

These definitions to distinguish "other male" from "male" seemed rather arbitrary, so it was felt best to combine the "other male" and "male" category. In combining these categories, it was assumed that the percent of the "other males" under 65 population which is below 45 was the same as for the "male category".

As mentioned in Section 2.3, the "other male" categories form a very small percent of the population. To test the sensitivity of the results of Tables 2 and 3 to the inclusion of the
"other male" category with the "male" category, log-normal equations were run on the "male" categories alone. As expected, in all cases, there were negligible differences between these results and those reported in Tables 2 and 3 for the "male" subgroups (i.e., subgroups 1, 2, and 3 of Table 1).
APPENDIX 3 - Graph Between Inflation and Unemployment across SMSA's in 1970

Change in the CPI between 1967 and 1970

19.0
18.0
17.0
16.0
15.0
14.0

.3 .5 10 15

1970 Deviation from a three year average of unemployment.

Appendix 4 - Derivation of Eqs (23) and (24)

Section 2.5

Using the notation of $\Lambda$, consider

$$f_o(x) = c_1 f_1(x) + c_2 \cdot f_2(x)$$

Then,

$$u_o - E_o(x) - f_o(x) \cdot x = c_1 \cdot x \cdot f_1(x) \ dx + c_2 \cdot x \cdot f_2(x) \ dx$$

or

$$u_o = c_1 u_1 + c_2 u_2.$$ 

Also,

$$E_o^2 = \mathbb{E}(x - u_o)^2 = \int (x - (c_1 u_1 + c_2 u_2))^2 \cdot f_o(x) \ dx$$

$$= \int [x - u_1 + (1-c) u_1 - c_2 u_2]^2 \cdot c_1 \cdot f_1(x) \ dx +$$

$$\int [x - u_2 - c_1 u_1 + (1-c_2) u_2]^2 \cdot c_2 \cdot f_2(x) \ dx$$

$$= c_1 \int (x - u_1)^2 \cdot f_1(x) \ dx + \int [(1-c_1) u_1 - c_2 u_2]^2 \cdot c_1 \cdot f_1(x) \ dx$$

$$+ 2c_1 \int (x-u_1)[(1-c_1) u_1 - c_2 u_2] \cdot f_1(x) \ dx +$$

$$c_2 \int (x-u_2)^2 \cdot f_2(x) \ dx + \int [c_1 u_1 + (1-c_2) u_2]^2 \cdot c_2 \cdot f_2(x) \ dx$$

$$+ 2c_2 \int (x-u_2)[c_1 u_1 + (1-c_2) u_2] \cdot f_2(x) \ dx.$$ 

But since $\int (x-u_1) f_1 \ dx = 0$, and $1-c_2 = c_1$, we obtain that
\[ \sigma^2_0 = c_1 \sigma^2_1 + c_2 \sigma^2_2 + c_1 c_2 (u_2 - u_1)^2 + c_1^2 c_2 (u_2 - u_1)^2 \], or

\[ \sigma^2_0 = c_1 \sigma^2_1 + c_2 \sigma^2_2 + c_1 c_2 (u_2 - u_1)^2. \]
Income Elasticities, Estimation of $\alpha$, and Construction of Model Households

3.1 Modifications in Income Elasticity Assumptions

A thesis of the Joint Center Study is that income elasticities are significantly below 1. The results of other sections of the Joint Center Study (see chapter 4) indicate that the income elasticities across different groups are fairly constant and are approximately equal to .47. (1)

In order to modify the income elasticities in the model, we adopt the approach of modifying the first two terms in the utility function used in the model. The current version of the utility function is of the form

$$U = (X_1 - A_1) \alpha_1 (X_2 - A_2) \alpha_2 \ldots \ldots$$

which is a Cobb-Douglas type utility function for which the $\alpha$'s are positive, the $X$'s represent quantities of goods and the $A$'s represent expectations. (Note that the utility function is undefined unless some minimum amount of each good is obtained.)

Presently, the model assumes $\alpha_2 = 1 - \alpha_1$, $X_1 = Q$, $X_2 = Y - P \cdot Q$, $A_1 = \frac{Y_1}{h}$, $A_2 = \frac{Y_1}{h} Y$, where $P_H$ is the price (per unit of service) of new housing, $P$ is the price per unit of the house chosen and $Y$ represents permanent income. When $P$, the per unit price of housing relative to the price of other goods, is constant and equal to $P_H$, then

$$(3, 1) P A_1 + A_2 = \frac{Y_1}{h} Y$$

so that $\frac{Y_1}{h}$ may be interpreted as some "minimum acceptable expenditure level"

(1) For other studies which indicate income elasticities below 1, see the discussion in Chapter 4 of the Joint Center Study.
which rises linearly with income.

For this constant P situation, the partial of U with respect to \( Q \) becomes

\[
\frac{\partial U}{\partial Q} = \alpha \frac{Q}{Q - A_1} \left( Y - PQ - A_2 \right)^{-1} + \alpha l
\]

Setting \( \frac{\partial U}{\partial Q} = 0 \) and substituting for \( A_2 \) from eq (3.1), the currently implied demand equation for housing is

\[
(3.2) \quad Q = \frac{\alpha l Y (1 - Y)}{P} + A_1
\]

Since \( A_1 = Y / P \) for the UI model, (1) the quantity of housing demanded remains proportional to \( Y \) implying an income elasticity equal to 1.0.

The expectation terms, \( A_1 \) and \( A_2 \), can be altered to allow income elasticities below one. Eq. (3.1) is maintained since a linear form for the minimum expenditure level is attractive, but a constant term is added to \( A_1 \) so that

\[
(3.3) \quad A_1 = \alpha \frac{Y}{P} + \frac{\alpha l b}{P}
\]

The first term in (3.3) reflects expectations that rise linearly with income, while the second term states that there is a component of minimum expectations which is independent of an individual's income. The value of "b" could depend on the particular demographic group to which the individual belongs.

Using the new definition of \( A_1 \), given in eq. (3.3), the demand equation now becomes

\[
(3.4) \quad Q = \frac{Y}{P} + \frac{b}{P}
\]

implying a preferred rental expenditure of

(1) Henceforth, UI will be the abbreviation used for the Urban Institute.
(3.5) \[ \text{Rent} = PQ = \omega_1 (Y + b), \]

holding accessibility, and neighborhood externalities and racial composition constant.

Given that (3.4) defines the demand curve that applies for constant P conditions, the income elasticity \( E \) becomes

\[
(3.6) \quad E = \frac{Y}{Q} \cdot \frac{\partial Q}{\partial Y} = \omega_1 \frac{Y}{P} \omega_1 (Y + P) = \frac{Y}{Y + b}.
\]

As long as \( b > 0, E < 1 \), as desired. (One drawback of this approach is that \( E \) increases steadily to one as income rises.)

Although the introduction of the constant \( \frac{\omega_1 \omega_1}{P} \) to the expectation term \( A \) has a justifiable interpretation, it should be emphasized that the \( A \) term was altered mainly to allow for income elasticities below one, and not necessarily, to model expectations more closely. The linear demand equation, which follows directly from the form of the utility function, is not the most reasonable specification for the housing demand relationship. (A log-log specification, as used in most empirical studies, would be preferred). Such a linear relationship should be regarded as a useful approximation which arises from choosing a form for the utility function which is manageable in a computer model. Rather than debating about whether \( A_1 \) accurately reflects the formation of expectations, it is much more important to choose \( A_1 \) so that the market behavior in the model is consistent with those measures of market behavior about which we do have a great deal of evidence, namely income elasticities. The purpose of this modelling effort is to make enough simplifying assumptions (e.g. about the form of the utility function, and \( A_1 \)) to make the model, manageable, yet still enable the model to accurately reflect important market phenomena, like income elasticities.
3.2 Estimation of $\kappa$

An important part of calibrating the model is to estimate the $\kappa$ parameters for different household types. The demand equation (3.4) and its equivalent statement about rental expenditure in (3.5) provide a method for econometrically estimating $\kappa$ for the model specification which incorporates the modified assumptions about the terms $A_1$ and $A_2$ in the utility function.

Let individual $i$ be a member of household type $j$, and let $Y_{m,j} =$ mean current income of group $j$, $Y_i =$ permanent income of individual $i$, $q =$ number of units of service demanded, $b_j =$ value of $b$ for group $j$, and $p =$ price per unit $q$. Accordingly, eq (3.4) can be expressed as:

$$(3.7) \text{Rent}_i = \kappa_i (Y_i + b_i) + e_i,$$

which may be written as:

$$(3.8) \text{Rent}_i = \kappa_i (Y_{m,j} + b_j) + \kappa_i (Y_i - Y_{m,j}) + e_i$$

The new term, $e_i$, is used to represent a random disturbance term, whose expected value is taken to be zero. If the term $Y_i - Y_{m,j}$ has a mean $\mu_j$, then eq (3.8) can be rewritten as:

$$(3.9) \text{Rent}_i = \kappa_i (Y_{m,j} + b_j + \mu_j) + n_i$$

where $n_i$ is a stochastic error term, with zero mean. Taking expectations of both sides of eq (3.9) over all individuals $i$ in group $j$, the following equation is obtained:

$$(3.10) \text{Rent}_j = \kappa_j (Y_{m,j} + b_j + \mu_j),$$

where $\text{Rent}_j$ stands for average rent of group $j$.

$$(1) \quad n_j = E_i + \kappa_j (Y_i - Y_{m,j} - \mu_j)$$
The value of the variables of the right-hand-side of eq. (3.11) is known or estimable for each of the demographic groups Table 1 of Part 2. From the 1960 and 1970 Census Data for Pittsburgh (1), series HC 21-168.1, estimates of $Y_{ij}$ and Rent $m_i$ can be formed. The term $\mathcal{u}_j$ is interpretable as the negative of the expected value of transitory income for group $j$. (2) Since, the distributions of current and permanent income have been estimated in the previous chapter, estimates of $\mathcal{u}_j$ can easily be formed for each relevant demographic group. The value of $b_j$ was chosen so that the income elasticity $E_j$ of group $j$ equaled .47 when evaluated at the median income of the group. Recall that a value of .47 for $E$ is consistent with the findings of other work done by the Joint Center.

Using eq (3.11) the values of $\mathcal{K}_j$ were computed for each of the demographic groups of Table 1 of Part 2. Then, to obtain the value of $\mathcal{K}$ for the more aggregate groupings used in the UI model, a weighted average of the $\mathcal{K}$ values was formed for each of the aggregated groupings, with the weights proportional to the relative size of each particular group in its aggregate group. The values obtained for in 1960 and 1970 for the four aggregated household types used in the model are presented below.

<table>
<thead>
<tr>
<th>Household Type</th>
<th>$\mathcal{K}$ value in 1960</th>
<th>$\mathcal{K}$ value in 1970</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. white non-elderly</td>
<td>.086</td>
<td>.073</td>
</tr>
<tr>
<td>2. white elderly/single</td>
<td>.140</td>
<td>.140</td>
</tr>
<tr>
<td>3. black non-elderly</td>
<td>.108</td>
<td>.102</td>
</tr>
<tr>
<td>4. black elderly/single</td>
<td>.143</td>
<td>.149</td>
</tr>
</tbody>
</table>

(1) This is the same source as the Urban Institute uses.

(2) To see this point, observe that (using the notation of Part 2) since $Y_{ij} = E(Y_{ij})$ for individual $i$ in group $j$, then $u_j = E(Y_{ij}) - E(Y_{ij})$, which is simply the negative of the expected value of transitory income in group $j$. Recall, from Part 2, that the value of $u_j$ will depend on economic conditions, and that in 'normal' economic conditions, $u_j = 0$. 

The relative ordering, of the $\lambda$ parameters across the four household types is the same as the ordering implied by $\lambda$ estimates of the Urban Institute which uses the expression (3.3) for $A_{ij}$ with $b$ set equal to 0. The values of $\lambda$ in Table (3-1) are roughly one half the values estimated by the Urban Institute. This relative ratio of one-half is just what should be expected. Eq (3.11) shows that if $\lambda_j$ is estimated from rent-income ratios, then $b_j = u_j = 0$. Since $b_j$ is of the same order of magnitude as $Y_m$, it immediately follows that the estimates of $\lambda$ in Table (3, 1) will be about one-half of those that are based on rent-income ratios (as are the UI estimates of $\lambda$).

3.3 Construction of Model Households

In order to run the housing model, it is necessary to divide the population of Pittsburgh into a small number of "model" households for both 1960 and 1970. In this study, we will use the same number of "model" households per household type as did the Urban Institute in their Pittsburgh study.

To obtain an estimate of the permanent income of each "model" household, the results of the permanent income estimation performed earlier are used. For each of the four aggregate household types used in the model, a weighted average is formed of the cumulative distribution function of permanent income for each of the demographic groups (of Table 1 of Part 2) comprising each aggregate household type. The weights are proportional to the relative size of each demographic group in its aggregate group. In this way, cumulative distribution functions of permanent income were formed for each of the four aggregate household types. Using the same number of "model" households for each household type as the Urban Institute, each of these
four cumulative distributions was divided into fractiles (each fractile corresponding to one "model" household), and the mean permanent income of each fractile was then determined. These 39 and 41 estimates of permanent income represent the estimated values of permanent income in 1960 and 1970, respectively, for the 39 and 41 model households used in the 1960 and 1970 runs of the model.

Two types of minor alterations in these values for permanent income were performed. In each case, the alterations involve behavior at the extremes of the income distribution.

First, it was felt that the influence of a few very rich households would be overrepresented if the mean income were used to represent purchasing power in the richest fractile. The underlying reasoning is that the very high purchasing power of the few households at the extreme upper end of the income distribution should not significantly affect the entire demand behavior of the richest fractile. Using the mean income of the richest fractile would cause an overestimate of the purchasing power for housing for this richest fractile. For this reason, it was felt that the median income of that richest fractile would provide a more accurate indicator of housing demand than would the mean income of that fractile.

Secondly, for the three poorest type 2 (white elderly/single) households in 1960 and the two poorest type 2 households in 1970, the estimated permanent income of each household is below $2,000. Although it is reasonable for their incomes to be below $2000, the predicted incomes were much too low to be credible. Whenever incomes are predicted to be below $2000, one must exercise caution in interpreting the results because the data source used, Metropolitan Housing Characteristics, does not provide any information on the income distribution below $2000. Also, any

(1) A description of household types 1, 2, 3, 4 appears in Table 3.1. We did not feel this "overrepresentation" problem would arise for the black households, i.e. household types 3 and 4.
assumptions (log-normal or otherwise) about the distribution of income are bound to be violated in the very low income range because of a) the substantially different mechanisms of income formation for the very poor, and b) the income "floor" which is created by the government's welfare programs. For these reasons, it was felt best to modify the estimates of permanent income of these poorest type 2 households.

Instead of accepting the predicted figures for permanent income, which were felt to be unrealistically low, the following distributional assumptions were made. In 1960, it was assumed that the probability density function of income below $2071, was a straight line beginning at $375 per year. In 1970, the density function of income below $1993 was again assumed to be a straight line, beginning at $5000 per year. Using these new distributional assumptions, the permanent income of the bottom fractiles of type 2 households was estimated. The final figures used to represent the permanent income of the households in 1960 and 1970 are presented in Table (3.2). For comparison purposes, the permanent income figures used by the Urban Institute are also presented in Table (3.2). Although the income figures appear somewhat similar for the two columns of Table (3.2), close examination shows that the Joint Center figures imply a greater dispersion of permanent income, especially for the richer households of type 1 and 2, than do the Urban Institute figures.  

(1) Because of our use of the log normal assumption for current income, our current income distribution (not reported) for "model" households also differs from that of the Urban Institute.
<table>
<thead>
<tr>
<th>Household Type</th>
<th>Permanent Income used by Joint Center</th>
<th>Permanent Income used by U.I.</th>
<th>Permanent Income used by Joint Center</th>
<th>Permanent Income used by U.I.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1778</td>
<td>1811</td>
<td>2</td>
<td>1021</td>
</tr>
<tr>
<td>1</td>
<td>2600</td>
<td>2834</td>
<td>2</td>
<td>1550</td>
</tr>
<tr>
<td>1</td>
<td>3059</td>
<td>3617</td>
<td>2</td>
<td>1905</td>
</tr>
<tr>
<td>1</td>
<td>3423</td>
<td>4107</td>
<td>2</td>
<td>2358</td>
</tr>
<tr>
<td>1</td>
<td>3740</td>
<td>4502</td>
<td>2</td>
<td>3034</td>
</tr>
<tr>
<td>1</td>
<td>4033</td>
<td>4749</td>
<td>2</td>
<td>3868</td>
</tr>
<tr>
<td>1</td>
<td>4309</td>
<td>4987</td>
<td>2</td>
<td>5016</td>
</tr>
<tr>
<td>1</td>
<td>4577</td>
<td>5191</td>
<td>2</td>
<td>9611</td>
</tr>
<tr>
<td>1</td>
<td>4840</td>
<td>5382</td>
<td>2</td>
<td>11436</td>
</tr>
<tr>
<td>1</td>
<td>5102</td>
<td>5564</td>
<td>3</td>
<td>2280</td>
</tr>
<tr>
<td>1</td>
<td>5366</td>
<td>5724</td>
<td>3</td>
<td>6260</td>
</tr>
<tr>
<td>1</td>
<td>5633</td>
<td>5882</td>
<td>4</td>
<td>2654</td>
</tr>
<tr>
<td>1</td>
<td>5909</td>
<td>6032</td>
<td>5</td>
<td>2479</td>
</tr>
<tr>
<td>1</td>
<td>6193</td>
<td>6220</td>
<td>6</td>
<td>2792</td>
</tr>
<tr>
<td>1</td>
<td>6491</td>
<td>6418</td>
<td>7</td>
<td>2972</td>
</tr>
<tr>
<td>1</td>
<td>6803</td>
<td>6586</td>
<td>8</td>
<td>3385</td>
</tr>
<tr>
<td>1</td>
<td>7138</td>
<td>6817</td>
<td>9</td>
<td>2479</td>
</tr>
<tr>
<td>1</td>
<td>7497</td>
<td>7129</td>
<td>10</td>
<td>2479</td>
</tr>
<tr>
<td>1</td>
<td>7889</td>
<td>7419</td>
<td>11</td>
<td>2479</td>
</tr>
<tr>
<td>1</td>
<td>8322</td>
<td>7717</td>
<td>12</td>
<td>2479</td>
</tr>
<tr>
<td>1</td>
<td>8813</td>
<td>8003</td>
<td>13</td>
<td>2479</td>
</tr>
<tr>
<td>1</td>
<td>9378</td>
<td>8279</td>
<td>14</td>
<td>2479</td>
</tr>
<tr>
<td>1</td>
<td>10053</td>
<td>8669</td>
<td>15</td>
<td>2479</td>
</tr>
<tr>
<td>1</td>
<td>10898</td>
<td>9478</td>
<td>16</td>
<td>2479</td>
</tr>
<tr>
<td>1</td>
<td>12038</td>
<td>10301</td>
<td>17</td>
<td>2479</td>
</tr>
<tr>
<td>1</td>
<td>13823</td>
<td>11326</td>
<td>18</td>
<td>2479</td>
</tr>
<tr>
<td>1</td>
<td>17671</td>
<td>16707</td>
<td>19</td>
<td>2479</td>
</tr>
</tbody>
</table>
Table 3.2

1970 Permanent Incomes of Model Households

<table>
<thead>
<tr>
<th>Household Type</th>
<th>Permanent Income used by Joint Center</th>
<th>Permanent Income used by U.I.</th>
<th>Household Type</th>
<th>Permanent Income used by Joint Center</th>
<th>Permanent Income used by U.I.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2832</td>
<td>2387</td>
<td>2</td>
<td>1193</td>
<td>1546</td>
</tr>
<tr>
<td>2</td>
<td>4257</td>
<td>3851</td>
<td>2</td>
<td>1768</td>
<td>2044</td>
</tr>
<tr>
<td>1</td>
<td>5063</td>
<td>4862</td>
<td>2</td>
<td>2304</td>
<td>2471</td>
</tr>
<tr>
<td>2</td>
<td>5699</td>
<td>5707</td>
<td>2</td>
<td>2967</td>
<td>2988</td>
</tr>
<tr>
<td>1</td>
<td>6251</td>
<td>6420</td>
<td>2</td>
<td>3666</td>
<td>3861</td>
</tr>
<tr>
<td>2</td>
<td>6756</td>
<td>7020</td>
<td>2</td>
<td>4451</td>
<td>4845</td>
</tr>
<tr>
<td>1</td>
<td>7233</td>
<td>7465</td>
<td>2</td>
<td>5383</td>
<td>5784</td>
</tr>
<tr>
<td>2</td>
<td>7694</td>
<td>7865</td>
<td>2</td>
<td>6565</td>
<td>6652</td>
</tr>
<tr>
<td>1</td>
<td>8144</td>
<td>8206</td>
<td>2</td>
<td>8208</td>
<td>7557</td>
</tr>
<tr>
<td>2</td>
<td>8593</td>
<td>8551</td>
<td>2</td>
<td>10924</td>
<td>8864</td>
</tr>
<tr>
<td>1</td>
<td>9044</td>
<td>8874</td>
<td>2</td>
<td>17360</td>
<td>12668</td>
</tr>
<tr>
<td>2</td>
<td>9502</td>
<td>9192</td>
<td>3</td>
<td>3388</td>
<td>4682</td>
</tr>
<tr>
<td>1</td>
<td>9971</td>
<td>9521</td>
<td>3</td>
<td>10986</td>
<td>9107</td>
</tr>
<tr>
<td>2</td>
<td>10456</td>
<td>9829</td>
<td>4</td>
<td>3611</td>
<td>3866</td>
</tr>
<tr>
<td>1</td>
<td>10962</td>
<td>10221</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>11495</td>
<td>10643</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>12061</td>
<td>11049</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>12670</td>
<td>11449</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>13334</td>
<td>11841</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>14068</td>
<td>12227</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>14894</td>
<td>12611</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>1548</td>
<td>13069</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>16985</td>
<td>14259</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>18403</td>
<td>15568</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>20313</td>
<td>16826</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>23291</td>
<td>18255</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>29673</td>
<td>25138</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
3.4 Value of b To Use For Each "Model" Household in the Modified Utility Function

Once the income and $\alpha$ parameters of each "model" household are specified, it remains to decide which value of "b" to use in the expectation term $A_1$ of eq (3.3) for each model household. It has previously been mentioned that $b$ is related to the income elasticity by the formula $E = \frac{Y}{y + b}$, and that income elasticities across various demographic groups were found to be fairly constant, and approximately equal to .47. In light of these facts, the value of $b$ was chosen for each household so that the implied income elasticity was .47, when evaluated at the permanent income of that household.
Estimation of Other Parameters of Household Behavior

4.1 - Estimation of the \( \gamma \) Parameters

As discussed in the introduction, the other parameters of the utility function, \( \gamma_1, \gamma_2 \) and \( \gamma_3 \), are estimated by running the housing model on 1960 data and choosing that set of \( \gamma \)'s which best reproduces the configuration of the SMSA in 1960. In running the 1960 simulation, the housing stock is taken as fixed. This means the supply parameters take the values \( \beta_1 = 1, \beta_2 = 0 \). To decide on the "best" fit, three criteria are used, which are similar to the criteria used by the Urban Institute. These criteria involve zonal averages of housing prices, household incomes, and racial proportions. The predicted zonal averages are then compared to "base" figures which the Urban Institute has estimated to be the actual values in 1960.

In this calibration effort, a departure was made from the Urban Institute's procedure of estimation. The Urban Institute, sets the price of new housing at 10% above its actual price in order to avoid having households pick new homes in 1960. Although it is clear that such a procedure will help insure that no new dwellings are desired in 1960, (so that all households are content with the existing stock), such a procedure will also raise prices, artificially, for the entire city. It was felt that if 1960 represents a market equilibrium (as is assumed in the model), then for the correct set of parameter values, and at the existing price of new housing, all households should be satisfied when living in the existing 1960 dwellings. If the model indicates that some households desire new housing in 1960, then that result should be regarded as an error, just like the other error measures described above. In calibrating the model, one should choose that parameter set that performs
best with respect to all the error measures. It seems improper to eliminate this "new house" error by artificially raising the price of new housing above its actual value.

Another modification of the Urban Institute's procedure involved the construction of the error in average zonal income. Because, in this modified version of the UI model, households have only a permanent income \( Y_p \), associated with them, it is necessary to translate the Urban Institute's "base" value of zonal current income into "base" values of zonal permanent income, in order to construct a meaningful income error figure. It is not possible to associate both a permanent and current income figure with each household, because as seen in Part 2, it is inappropriate to assume a one-to-one relationship between current income and permanent income. The true relation is definitely stochastic, and cannot be adequately represented as a deterministic one-to-one relationship.

As described in the section of this study dealing with the " \( \alpha \) " estimation, it is quite simple to form \( E(Y_c) - E(Y_p) = \hat{Y}_c \) (1) for each demographic group in Table 1 in Part 2. A weighted average of these \( \hat{Y}_c \) values can then be used to estimate the expected value of transitory income for the population as a whole. This figure for average transitory income is then subtracted from each of the four "base" zonal measures of current income per household (more exact corrections are possible if it is known where the different household types reside). These corrected "base value" figures should represent average zonal permanent income. However, when compared to the average of the estimated permanent income of our 39 households in 1960, these constructed "base" values indicated a $25 higher SMSA-wide monthly permanent income.

(1) Recall that the value of \( E(Y_c) - E(Y_p) \) depends on economic conditions, and equals zero in "normal" economic circumstances.
This $25 discrepancy could arise for two reasons. First, whenever the expected value of \( X \) is used to predict the numerical average value of \( X \), statistical uncertainty can result. (This problem arises in the predictions of average transitory income for the population). Secondly, the Urban Institute's figures are admittedly only estimates which are based on specific assumptions, and could be subject to the usual statistical error.

It was decided that whatever the reason for the $25 discrepancy, it is best to use as "base" values, the zonal permanent income averages that are consistent in the aggregate, with the permanent income distribution that was estimated from the Pittsburgh Census data, and which was used in the model to construct estimates of the permanent income of each household. Hence, each "base" zonal permanent income estimate derived above was corrected for this $25 bias.

For each of the three criteria, zonal income, race and price, the root mean square of the zonal differences from the "base" value is computed. Each of these three errors is divided by the relevant SMSA mean value. So, for example, the income error is divided by mean SMSA permanent income.

Before turning to the specific results, it is useful to first consider which of the three errors deserve the most attention. Since there are only three black households in the model, the racial composition of a neighborhood is heavily influenced by the "discreteness" problem, \(^{(1)}\) and, hence, could be quite sensitive to minor changes in underlying parameters.

\(^{(1)}\) For example, if one black household moves into zone 2, the black population jumps from 0 to 25%.
It appears that errors in income and price should have primary importance. It would also seem that, in the formation of each of the three aggregate errors from the zonal errors, the zonal errors should not simply be arithmetically averaged (as is currently done by the Urban Institute), but should be weighted by the population of each zone. Refinements in the error formation was, however, beyond the scope of this study.

Table (4.1) presents the zonal errors that were generated for various values of \( \gamma_1, \gamma_2 \) and \( \gamma_3 \). In searching over the \( \gamma \)'s, it was found that the model's solution responded strongly to changes in \( \gamma_1 \), moderately to changes in \( \gamma_3 \), and insignificantly to changes in \( \gamma_2 \). Other lessons learned from doing these simulations are discussed later. As can be seen from the results of Table 4.1, the parameter set \( \gamma_1 = .8, \gamma_2 = .3, \gamma_3 = .3 \) dominates all other parameter values.

For comparison purposes, the errors which the Urban Institute found for their 1960 Pittsburgh simulation \(^{(1)}\) are presented on the following page. Because of the different treatments of income in the Urban Institute model, their error results are not perfectly comparable to those of Table (4.1). Nevertheless, a comparison of the two error results can provide one with general ideas about the relative merits of the Urban Institute model and the modified Urban Institute model used in this study.

\(^{(1)}\) DeLeeuw F., and Struyk, R., Urban Institute Housing Model I: Application to Pittsburgh, Pennsylvania. The Urban Institute was in the process of reestimating the model for Pittsburgh, when this report was written. The values above may not represent their "new" best fit.
Table 4.1

Errors for the 1960 Model Simulation

<table>
<thead>
<tr>
<th>Value of Parameters</th>
<th>Income Error</th>
<th>Race Error</th>
<th>Price Error</th>
<th>Average of the three Errors</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \gamma_1 \gamma_2 \gamma_3 )</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.5 0.3 0.7</td>
<td>0.119</td>
<td>0.085</td>
<td>0.165</td>
<td>0.123</td>
</tr>
<tr>
<td>0.7 0.3 0.7</td>
<td>0.119</td>
<td>0.085</td>
<td>0.135</td>
<td>0.113</td>
</tr>
<tr>
<td>0.8 0.3 0.7</td>
<td>0.057</td>
<td>0.085</td>
<td>0.125</td>
<td>0.089</td>
</tr>
<tr>
<td>0.8 0.3 0.4</td>
<td>0.045</td>
<td>0.034</td>
<td>0.133</td>
<td>0.070</td>
</tr>
<tr>
<td>0.8 0.3 0.2</td>
<td>0.045</td>
<td>0.034</td>
<td>0.125</td>
<td>0.068</td>
</tr>
<tr>
<td>0.8 0.7 0.3</td>
<td>0.045</td>
<td>0.034</td>
<td>0.128</td>
<td>0.069</td>
</tr>
<tr>
<td>0.8 3.0 3</td>
<td>0.049</td>
<td>0.034</td>
<td>0.130</td>
<td>0.171</td>
</tr>
<tr>
<td>best fit</td>
<td>0.045</td>
<td>0.034</td>
<td>0.118</td>
<td>0.065</td>
</tr>
</tbody>
</table>
Table 4.2 - Error Comparison

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Income Error</th>
<th>Race Error</th>
<th>Price Error</th>
<th>Average of the Three Errors</th>
</tr>
</thead>
<tbody>
<tr>
<td>U.I. model</td>
<td>.104</td>
<td>.082</td>
<td>.075</td>
<td>.087</td>
</tr>
<tr>
<td>( \gamma_1 = .6 )</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \gamma_2 = .6 )</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \gamma_3 = .3 )</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>modified U.I. model</td>
<td>.045</td>
<td>.034</td>
<td>.118</td>
<td>.065</td>
</tr>
<tr>
<td>( \gamma_1 = .8 )</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \gamma_2 = .3 )</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \gamma_3 = .3 )</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

As Table (4.2) indicates the modified model outperforms the original model in all categories except price. The prices predicted by the modified model always were underestimates of the zonal "base" values of the prices estimated by the Urban Institute.

Recall that in doing the 1960 simulations, the Urban Institute uses a 10% higher price for new housing to insure that no households will choose new housing in 1960 and, hence, that households will be content with the existing housing stock. As mentioned earlier, the use of an inflated price of new housing, does not seem theoretically well justified, though it is easy to envision circumstances where it is the only way to make the 1960 households satisfied with the existing 1960 stock.
Such a situation must prevail if the housing market in 1960 is in equilibrium in 1960. For the reasons discussed previously an inflated price of new housing was not used in our simulations. Fortuitously, even though the actual price of new housing was used in our simulation in 1960, the problems of households desiring new homes never arose.

By using a price for new housing above its actual price, the unmodified U.I. model is able to raise all the prices of all housing, and in this way avoid some of the price underestimation which characterized the modified model's solution. In fact, the modified model with the best parameter set \( V_1 = .8, V_2 = .3, V_3 = .3 \) was run with an inflated price of new housing, the price underestimation problem disappeared, and the price error fell to .099, while the average error fell to .062.

Such results strongly suggest that the lower price error of the unmodified U.I. model was due to their artificial price "markup" in new housing, and indicates that the price error of the modified model could be further reduced by trying different price markups on the actual price of new housing. Such experiments were not performed, because it was felt that the theoretically correct simulation to use was the one involving no price markup.

Although the modified model matches the 1960 Pittsburgh housing market more closely than the unmodified model, it would be premature to pass judgment on the relative merits of the two models. First, the error results between the two models are not strictly comparable for the reasons mentioned above. Second, small shifts in equilibrium assignments can sometimes dramatically change the errors. Finally the results from a single simulation provide too scanty evidence upon which to reach definite conclusions, and can only be regarded as strongly suggestive of the relative merits of the modified model as compared to the original model.
4.2 Lessons Learned from the 1960 Simulations

There were several patterns that emerged from doing these simulations in 1960. First, the errors improved with increasing $\gamma_1$. However, for values of $\gamma_1$ of .9, the algorithm would not converge even when allowed 200 iterations. (The algorithm always converged in less than 100 iterations for all the parameter values displayed in table 4.2.) Based on the relation $PA_1 + A_2 = \gamma_1 Y$, a high value of $\gamma_1$ implies very high expectation levels. The higher the level of $\gamma_1$, the less flexibility a consumer has in his buying decisions, if he is to receive positive utility. The consumer must spend at least $PA_1$ on housing and $A_2$ on other goods to receive positive utility. Therefore, although the convergence problems with high values of $\gamma_1$ are disturbing, such values of $\gamma_1$, seen unrealistically high, and would seriously limit the range of housing that a consumer could choose. Note that a value of $\gamma_1 = .8$, instead of $\gamma_1 = .9$, allows the consumer to have twice the flexibility in his buying decisions.

The equilibrium assignments of households to dwellings seemed insensitive to the values chosen for $\gamma_2$ and $\gamma_3$. Closer study of this insensitivity revealed that when the $\gamma_2$ term (which has to do with neighborhood externalities) is restricted to lie in the $(0, 1)$ range, the importance of neighborhood externalities practically vanishes. In the utility function, the term which measures neighborhood effects ranges between .5 and 1.5. However, the exponent of this term in the utility function is not $\gamma_2$, but rather $.01 \cdot \gamma_2$. For $\gamma_2$ in the range $(0, 1)$, this neighborhood term has practically no effect on a person's utility. A few radically different

(1) See the Appendix to this report for a discussion of other lessons learned from experimenting with the model.
value of \( \gamma_2 \) (i.e., 3, 15, 30) were able to make externalities important, and had the expected effect of raising prices in the wealthier neighborhoods. As Table 4.1 indicates, for the parameter set (.8, 3, .3), the associated average error was .071. Such a result indicates that drastically different values for \( \gamma_2 \) are able to produce quite reasonable results, and should not a priori be excluded from consideration. A more complete examination into the effect of various values of \( \gamma_2 \) on the solution assignment was beyond the scope of this research. Suffice it to say that the restriction, that \( \gamma_2 \) necessarily lie in the range (0, 1) seems improper and implies that neighborhood externalities have at most a minor effect in determining the market equilibrium of the housing market.
Estimation of the Parameters of the Supply Curve

The two parameters of the supply curve, $\beta_1$ and $\beta_2$, are estimated by running model simulations over the 1960-70 decade, and choosing that set of $\beta$'s which "best" fit the SMSA in 1970.\(^1\) As in the estimation of the $\gamma$'s, the "best" fit is determined by minimizing certain errors which represent the discrepancy between predicted and actual 1970 zonal values. The basic criteria used for the $\beta$-estimation are zonal income, housing expenditures, racial composition and the utilization rate of the 1960 housing stock.\(^2\)

As was mentioned in the discussion on the $\gamma$'s estimation, each error criteria should not necessarily receive equal weight. Furthermore, in constructing each of the four aggregate errors from the zonal errors, some weighting scheme, according to the size of each zone, seems appropriate. Although more refined error criteria were not created for the $\gamma$ estimation, in this $\beta$-estimation, the sharp distinction between new housing and existing housing motivated forming additional error criteria other than those just mentioned. The two criteria that were added both involved distinguishing between new and old housing.

First, the predicted number of new "model" houses built in 1970 was reported for each simulation. Since the housing allowance program will hopefully affect vacancy rates is low quality housing, and since vacancies can only occur in the model if there is new construction, it seemed important that the model should be able

\(^1\) In doing these simulations, the $\gamma$ values that are used are the ones which were estimated in Part 4.

\(^2\) These are the same criteria that the Urban Institute uses.
to correctly assess the number of homes left vacant, and the demand for new homes.

As the model is formulated, at least two new homes have to be built in 1970, since there are 41 households and 39 existing dwellings in the 1970 simulation. The "actual" number of new homes built in 1970 is estimated to be four.

Secondly, an additional housing expenditure error was constructed to include only expenditure on the existing stock. It is much more important to be concerned with this housing expenditure error, than with the one which includes expenditure on new housing, since the objective is not to accurately predict the money spent on new housing, but rather to predict the resulting market equilibrium for the entire housing stock, which is dominated by the existing stock.

In ranking the various errors by their importance, the following judgements seemed reasonable. The income, and existing housing expenditure error both stand out as crucially important. Next in importance is the error in the number of new homes built (or equivalent, the number of homes left vacant), and in the utilization rate of the 1960 housing stock. The racial composition error, though reflecting an important phenomenon, should not receive the same weight as the errors, because of the sensitivity of the zonal racial composition, which results from the "discreteness" of the model. Finally, the least important error to focus on is the total housing expenditure error, which includes expenditure on new housing.

Unlike the case of the estimation, in the search over the various values of $\beta_1$ and $\beta_2$, certain relationships between the $\beta$'s could be used to restrict the

(1) In computing the income error, the same corrections, as described in Part 4, were performed to derive a "base" average permanent income for each zone. Unlike the case in Part 4, where a correction for a large bias of $25 was made, in this case, the "base" values had to be corrected for a small bias of only $3.
parameter space that is being searched. Recall that the supply curve is

\[ Q = \left( \beta_{1} + \frac{2}{3} \beta_{2} \frac{(P - Po)}{Pc} \right) \cdot Q_{0} \],

where

Po = operating cost per unit Q,

Pc = construction cost per unit Q,

Q0 = initial quantity of services in 1960, and

Q = services offered in 1970.

Depreciation equals 1 - \( \beta_{1} \), while the slope of the supply curve is related to \( \beta_{2} \). As \( \beta_{1} \) increases ( \( \beta_{2} \) is held constant), the services that existing units can offer increases. Hence, increasing \( \beta_{1} \) makes existing units more competitive with new housing, and should reduce the demand of new housing. Similarly, increasing \( \beta_{2} \) are holding \( \beta_{1} \) constant allows a unit to increase the amount of services offered at any price, and hence should reduce the demand for new housing. Thus, ( \( \beta_{1}, \beta_{2} \) ) can be systematically varied, so as to attempt to keep the number of "model" houses demanded (i.e. the number of vacancies) around the estimated actual number of four.

There is another relation involving \( \beta_{1} \), and \( \beta_{2} \), which is useful, ex-post, in deciding whether a particular set of values for ( \( \beta_{1}, \beta_{2} \) ) is reasonable. It is evident, from eq (5.1) of the supply curve that if \( P = \bar{x} = \left( \frac{1}{2} \left( 1 - \beta_{1} \right) \right) \cdot 1 \cdot Pc + Po \), then \( Q = Q_{0} \). In other words, at a price equal to \( \bar{x} \), a landlord will have no incentive to alter the quantity of services that he offers between 1960 and 1970. A reasonable range for \( \bar{x} \) would be between the price of abandonment (.54) and the price of new construction (1.33). For unchanging demand conditions, \( \bar{x} \) values near .54, indicate that considerable upgrading in the housing stock is likely to occur over the decade, while \( \bar{x} \) values near 1.33 indicate that drastic deterioration is likely to occur.
value of \( \bar{x} \) somewhere between these two extremes appears most reasonable.

The table below presents the results for several of the sets of \((\beta_1, \beta_2)\) which performed reasonably well. In every single case (including those cases not reported), the predicted number of vacancies (or number of new homes demanded) behaved in the manner described above, with high values of \( \beta_1 \) and \( \beta_2 \) discouraging new housing. (Another result, not obvious from the table, is that high values of \( \beta_2 \) tend to be associated with higher prices of housing.) A reassuring fact about the table is that it shows that for every \((\beta_1, \beta_2)\) pair which appeared mildly plausible (in the sense of predicting between 2 to 6 new homes), the implied value of \( x \) always fell within the (1.54, 1.33) range.

Unlike the case of the \( C \) estimation, no one parameter set dominates all the rest. The three main contenders for "best" parameter set are \((\beta_1, \beta_2) = (.9, .3), (.7, .6)\) and \((.8, .5)\). The last parameter set \((.8, .5)\) was rejected, because it had a relatively large error in the important housing expenditure category, (i.e. the one which includes expenditure on the existing 1960 stock only), and it underestimated (though only by one) the number of vacant homes in 1970.\(^{(1)}\)

Based on the results of Table 5.1, it is very difficult to decide between the parameter sets, \((\beta_1, \beta_2) = (.9, .3)\) and \((.7, .6)\). Both parameter sets generated identical equilibrium assignments of households to dwellings. Both parameter sets predict plausible values for \( \bar{x} \), though the higher \( \bar{x} \) value of 1.13 for \((\beta_1, \beta_2) = (.7, .6)\) does seem slightly more reasonable than the \( \bar{x} \) value of .93 for \((\beta_1, \beta_2) = (9, .3)\). The parameter set \((\beta_1, \beta_2) = (.9, .3)\) had a slightly lower error in the

\(^{(1)}\) Recall that these should be two vacant "model" houses in 1970.
### Table 5.1

Errors for the 1970 Model Simulations

<table>
<thead>
<tr>
<th>Value of Parameters</th>
<th>Income Error</th>
<th>Race Error</th>
<th>Utilization Rate Error</th>
<th>Total Housing Expenditure Error</th>
<th>Housing Expenditure not including new housing Error</th>
<th>Average of Errors 1 2 3 4</th>
<th>Average of Errors 1 2 3 5</th>
<th>No. of new &quot;model&quot; homes predicted</th>
<th>Implied χ² value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\hat{\beta}_1$</td>
<td>.4</td>
<td>.10</td>
<td>.17</td>
<td>.08</td>
<td>.04</td>
<td>.33</td>
<td>.11</td>
<td>.16</td>
<td>5</td>
</tr>
<tr>
<td>$\hat{\beta}_2$</td>
<td>.5</td>
<td>.10</td>
<td>.21</td>
<td>.06</td>
<td>.08</td>
<td>.28</td>
<td>.19</td>
<td>.16</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>.6</td>
<td>.10</td>
<td>.21</td>
<td>.08</td>
<td>.04</td>
<td>.32</td>
<td>.09</td>
<td>.16</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>.7</td>
<td>.10</td>
<td>.10</td>
<td>.06</td>
<td>.08</td>
<td>.55</td>
<td>.17</td>
<td>.19</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>.8</td>
<td>.10</td>
<td>.08</td>
<td>.13</td>
<td>.08</td>
<td>.36</td>
<td>.12</td>
<td>.16</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>.8</td>
<td>.10</td>
<td>.26</td>
<td>.11</td>
<td>.08</td>
<td>.29</td>
<td>.08</td>
<td>.19</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>.9</td>
<td>.10</td>
<td>.10</td>
<td>.06</td>
<td>.08</td>
<td>.55</td>
<td>.18</td>
<td>.20</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>.9</td>
<td>.10</td>
<td>.13</td>
<td>.08</td>
<td>.05</td>
<td>.30</td>
<td>.07</td>
<td>.14</td>
<td>4</td>
</tr>
<tr>
<td>best fit</td>
<td>.7</td>
<td>.10</td>
<td>.13</td>
<td>.08</td>
<td>.08</td>
<td>.30</td>
<td>.10</td>
<td>.14</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>.10</td>
<td>.14</td>
<td>.07</td>
<td>.08</td>
<td>.07</td>
<td>.30</td>
<td>.10</td>
<td>.14</td>
<td>4</td>
</tr>
</tbody>
</table>

Implied χ² value = 1.13
important housing expenditure category. However, in view of the sensitivity of the error results to slightly different parameter values, it was felt that Table 5.1 did not provide sufficient evidence to justify choosing either (.9, .3) or (.7, .6) as the "best" set of parameters.

The values $\beta_1 = .9$, $\beta_2 = .3$ imply a very low rate of depreciation, and a steeply sloping supply schedule. One would hope that the values $\beta_1 = .7$, $\beta_2 = .6$ imply more reasonable assumptions about depreciation and supply responsiveness over the decade.

To investigate the robustness of both the (.9, .3) and (.7, .6) parameter sets, the following experiments were performed. The 1970 simulations were run, except now a very low (1%) tax with high exemption levels was imposed such a small tax should have insignificant effects on the housing market. However, the (.9, .3) equilibrium solution, in the case of a 1% tax changed drastically from the no tax 1970 equilibrium solution. The number of new "model" houses dropped to 3 (recall the estimated actual value in 1970 was 4), and the equilibrium assignments of households to dwellings changed considerably from the no tax run. On the other hand, the (.7, .6) equilibrium solution was virtually unaffected by the imposition of the small tax. The number of new homes remained at 4, and the equilibrium assignments of households to dwellings were almost identical to those in the no-tax equilibrium solution, (two households swapped houses). Of course, the quantity of housing was slightly lowered due to the 1% reduction in income.

It seems extremely unlikely that a 1% tax should drastically change housing patterns, and reduce the amount of new housing by 25%. In light of the results of this experiment, and in light of the previous feelings that $\beta_1 = .7$, $\beta_2 = .6$
represented more reasonable supply behavior than \( \hat{\beta}_1 = .9, \hat{\beta}_2 = .3 \), the
parameter set (.7, .6) appears to be the "best" one. This parameter set will be
used in the simulations of various housing allowance policies. Because the effects
of a housing allowance will depend crucially on the slope of the supply curve,
some policy simulations with the parameter set (.9, .3) will also be performed so
as to obtain a better feeling of the range of effects that can be expected under
various assumptions of supply behavior.

A comparison of the 1970 errors, reported in Table 5.1, with the 1970 errors
of the Urban Institute, as reported in their Pittsburgh study, \(^{(1)}\) is not performed
for two reasons. First, the Urban Institute does not use the more refined error
categories that were used in this study. (It should be mentioned, though that
using errors for income, race, total housing expenditure, and utilization of 1960
stock, the average error of the Urban Institute's best fits about .096. - which is
quite low). Secondly, it is our understanding that in the Urban Institute's recent
reestimation of the Pittsburgh model, that the "best" supply parameters were changed
considerably from those reported in the Pittsburgh study, and are now in the range
\( \beta_1 = 5, \hat{\beta}_2 = .6 \) (instead of either \( \beta_1 = .9, \hat{\beta}_2 = .3 \)
\( \beta_1 = .7, \hat{\beta}_2 = .3 \) as reported in the Pitts-
burough study) These values for \( \beta \) are quite different from the ones which we estimated,
and imply a very high \( R \) value of 1.53. A discussion of this difference between
our \( \beta \) values and those of the Urban Institute is contained in Part 7.

\(^{(1)}\) DeLeeuw, Ew, and Stryk R., Urban Institute Housing Model I: Application
to Pittsburgh, Pennsylvania.
Market Simulations of Various Housing Allowance Programs

Once all the parameters of the model have been estimated, the effect of various housing allowance payment formulas can be examined. The model can provide information about the relative merits of various formulations of housing allowances. In particular, the model can indicate the price inflation and housing improvement that may occur as a result of the allowances. The effects on recipients and non-recipients can also be examined.

Although there are many possible income and rent conditioned hybrid forms of allowance formulas, we decided to analyze the "pure" forms of the payment formulas so as to highlight their differential effects. Basically, two types of housing allowance programs are investigated in this study. These are the income gap, and percent of rent formulas. For each of these basic programs, the effects of the magnitude of the allowance, of eligibility requirements, and of earmarking are examined.

Although the model can predict the results of the interaction of supply and demand in the housing market, the model is not well suited to answer detailed questions on racial and locational shifts, caused by the allowance. The small number of black "model" households (only three "model" households), together with the "discreteness" problem, make any answers to these questions very uncertain. Recall from the introduction that the "discreteness" problem arises because each "model" household represents about 14,000 people. Slight changes in the equilibrium assignments of "model" households to "model" dwellings can significantly alter the locational implications of a particular equilibrium assignment. For example, as mentioned in Part A, if a black "model"
household moves into zone 2, then the black percentage in that zone jumps from 0 to 25%. Because of the usual statistical uncertainty associated with any equilibrium assignment, it would be unwise to use the model to confidently predict detailed locational effects. Furthermore, the model is not intended to address questions regarding the administrative complexities of the allowance program, possible incentives for fraud, and altered work incentives.

In this Part 6 of the study, the exact form of the allowances that are used in the policy simulations are described. The specific results of each simulation are presented. In Part 8 of this study, the results presented in this part will be examined and compared in detail. Based on this detailed comparison, policy recommendations on the form of the housing allowances will be made.

For comparison purposes, each policy is run under two different assumptions about income elasticity. The first assumption implies that the housing allowance will be treated as ordinary income, and hence the income elasticity of .47, that other Joint Center research has found, should apply. On the other hand, the second assumption assumes that recipients treat housing allowance income differently than ordinary income, and that recipients will behave with an income elasticity of unity to this housing allowance income. It would seem most reasonable to assume that housing allowance income will not be treated differently by recipients, so that the correct elasticity to use is .47. The results of the simulations with the assumption of unitary elasticity will indicate how different the conclusions about policy recommendations can be, if the elasticity is taken to be one, instead of .47. Since other housing allowance studies - notably those of the Urban Institute assume a unitary income elasticity, this comparison between policy
recommendations should provide some idea of the sensitivity of their recommendations
to their assumption of a unitary income elasticity.

6.1 - Simulating Housing Allowances

Before the results of the simulations can be described, we must first discuss three
questions: 1) How is the effect of the allowances on permanent income computed?
2) How do expectations change? and 3) How can the assumption of an income
elasticity of one out of the income received from the housing allowance be incorporated
into our modified housing model?

By adding to the purchasing power of a recipient for a presumably long period of
time, any housing allowance program will raise the permanent income of the consumer.
In order to predict the effect of a housing allowance on housing expenditures, it is
necessary to predict the effect of the housing allowance on permanent, not current, income,
since it is generally accepted that permanent income represents the measure of income
on which consumers base their housing expenditures. In the case of a percent of rent
program, the computation of the new level of permanent income is straightforward.
If \( F(R) \) represents the allowance a person receives if he lives in a house with rent \( R \),
then that person's new permanent income becomes

\[
Y_p = Y_p \text{ old} + F(R)
\]

Since \( R \) depends on permanent, not current income, the rent, \( (R) \), and hence
the allowance \( F(R) \), is unchanged, when permanent income remains constant, but
current income fluctuates. (Recall that permanent income is the expected value of
current income). The fact that current income will fluctuate around \( Y_p \), will not alter
the allowance a person receives, since the allowance depends on rent, which will not
change unless permanent (not current) income does. Hence, for an allowance program
that depends only on rent, eq. (6.1) gives the exact expression for the new permanent
income of the recipients.

On the other hand, suppose that some rent based formula is used for an allowance program, but that eligibility for this rent based program is based on current income. Then, fluctuations in current income around $Y_p$ will have an effect on the allowance received, because as current income fluctuates above some cutoff eligibility income, the allowance would no longer be $F(R)$, but rather would immediately drop to zero. In such a case, the discussion below becomes relevant.

In the case of an income gap program, the problem can arise that fluctuations in $Y_c$ will affect the amount of the allowance received. The basic form of an income gap allowance is

$$\text{(6.2) \ Allowance} = C^* - b \ Y_c, \ Y_c \leq Y_{\text{max}}$$

where $C^*$ and $b$ are parameters of the program, and $Y_c$ is a measure of current income.

The value for $Y_{\text{max}}$ (which is the income eligibility cut off) is usually $C^*/b$. In any particular housing allowance policy, the allowance that actually will be paid in any one year will probably be based on the current income of that year. The question still arises, though, as to whether, in the 1970 simulation, the actual allowance (based on actual $Y_c$) paid in 1970 should be used, or whether the expected allowance (based on $E(Y_c)$ paid in 1970 should be used.

Since it is generally accepted that people base their housing consumption on

---

1. This is not the only option. It is quite possible to base the allowance on some average of past year's income. This average would attempt to measure permanent income.
permanent income, it is necessary to compute the new value of permanent income, given that there is a housing allowance. By definition, \( Y_p = E(Y_c) \), \(^1\) in "normal" economic conditions. Once the allowance program is instituted, the individual calculates his new permanent income as

\[
(6.3) \quad Y_p = E(Y_c) + E(\text{Allowance}).
\]

Eq. (6.3) shows that it is the expected allowance that will enter the individual's calculation of his new permanent income, upon which he bases his housing decisions.

Suppose in eq. (6.2), that \( Y_{\text{max}} \) is very large (so that everyone is eligible). then

\[
(6.4) \quad E(\text{Allowance}) = E(C^* - bY_c) = C^* - bE(Y_c).
\]

In other words, the allowance that the person expects to receive is a simple function of the income that a person expects to receive in any "normal year," namely \( E(Y_c) = Y_p \).

Now, \( Y_{\text{max}} \) (i.e. the income eligibility cutoff) is usually \textit{not} very large; (it usually equals \( C^*/b \)). This creates a discontinuity problem at \( Y_c = Y_{\text{max}} \), in that a person whose income, \( Y_c \), fluctuates around \( Y_{\text{max}} \), will jump on and off the allowance. (A mitigating factor is that if a person's income is near \( Y_{\text{max}} \), work disincentives exist to encourage a worker to keep \( Y_c \) below \( Y_{\text{max}} \).) Using eq. (6.4) to predict the expected allowance of this person, will \textit{understate} the amount that the person expects to receive from the allowance.

---

1. The "E" stands for expected value.

2. This discontinuity problem is sometimes referred to as the "notch effect."
To understand why this underestimation occurs, note that the exact expression for the expected value of the allowance is

\[(6.5)\] \[E (\text{Allowance}) = E (\text{Allowance/person is eligible}) \cdot Pr (\text{eligibility}) + E (\text{Allowance/person is non-eligible}) \cdot Pr (\text{non-eligibility}).\]

Eligibility is determined according to whether the person's current income is above or below \(Y_{\text{max}} = C^*/b\). In using eq. (6.4) to compute the expected allowance, it is implicitly being assumed that the expression for the allowance received by a non-eligible (i.e., the term \(E (\text{Allowance/person is non-eligible})\) in eq. (6.5)) is \(C^* - b \cdot Y_c\), when, in reality, it is zero, since a non-eligible person receives a zero allowance.

Since the value of the term \(C^* - b \cdot Y_c\) is negative for \(Y_c > C^*/b\) (i.e., for non-elitigibles), using eq. (6.4) (where the allowance for non-elitigibles is incorrectly assumed to be negative), instead of eq. (6.5) (with the allowance for non-elitigibles correctly set to zero) will understate the amount of the allowance that a person can expect to receive.

This understatement will not be a serious problem when dealing with income gap programs, since 1) we are most interested in seeing what happens to the poorest people on the allowance, 2) the magnitude of the allowance is smallest around \(Y = Y_{\text{max}} = C^*/b\), and 3) it does not appear that the behavior of the "model" households with income \(Y_c\) near \(Y_{\text{max}}\) are strongly affected by the allowance. (These last two reasons do not apply to the previously discussed per cent of rent formulas when eligibility is determined by current income. There, the allowance is largest for those who pay the highest rent. In our simulations with per cent of rent formulas, we chose, beforehand, the households which were eligible for the program. To facilitate comparison between

\(^{1}\) "Pr" stands for "probability of".
income-based and rent-based programs, the households who were eligible for the
income-based programs were the ones chosen to be eligible for the rent-based programs.)

One way around this discontinuity problem (or "notch effect") would be to
use the probability distribution of \( Y_c \) for each "model" household. One could then
compute the exact expression \( E(\text{Allowance}, Y_c \leq Y_{\text{max}}) \).
\( P_r(Y_c \leq Y_{\text{max}}) \).
(Note that this expression is just eq. (6.5), with the second term on the right hand
side of eq. (6.5) set to zero.) Although exact, this correction would have involved
computing partial integrals of the distribution of \( Y_c / Y_p \). (Recall that the distribution of
\( Y_c / Y_p \) was derived in Part 2. This correction was not made because we felt that the small
underestimation bias introduced by using eq. (6.4), to compute \( E(\text{Allowance}) \) for
income gap formulas was well within the statistical confidence regions of the estimates
of the permanent income, \( Y_p \), of each model household.

As \( Y_p \) is increased by any form of the housing allowance, then the "minimum
expected expenditure level," \( Y_p \), is automatically increased. This stands
to reason since increasing \( Y_p \) will increase aspiration levels. As the "minimum"
expectations for both housing and non-housing goods rise, the minimum amount of each good
that the person must receive to obtain positive utility, also rises.

Finally, as mentioned previously, several studies have found that income
elasticities are well below 1.0, and fairly constant across different demographic groups.
Other Joint Center research have found this elasticity to be around .47. Recall that
the formula for the income elasticity was found, in Part 3, to be \( E = \frac{Y}{Y + b} \), where the
value of \( b \) was chosen so that \( E = .47 \) for each "model" household. (Recall from Part 3,

I. The results of Part 2 could be used to do this.
that "b" appears as a constant term in the expression for "minimum" expectations for housing. See Part 3 for a more detailed discussion of "b"). If the value of b is kept fixed, then as $Y_p$ is increased by the housing allowance, the income elasticity will stay well below 1.0, though it will rise above .47. (The fact that the income elasticity cannot be held constant at .47 as income increases arises from the particular functional form of the utility function. In any case, the largest value E ever achieves under the housing allowance programs considered here, is .60). Holding b fixed corresponds to the assumption that added income received from the housing allowance program will be treated like a permanent increase in ordinary income.

The value of "b" can be adjusted to incorporate the assumption that the income elasticity out of income received from the housing allowance is one, but that the income elasticity out of regular (permanent) income is .47. If $Y_{p, new}$ is the new permanent income of a recipient after the housing allowance, then if b is chosen so that

$$\frac{Y_{p, new}}{Y_{p, new} + b} = .47$$

then the implication is that at the new permanent income, $Y_{p, new}$, the income elasticity out of ordinary (permanent) income is .47. However, since the value of b is being changed only when income from the housing allowance is received, it follows from eq. (6.6) that for any housing allowance which alters $Y_{p, new}$, that

$$\frac{(Y_{p, new} + b)}{Y_{p, new}} = \frac{2.13 Y_{p, new}}{.47}$$

Going back to the expression for $A_T$ in Part 3, (recall, from Part 3, that $A_T$ represents "minimum" expectations for housing), eq. (6.7) implies that expectations are proportional
to $Y_{p_{\text{new}}}$. Based on the discussion in Part 3, we know that whenever expectations are proportional to $Y_{p_{\text{new}}}$ (which changes only when income is received from the housing allowance program), then the income elasticity out of income received from a housing allowance program will be unity. Hence, by adjusting "b" so that eq. (6.7) is always satisfied, each consumer will have a unitary income elasticity out of income received from the housing allowance, but an income elasticity of .47 out of other increases in income. Let us now consider the specific housing allowance programs that were simulated.

6.2 Policy Simulations

The specific values for the parameters of the various housing allowance programs were chosen so as to illustrate a range of effects for programs that varied with respect to eligibility, cost level and earmarking levels. In general, programs were chosen so that either 22 percent or 15 percent of the population was eligible. In terms of the model, this meant that either 9 or 6 out of the 41 model households were eligible for particular allowance programs.

Several criteria are used to summarize the effect of each program. Net cost,
average price, average quantity, average price faced by those eligible, average quantity of those eligible are all reported. Two additional measures are also reported. It turned out that the recipients of all allowance programs always were among the "bottom" 14 households when ranked in terms of their demand for housing. Furthermore,

1. All programs include a 1% income tax, for financing purposes. The exemption levels are $2,550 for multiple families, and $980 for elderly and single families. Households who are eligible for a housing allowance program do not pay any financing tax.
2. Their "demand for housing" is measured by the size of their preferred new house.
various allowance programs affected the "bottom" 7 in significantly different ways. (These "bottom" 7 are the ones in the worst housing). Therefore, the average quantity of the housing consumed by the "bottom" 7 and 14 are reported. (Henceforth, the bottom 7 and 14 will be referred to as the bottom 1/6 and 1/3 of the population.)

The last summary measure that is reported has to do with "expenditure ratios." An expenditure ratio (abbreviated as ER) is defined as the total increase in rent paid by recipients divided by the total allowance. It is a measure of the fraction of the allowance that was spent on additional housing. Although a high ER is desirable, it does not necessarily indicate how good a program is. Increases in the quantity \( Q \) of housing consumed should measure the success of a program. High price inflation could cause ER to be high, while the actual increase in \( Q \) could be quite low.

6.3 Income Gap Programs

The basic form of an income gap allowance is

\[
\text{Allowance} = C^* - b \, Y, \quad Y \leq C^*/b
\]

The parameter \( b^2 \) is taken to be the same for each household type, while the parameter \( C^* \) is the same for household types 1 and 3, and for types 2 and 4. These different \( C^* \) values for the four household types reflect the varying needs of single/elderly families, and those of multiple families. Furthermore, certain programs require that a person is eligible for a housing allowance only if he lives in a house of minimum acceptable quality.

1. An "expenditure ratio" is sometimes referred to as the "earmarking ratio." Earmarking, in this study, will be used to refer to the minimum standard housing that a person must consume if he is to receive the allowance.
2. Do not confuse the "b" of the allowance program with the "b" (discussed in Part 3) which effects expectations. The context should make it obvious which "b" is being discussed.
3. Recall that in the model, there are four household types. Type 1 is white nonelderly multiple households. Type 2 is white single and elderly. Type 3 is black nonelderly multiple households. Type 4 is black single and elderly.
These quality levels, like the $C^*$ parameters, are the same for household types 1 and 3, and 2 and 4. These minimum quality levels are referred to as the earmarking requirements. The basic programs that were considered are reported in Table 6.1. In that table, $C^*$ and "earmarking requirements" are completely described by a pair $(a, b)$, where $a$ refers to household types 1 and 3, while $b$ refers to types 2 and 4.
<table>
<thead>
<tr>
<th>Program</th>
<th>Abbreviation</th>
<th>Parameters</th>
<th>Earmarking Requirements</th>
<th>No. of Model household eligible</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>$C^*$</td>
<td>Minimum $Q$</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(multiple single)</td>
<td>(multiple single)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(families, elderly)</td>
<td>(families, elderly)</td>
<td></td>
</tr>
<tr>
<td>High payment, no earmarking</td>
<td>(H, O)</td>
<td>125,100</td>
<td>.3</td>
<td>None</td>
</tr>
<tr>
<td>Medium payment, no earmarking</td>
<td>(M, O)</td>
<td>(82.7, 66.4)</td>
<td>.2</td>
<td>None</td>
</tr>
<tr>
<td>Medium payment, medium earmarking</td>
<td>(M, M)</td>
<td>(82.7, 66.4)</td>
<td>.2</td>
<td>(53, 44)</td>
</tr>
<tr>
<td>Medium payment, only 6 eligible, no earmarking</td>
<td>(M*, O)</td>
<td>(135, 108)</td>
<td>.4</td>
<td>None</td>
</tr>
<tr>
<td>Low payment, no earmarking</td>
<td>(L, O)</td>
<td>(62, 50)</td>
<td>.2</td>
<td>None</td>
</tr>
<tr>
<td>Low payment, low earmarking</td>
<td>(L, L)</td>
<td>(62, 50)</td>
<td>.2</td>
<td>(50, 41)</td>
</tr>
</tbody>
</table>

These payment formulas were chosen so as to illustrate the effects of increased payments, number eligible, and earmarking on the resulting housing market equilibrium. For example, plans (M, O) and (M*, O) spend about the same amount of money, but simply distribute it to different percentages of the population. Several earmarking levels were tried to see which ones worked best. The ones which worked best are reported in the table above.
The diagram below graphically illustrates the allowances that types 1 and 3 households can receive under the various plans. (Similar diagrams hold for types 2 and 4).
Table 6.2 presents the results of the simulation of each of the allowance programs of Table 6.1 under the assumption of a .47 income elasticity. For comparison purposes, the "base" 1970 run is reported on the last line of Table 6.2. This "base" run is the one from Part 5 which best fit the Pittsburgh housing market in 1970. Note that with plans (M, M) and (L, L), not all those eligible for the housing allowance program choose to participate. Some poor people will choose not to participate in a housing allowance program with earmarking, because they are not willing to sacrifice their non-housing consumption in order to meet the minimum housing requirements of the allowance program. In each case, it it those households with the lowest housing consumption who choose not to participate. The figures of Table 6.2 will be examined more closely in Part 8.
### Table 6.2 - Housing Allowance Simulations for Income Gap Formulas
Under the Assumption of an Income Elasticity of .47

<table>
<thead>
<tr>
<th>Income Plan Abbreviation</th>
<th>Price Per Unit of Housing Service</th>
<th>Quantity of Housing Services</th>
<th>Total Allowance Paid To Eligible &quot;Model&quot; Households</th>
<th>No. of &quot;Model&quot; Households Eligible for Allowance</th>
<th>No. of &quot;Model&quot; Households Who Accept Allowance</th>
</tr>
</thead>
<tbody>
<tr>
<td>H, O</td>
<td>$1.22 $1.14 $1.10</td>
<td>114.44</td>
<td>56.30 44.87 52.85 .22</td>
<td>$325.44</td>
<td>9</td>
</tr>
<tr>
<td>M, O</td>
<td>1.21 1.08 1.06</td>
<td>114.13</td>
<td>55.44 43.55 51.52 .21</td>
<td>213.48</td>
<td>9</td>
</tr>
<tr>
<td>M, M</td>
<td>1.23 1.13 1.11</td>
<td>114.49</td>
<td>56.69 45.56 53.20 .47</td>
<td>119.70</td>
<td>9</td>
</tr>
<tr>
<td>M*, O</td>
<td>1.21 1.09 1.05</td>
<td>114.12</td>
<td>55.47 44.17 45.42 .20</td>
<td>220.32</td>
<td>6</td>
</tr>
<tr>
<td>L, O</td>
<td>1.20 1.05 .98</td>
<td>113.77</td>
<td>54.54 42.52 43.78 .19</td>
<td>83.16</td>
<td>6</td>
</tr>
<tr>
<td>L, L</td>
<td>1.20 1.06 1.00</td>
<td>113.86</td>
<td>54.82 43.13 44.75 .53</td>
<td>38.24</td>
<td>6</td>
</tr>
<tr>
<td>base</td>
<td>1.19 1.03 -</td>
<td>113.85</td>
<td>54.19 41.56 -</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Notation:
- $\bar{P}$ = average price per unit of housing service for the entire market.
- $\bar{P}_{1/3}$ = average price per unit for the bottom 1/3 of the population.
- $\bar{P}_{1/6}$ = average price per unit for the bottom 1/6 of the population.
- $\bar{P}_{\text{eligible}}$ = average price per unit for those who are eligible for the allowance (even those who do not take the allowance).

Similar interpretations for $\bar{Q}$, $\bar{Q}_{1/3}$, $\bar{Q}_{1/6}$, $\bar{Q}_{\text{eligible}}$.

ER = expenditure ratio

Total Allowance Paid = total allowance paid to the "model" households who take the allowance.
Table 6.3 - Housing Allowance Simulations for Income Gap Formulas under the Assumption of a Unitary Income Elasticity

<table>
<thead>
<tr>
<th>Income Plan Abbreviation</th>
<th>Price Per Unit of Housing Service</th>
<th>Quantity of Housing Services</th>
<th>Total Allowance Paid to Eligible &quot;Model&quot; Households</th>
<th>No. of &quot;Model&quot; Households</th>
<th>No. of &quot;Model&quot; Households Eligible for Allowance Who Accept Allowance</th>
</tr>
</thead>
<tbody>
<tr>
<td>H2O</td>
<td>$1.28</td>
<td>$1.30</td>
<td>115.65</td>
<td>60.06</td>
<td>50.60</td>
</tr>
<tr>
<td>M1,O</td>
<td>1.24</td>
<td>114.82</td>
<td>57.7</td>
<td>47.15</td>
<td>55.24</td>
</tr>
<tr>
<td>M2,M</td>
<td>1.28</td>
<td>115.73</td>
<td>60.27</td>
<td>51.12</td>
<td>58.16</td>
</tr>
<tr>
<td>M2,O</td>
<td>1.26</td>
<td>115.26</td>
<td>58.57</td>
<td>49.59</td>
<td>51.40</td>
</tr>
<tr>
<td>L1,O</td>
<td>1.21</td>
<td>113.86</td>
<td>55.12</td>
<td>43.86</td>
<td>45.04</td>
</tr>
<tr>
<td>L2,L</td>
<td>1.22</td>
<td>114.2</td>
<td>56.03</td>
<td>45.53</td>
<td>46.78</td>
</tr>
</tbody>
</table>

* See Table 6.2 for an explanation of the notation.
Table 6.3 presents results of the simulations for the same income allowance programs, under the assumption that the income elasticity out of housing allowance income equals one. As a quick comparison shows, the tables differ markedly in their predicted effects of an income allowance plan. As expected, Table 6.3 suggests much larger effects on housing consumption than does Table 6.2. The nature and implications of the differences between Tables 6.2 and 6.3 will be examined more closely in Part 8.

6.4 Percent of Rent Programs

The basic form of percent of rent programs is Allowance = \( a \) Rent, where \( a \) is a parameter that could vary according to the household type and income of each household. The parameter \( a \) was chosen to be the same for all those eligible for the allowance program. Unlike the income gap programs, percent of rent programs do not give the largest allowances to the poorest members of the population. In order to facilitate comparison with the income gap simulations, the same 9 and 6 households who were eligible for the income gap programs were chosen to be eligible for the percent of rent programs. Table 6.4 lists the percent of rent programs that were successfully simulated.

<table>
<thead>
<tr>
<th>Program</th>
<th>Abbreviation</th>
<th>Value of ( a )</th>
<th>Earmarking Requirements</th>
</tr>
</thead>
<tbody>
<tr>
<td>30% of Rent no earmarking</td>
<td>(30, O)</td>
<td>30%</td>
<td>None</td>
</tr>
<tr>
<td>30% of Rent low earmarking</td>
<td>(30, L)</td>
<td>30%</td>
<td>(50, 41)</td>
</tr>
<tr>
<td>40% of Rent no earmarking</td>
<td>(40, O)</td>
<td>40%</td>
<td>None</td>
</tr>
</tbody>
</table>

Table 6.4 - Percent of Rent Programs

<table>
<thead>
<tr>
<th>Program</th>
<th>Abbreviation</th>
<th>Value of ( a )</th>
<th>Earmarking Requirements</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Minimum Q</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(multiple, single)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(families, elderly)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>No. of Model Households Eligible</th>
</tr>
</thead>
<tbody>
<tr>
<td>9</td>
</tr>
<tr>
<td>6</td>
</tr>
<tr>
<td>9</td>
</tr>
</tbody>
</table>
It should be mentioned that several additional percent of rent programs were tried, but in each case the algorithm failed to converge in 200 iterations. Tables 6.5 and 6.6 present the results of the simulations of the percent of rent programs, under the assumption of an income elasticity of .47, and under the assumption of a unitary income elasticity out of income received from the housing allowance.
Table 6.5 - Housing Allowance Simulations for Percent of Rent Formulas under the Assumption of an Income Elasticity of 0.47

<table>
<thead>
<tr>
<th>Rent Program Abbreviation</th>
<th>Price Per Unit of Housing Service</th>
<th>Quantity of Housing Services</th>
<th>Total Allowance Paid to Eligible &quot;Model&quot; Households</th>
<th>No. of &quot;Model&quot; Households Eligible for Allowance</th>
<th>No. of &quot;Model&quot; Households Who Accept Allowance</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$1.20</td>
<td>$1.05</td>
<td>$1.01</td>
<td>114.03</td>
<td>55.02</td>
</tr>
<tr>
<td>(30, O)</td>
<td>1.20</td>
<td>1.06</td>
<td>0.97</td>
<td>113.92</td>
<td>55.04</td>
</tr>
<tr>
<td>(40, O)</td>
<td>1.21</td>
<td>1.07</td>
<td>1.03</td>
<td>114.25</td>
<td>55.35</td>
</tr>
</tbody>
</table>

Table 6.6 - Housing Allowance Simulations for Percent of Rent Formulas under the Assumption of a Unitary Income Elasticity

<table>
<thead>
<tr>
<th>Rent Program Abbreviation</th>
<th>Price Per Unit of Housing Service</th>
<th>Quantity of Housing Services</th>
<th>Total Allowance Paid to Eligible &quot;Model&quot; Households</th>
<th>No. of &quot;Model&quot; Households Eligible for Allowance</th>
<th>No. of &quot;Model&quot; Households Who Accept Allowance</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$1.23</td>
<td>$1.12</td>
<td>$1.07</td>
<td>114.82</td>
<td>56.99</td>
</tr>
<tr>
<td>(30, O)</td>
<td>1.23</td>
<td>1.12</td>
<td>1.02</td>
<td>114.70</td>
<td>56.88</td>
</tr>
<tr>
<td>(40, O)</td>
<td>1.23</td>
<td>1.13</td>
<td>1.08</td>
<td>114.97</td>
<td>57.67</td>
</tr>
</tbody>
</table>

* See Table 6.2 for an explanation of notation.
6.5 Simulations under Different Supply Assumptions

Recall that in Part 5, the supply parameter sets (.9, .3) and (.7, .6) performed equally well in terms of error measures. The (.7, .6) combination was chosen because of its robustness and more plausible implications for supply behavior. In order to see how much the results of the policy simulations would change under different supply assumptions, the income gap policies were run under the supply parameter set (.9, .3).

As could be expected with the steeper supply curves represented by the parameter set (.9, .3), each housing allowance program now created much more inflation, than with the (.7, .6) parameter set. Yet the increase in the quantity of housing consumed was less than in the (.7, .6) case. In general, each program looked much worse with the (.9, .3) parameter set.

Some of the results of the policy simulations with the (.9, .3) parameter set confirm the implausibility of (.9, .3) as a reasonable parameter set. In all the income gap policy simulations made under the assumption of a .47 income elasticity, the quantity of housing consumed by the bottom 1/6, 1/3, and by the entire population actually fell from the no allowance situation, even under the most expensive allowance plan. The idea that a 1 percent financing tax could so overwhelm the effects of a generous income transfer program at every point in the housing market is extremely unlikely, and provide further evidence on the inappropriateness of the parameter set (.9, .3). Such results strongly reinforce the choice of the supply parameter set (.7, .6),
instead of (.9, .3), as the one which should be used to best simulate the effects of
various housing allowance policies in Pittsburgh. The policy comparisons and
recommendations in Part 8 will be based on the results of the simulations with the
supply parameter set (.7, .6).
Part 7: The Appropriateness of the Modified Model

The two primary changes in the modified model involved the reduced income elasticity and the improved empirical technique for estimating the permanent income distribution of model households. Having calibrated the modified model for Pittsburgh and simulated several housing allowance payment formulas, we can review the results and reflect on the appropriateness of the modified model.

The permanent income modification amounted to an empirical change in the model input. It avoided the common identification problems associated with estimating permanent income from cross-sectional data and yet it replaced the ad hoc income adjustment in the original model with theoretically derived income distributions. As a result of making this change in purchasing power more explicit, the effect of reducing the income elasticity could be examined more confidently.

The technique for estimating permanent income appeared to work quite well, producing realistic changes in the means and variances of permanent income distributions for various households, and yielding model incomes that resulted in plausible model calibrations and low calibration errors.

The use of low income elasticities was motivated by our empirical findings -- see Chapter 4. It is important to note, then, that the experience with the modified model supports the low income elasticity estimates. The 1970 Pittsburgh calibrations yielded beta values that had more plausible market interpretations. Recall from Part 5 that the supply equation given in (1.1) implies that a house offering $Q_o$ units of service in 1960 would also offer $Q_o$ units at the end of the decade if the 1970 price $\bar{x}$ satisfied
\[ 1 - \beta_1 = \frac{2}{3} \beta_2 \left( \frac{\bar{x} - P_0}{P_0} \right). \]  

For the \( \beta_1 = 0.7, \beta_2 = 0.5 \) case, \( \bar{x} = 1.13 \), a price about 15 percent below the new price. Such a result is quite reasonable. Refering to Figure 8.1, it implies that most of Pittsburgh's housing stocks improved slightly during the decade but the poorest third of the market was allowed to deteriorate to a level of \( Q \) significantly below that in 1960.

Indeed every one of the tested sets of beta values in Table 5.1 that yielded 3, 4, or 5 new houses satisfied (7.1) for \( \bar{x} \) values between 0.95 and 1.25. The modified model had little difficulty in predicting the right amount of new housing (four units) using parameters with realistic supply side market interpretations. Those sets of values for beta which well matched the behavior of existing houses also predicted the right number of new houses.

On the other hand, the Urban Institute calibrations of the unmodified model tends to yield beta values which imply \( \bar{x} \) levels much higher than the new house price. Early Pittsburgh calibrations with (1.9, 1.1) and (1.7, 1.3) beta values implied \( \bar{x} \) values of 1.72. Recent Pittsburgh calibrations with (1.5, 1.6) values imply an \( \bar{x} = 1.53 \). The UI results for other cities also yield betas with implied \( \bar{x} \) values well above the new house price -- a result which seems quite implausible.

To insure that these differences in the type of beta values selected arose because of the income elasticity assumptions and not the different permanent income estimates used, the modified model was run using 1970 base conditions and beta values of (1.7, 1.5).
with income elasticities equal to unity (that is, the \( b \) term in the expectation terms was set to zero so that \( E_1 = Y / (y + b) = 1 \)). As expected, the model solution indicated only two new houses and no abandonments. This result indicates that it is more difficult to get people in new housing when \( E_1 = 1 \) if \( \bar{x} \) values are kept below the new house price. Other parameter values could certainly produce better fits. However, UI gamma values were similar to ours and, to increase the predicted amount of new housing by changing the supply side, one would have to alter the beta values in ways that increased the implied \( \bar{x} \) significantly and, we feel, unrealistically above \( P_N \).

To be sure, we studied the modified model only for one city. Perhaps in other cities the results would turn out differently. However, the reasonableness of this model calibration, especially in comparison to the Urban Institute experience, was additional evidence in support of the \( E_1 = 0.47 \) assumption.

Lowering the \( E_1 \) value appears to affect the calibration process more than one might expect. After all, the elasticity affects one’s behavioral response to increments in income. Clearly, misspecification of the income elasticity significantly alters one’s picture of where one begins as well as affecting the use of additional income. Even if housing allowances were treated by individuals as if their income elasticity were one (as in some of the cases discussed in Section 8.1), one would have to carefully specify the behavior that produced the base conditions in the market.

In other words, even if \( E_1 = 1 \) for housing allowance payments, as long as \( E_1 = 0.47 \) for other income, calibrating the model using the \( E_1 = 1 \) assumption will yield unrealistic parameter estimates.
Several other minor modifications were also made in selecting empirical inputs and criteria for model fits and in running the model algorithm. They were intended to improve the calibration and running of the modified model. Minor modifications not involving the solution algorithm have been discussed in other parts of the chapter.

A more extensive look at the solution algorithm is contained in the Technical Appendix. We felt that alternative solution techniques (still in the development stages) might greatly facilitate the sensitivity analysis of the model, enable one to have more confidence in the model's solutions, and permit one to avoid the "discreteness" problem of the current model thereby greatly enhancing its value. The Technical Appendix reviews our thoughts and experiences along these lines.

One final note on modifications. Our remarks should not be regarded as critical of the basic UI model. Similar changes would be difficult to accomplish using other models. Indeed, it is a credit to the model that it may be readily modified and/or disaggregated to accommodate alternative assumptions and study a variety of urban programs. The more we worked with the model the more we appreciated its flexibility and consistent underlying theory. Our major reservations concerned the solution algorithm and, in particular, the inability of such practical but ad hoc algorithms to identify solutions with consistent, known properties.
Part 8: Conclusions and Recommendations

The conclusions and recommendations are organized into three sections. The first outlines the approach used to interpret the results. The next interprets the housing allowance runs for the Pittsburgh calibration. Finally, policy recommendations are developed for designing payment formulas to match various housing market situations.

8.1 Interpreting the Results

The housing allowance simulations described in Part 6 used pure forms of income-gap and percent-of-rent payment formulas. As was illustrated in Tables 6.1 and 6.4, the parameters of the plans were selected to correspond to low, medium, and high assistance levels with and without different levels of minimum housing standards "earmarking". Two "medium" level income gap formulas were used -- one emphasizing higher payments per household and the other a larger number of eligible households.

The level of the medium income-gap plan was arbitrarily selected to cost an average of roughly $60 per household per year (exclusive of administrative expenses) -- a cost that could be financed in Pittsburgh by an income tax of approximately 0.8 percent.* For purposes of comparison, simulations of percent of rent formulas assumed that the same households were eligible as was the case for the comparable income gap formulas.

*We have not attempted to model the financing of housing allowance in any detail. The reference to break-even income tax rates is intended to illustrate the level of various allowances. In running the model, the income of non-eligible households is reduced by 1 percent as a rough approximation to the net effect of financing.
As is often stated in the literature, housing allowances may be designed to further a variety of (sometimes conflicting) goals. Of particular interest are the effects on housing quality, the rent burden of the poor, the efficiency of the subsidy and various accompanying effects, such as price inflation, geographic changes and distributional equity. The strong point of the ten year simulation is its simultaneous consideration of altered demand and supply responses in identifying price and housing changes within various submarkets.* In addition, we are particularly interested in the effect of our reduced income elasticity on the housing improvement and efficiency of housing allowance subsidies.

Accordingly, we shall interpret the Pittsburgh runs by comparing the inflation, housing improvement, efficiency and distributional effects of allowance payment formulas that use different forms, payment levels and minimum standard "earmarking" requirements. The results for the simulations where the income elasticity is assumed to be 0.47 are analyzed first. Then the conclusions are contrasted with those that would result if the income elasticity for housing allowance income were optimistically and, we believe, inappropriately assumed to be unity.

Relating the predicted absolute levels of rent and quantity of housing service to actual values in any one year is difficult. The runs simulate 1960 to 1970 changes in the city’s housing market that would have resulted if particular housing allowances had been in effect during those ten years. In the model, nominal income and rents are

*Recall that the discreteness problem mentioned earlier (and associated with integer assignments of model households to model dwellings) limits our ability to identify racial and other geographic patterns of change.
assumed. However, the prices correspond to some time average over the 1960-1970 decade that is simulated. Hence, the relative magnitudes rather than the absolute levels of price, earmarking and the like are particularly important when comparing housing allowance results with each other and with base conditions in 1970.

8.2 Housing Allowances in Pittsburgh

Housing allowance subsidies might conceivably improve the quality of housing in an urban area either by stimulating the repair and rehabilitation of existing dwellings (leaving assignments unchanged) and/or by increasing the mobility of the poor so that some recipients move to higher quality submarkets. For the second situation to occur, both additional new construction and additional abandonments are required — that is, the population improves its average housing condition by constructing new units at some quality level, relocating certain households and abandoning the poorest housing. Hereafter, we shall refer to the first situation as "upgrading" and the second as "upward movement." The distinction will be particularly important.

The 1970 Pittsburgh calibration for the modified model (and for the Urban Institute version as well) indicated substantial price "discounts" for the low quality end of the existing housing stock in the absence of housing allowances. For that third of the housing stock providing the least quantity of service* the price per unit service was 1.02, approximately 25 percent below the new house price of 1.33 and 20 percent below the 1.27 average price for the other two thirds of the housing stock in 1970.

*Recall that quantity of service, Q, combines size and quality measures. One unit of Q had an average market value of $1.00 in 1960.
Figure 8.1 graphs the Price-Quantity relationship for the 1970 base conditions in Pittsburgh.

For whatever reason, the bottom third of the housing stock was allowed to deteriorate between 1960 and 1970 reaching a 1970 price per unit of service below the 1.13 level needed to maintain Q at 1960 levels.* This situation, together with a minimum standard for new housing (of 66 units of service per month) that kept the poor out of the new house market, virtually eliminated any "upward movement" effect in all the Pittsburgh housing allowance runs. The poorest households could not afford new housing and their increased demand was not sufficient to stimulate much additional new construction by wealthier households not restricted by the 66 unit minimum. Substantial improvements in the low-quality housing stock were possible before the price would have reached the new construction price and make new construction competitive (for reasons other than externalities and flexibility in selecting Q) at the low to medium quality levels.

In reality, some filtering might occur. Changes in the housing to household ratio due to family grouping or the subdivision of dwelling units cannot be modeled. Moreover, the "discreteness" problem and the level of aggregation of the model prohibit distinguishing a change of less than 15,000 units of new construction (or abandonment). Nevertheless, the consistent absence of new construction and the fact that the most expensive program considered raised average prices in the bottom third only to 1.23 suggest that significant upward movement would not result (in Pittsburgh) from

*Recall the discussion in Parts 5 and 7 of the relationship among the betas, the new construction price, and $R$ (which in this case equals 1.13).
FIGURE 8-2

PRICE-QUANTITY RELATIONSHIP IN PITTSBURGH -- (11, 0) INCOME-GAP FORMULA

\[ P_N = 1.33 \]

Zones 2, 3, 4

Zone 1

\[ P_0 = 0.54 \]

Price Per Unit of Service

Quantity of Service

1.50 1.25 1.00 0.75 0.50

250 200 150 100 50 300
FIGURE 8.3
PRICE-QUANTITY RELATIONSHIP IN PITTSBURGH -- (40, 0) PERCENT-OF-RENT FORMULA

$P_N = 1.33$

$P_0 = 0.54$

Zone 1
Zones 2, 3, 4
income gap and percent of rent payment formulas such as those considered here.

Income Gap Allowances

As explained earlier (see Table 6.1), income gap allowances were considered which allowed either 15 percent (six model households) or 22 percent (nine model households) to be eligible. To at least partially account for the funding of such allowance programs, the allowance simulations reduced the income by 1 percent of all households ineligible for the allowances.

A comparison of the effects of the various housing allowances is presented in Table 6.1. The specific parameter values and allowance vs. income or allowance vs. rent graphs were presented earlier in Tables 6.1 and 6.4. The payment formula, parameter values and earmarking levels substantially affected the results. The results for each measure of effectiveness will be discussed separately in the next few paragraphs.

1. Inflation Effects and Housing Improvement: Without upward movement, improvements in housing must come from upgrading. In the model, this corresponds to landlords moving up their supply curves increasing price linearly with the quantity of service provided.* The allowances did alter several household assignments but did not change the number of new dwellings. Understandably, the P-Q changes were confined to the bottom third of the market with only small reductions in Q elsewhere (due to the 1 percent tax). There, P and Q changes moved together with a 1 percent

*Recall that all supply curves are linear with the same intercept at Q = 0 but shallower slopes for larger values of Q_0, the initial quantity of service provided in 1960.
Table 8.1 Comparing the Pittsburgh Results — $E_{HA} = 0.47$

<table>
<thead>
<tr>
<th>Payment Formula</th>
<th>Allowance Level Eligibility</th>
<th>Price Changes$^5$ Overall</th>
<th>Housing Changes$^6$ Bottom Third</th>
<th>Efficiency Measures</th>
<th>Housing Change per $100^9$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Cost$^4$ (per month x 10$^5$)</td>
<td>Bottom Third</td>
<td>Bottom Sixth</td>
<td>Participation</td>
<td>Expenditure Ratio$^7$ Bottom Third</td>
</tr>
<tr>
<td>Income Gap, Earmarking</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(High, O)</td>
<td>22% 4.88</td>
<td>2.5% 11.2%</td>
<td>3.9% 8.0%</td>
<td>100%</td>
<td>22% 9.5</td>
</tr>
<tr>
<td>(Med., O)</td>
<td>22 3.20</td>
<td>1.7 5.4</td>
<td>2.3 4.8</td>
<td>100</td>
<td>21 8.8</td>
</tr>
<tr>
<td>(Med.*, O)</td>
<td>15 3.30</td>
<td>1.7 6.3</td>
<td>2.4 6.3</td>
<td>100</td>
<td>20 9.0</td>
</tr>
<tr>
<td>(Med., Med.)</td>
<td>22 1.80</td>
<td>3.4 10.2</td>
<td>4.6 9.6</td>
<td>67</td>
<td>47 33.5</td>
</tr>
<tr>
<td>(Low, O)</td>
<td>15 1.25</td>
<td>0.8 2.4</td>
<td>0.6 2.3</td>
<td>100</td>
<td>19 7.9</td>
</tr>
<tr>
<td>(Low, Low)</td>
<td>15 0.57</td>
<td>0.8 3.4</td>
<td>1.2 3.8</td>
<td>67</td>
<td>53 50.8</td>
</tr>
<tr>
<td>% of Rent, Earmarking</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(40%, O)</td>
<td>22 3.01</td>
<td>1.7 4.4</td>
<td>2.5 2.2</td>
<td>100</td>
<td>22 10.3</td>
</tr>
<tr>
<td>(30%, O)</td>
<td>22 2.19</td>
<td>1.7 2.4</td>
<td>1.5 1.5</td>
<td>100</td>
<td>24 8.9</td>
</tr>
<tr>
<td>30%, Low)</td>
<td>15 1.00</td>
<td>0.8 3.4</td>
<td>1.5 2.0</td>
<td>67</td>
<td>40 25.4</td>
</tr>
<tr>
<td>1970 Base Conditions</td>
<td>-</td>
<td>1.19 1.025</td>
<td>54.19 41.56</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

$^4E_{HA}$ is the Elasticity of demand with respect to Housing Allowance Subsidies.

1. Terms in parenthesis refer to allowance level and minimum standard earmarking (see text and Table 6.1).

(continued on next page)
(Table 8.1 continued)

2. Terms in parenthesis refer to percent to rent subsidized and minimum standard earmarking (see text and Table 6.4).
3. Percent of all Pittsburgh households.
4. Approximate cost in millions of 1970 dollars per month (exclusive of administrative costs).
5. The average percent increase in the price per unit of service. "Bottom third" refers to the average over that third of the households demanding the lowest quantity of housing.
6. The average percent increase in the quantity of housing provided by those dwellings occupied by the "poorest" third and sixth of all households. The overall change is small since the better quality housing changes little (see Tables 6.1 and 6.4).
7. The increase in rental expenditure of eligible households as a percent of the total housing allowance subsidy.
8. The change in the average number of housing units consumed by the "poorest" third and sixth of all households. This change is measured per $100 per month transferred to the "poorest" third of all households.
rise in $Q$ accompanying an increase of about 3 percent in $P$. The $P - Q$ relationships resulting from income gap and percent of rent formulas are illustrated in Figures 8.1, 8.2, and 8.3. Figure 8.1 presents the 1970 base conditions. Figure 8.2 shows the medium level income gap case (Med. *, 0) without earmarking and with 15 percent of the households eligible for the allowance. Figure 8.3 graphs the 40 percent-of-rent case without earmarking.

(2) Level of Allowance Payment: Since the model finds ten-year equilibriums, it can identify long-run but not short-run effects of allowance payment levels. For Pittsburgh, with its initial price discount for low-quality housing, even sizable allowance programs do not drive $P$ above the new house price and induce upward movement due to new construction in some segment of the market. Even the 22 percent eligible program with an average allowance of $36 per month per household increases $P$ for the bottom third only to $1,14$ and stimulates no additional new housing.

If upward movement were induced for allowances above some critical size, dramatic changes in efficiency might result since upward movement could provide increased $Q$ at constant price. However, for Pittsburgh, the size of the income gap program needed to generate more positive effects due to such upward movement in substantial -- more than 6 or 7 million (1970) dollars per month. Such allowance levels probably imply marginal tax rates that result in too much work disincentives and/or too large an eligible population to be practical.

(3) Earmarking Minimum Housing Standards: As the results in Table 8.1 indicate, earmarking* can have a dramatic impact on the effectiveness of housing allowances.

*Only a minimum standard type of earmarking was modeled. Minimum expenditure standards would have similar effects though price changes complicate the determination of comparable rent and $Q$ standards. In reality the choice depends upon ease of administration, incentive to fraud, and other such practical considerations. Here we estimate the amount of change that can be induced.
When earmarking was specified for a particular allowance payment formula, Q continued to increase while total allowance payments dropped (because of lower participation). Comparing Q changes in the bottom third and sixth of the market, earmarking was able to induce an additional 2 to 5 percent increase thereby doubling the improvement in housing consumption that the unearmarked allowance could produce (see the (Med., Q) and (Med., Med.) results in Table 8.1).

However, setting the level for the minimum standard is critical. Levels too high created severe hardships for the poorest households who could not spend enough of their income to meet the standard but still faced significantly higher prices in the post-allowance market. Attempts to raise the minimum standard and force expenditure ratios above 0.5 were unsuccessful and produced lower participation rates and less housing improvement as well.

Unfortunately, the absolute levels of the more effective minimum housing standards used in the model are not easily interpreted.* They measure units of housing service originally standardized using 1960 market values. Also, the model stimulates average market conditions during the 1960-1970 decade. Hence, the actual 1974 (or even 1970) rent associated with specific model housing standards is difficult to estimate. The earmarking levels should be interpreted relative to the base conditions.

For Pittsburgh, the results indicate that minimum housing standards should be designed at most to match the Q increase that would result without earmarking. That

*The medium level earmarking corresponds to Q standards of 53 and 44 units of service for household types 1 and 3 and types 2 and 4, respectively.
is, one can expect to double the Q change listed in the no-earmarking rows of the "bottom third" column of Figure 8.1. Without earmarking, the medium level income gaps produced 2.3 percent and 2.4 percent changes in Q. With proper earmarking the figure doubled.

(4) Efficiency: Two measures were used to indicate how efficiently the allowance programs improved housing conditions. The expenditure or earmarking ratio, ER, measures the change in housing expenditures (among recipients) per dollar of allowance. The housing improvement ratios focus on changes in the quantity of housing provided by the bottom third and bottom sixth of the (occupied) housing stock. As an index of efficiency, we consider the average change in Q per dollar of net payments (allowance minus tax) to that third of the households desiring the least amount of housing.

The expenditure ratios (see Table 8.1) measures changes in rent and hence reflect price as well as quantity increases. Without earmarking, only slightly more than 20 percent of the allowance dollars were channeled into increased rents.

The housing improvement ratios measure only housing changes and do not reflect price increases. The housing improvement is normalized by the net average transfer to a household in the bottom sixth -- a measure of the influx of funds to that segment of households.

*If we valued the change in Q at the average post-allowance price, these ratios could be interpreted as dollars of improvement per dollar transferred. However, this was not done since the numbers would be misleadingly low. Improvements come about through upgrading -- a process that requires landlords in a competitive market to increase the price (per unit service) that is charged for each unit of service provided and not just the last few.
The expenditure and housing improvement ratios move together, as expected. The important point to note about the two improvement ratios is that income gaps help the poorest sixth of the households the most -- often more than twice as much. This result occurs since income gap formulas give the poorest the largest allowances.

Earmarking the allowance via minimum housing standards was able to increase the efficiency ratios 200 to 600 percent. However, the changes are somewhat misleading since they primarily reflected reduced allowance levels and not dramatic improvements in \( Q \). Households who refused the allowance because of the level of the minimum standard still faced the higher price levels that resulted in the market. Thus, the most efficient programs caused much inequity -- a part to be discussed shortly.

Even considering the income elasticities of 0.5, the expenditure ratios of 20 percent for the cases without earmarkings are quite low. The fact that the low quality housing was initially discounted and can improve only by moving along fairly inelastic supply curves limits the increase in housing expenditures and makes earmarking more important.

The expenditure ratios increased with the level of the allowance -- an interesting result. However, the expectation term of each household's utility assumes that income elasticities increase slightly as income rises (\( E = Y/(y + b) \)). Another explanation will be offered in the discussion of locational effects.

(5) Distribution of Benefits: The (Med., 0) and (Med.*, 0) programs spend about the same amount of money -- $3.2 and $3.3 million per month -- but distribute
it differently. The (Med., 0) program distributes allowances averaging $23.72 per month to nine of the 14 households. The (Med.*, 0) program distributes $36.72 per month to the six most needy households. A comparison of the results in Table 8.1 indicates that fewer people and more dollars per person is a better choice in this case. The (Med.*, 0) program emphasizes the poor but, through interaction of forces in the market, still causes the average Q for the entire bottom third to rise more than in the (Med., 0) case. (Of course, those who do not receive an allowance under the (Med.*, 0) program experience higher rent burdens and lower utilities.)

For such situations, giving the next dollar to the poorest segment of the market appears to be a good choice unless, (1) the level becomes so high that one must consider work disincentives, or (2) spending the money on increasing the number of eligible households will produce significant upward movement. Based on our earlier discussion, the second case is unlikely to arise first.

**Percentage of Rent Allowances**

As explained in Part 6, (see Table 6.4), percent of rent allowances comparable to the income gap forms were analyzed. They subsidized 30 or 40 percent of the rent paid by either six (15 percent) or nine (22 percent) eligible households. Once again, the allowances were balanced by a 1 percent "tax" on incomes of all households ineligible for the allowance.

(1) Inflation Effects and Housing Improvement: Since there is little movement to higher quality submarkets, \( Q \) increases continue to result from upgrading. The \( \Delta P/\Delta Q \) ratio is now slightly less than for income gaps since the percentage-of-rent formula focuses more of the allowance on the "wealthier" recipients where the supply curves are somewhat
more elastic. For the bottom third of the market $A^P/\Delta Q$ ratios are now 0.033 for the (40, 0) percent of rent program compared with 0.05 for the $(M^*, 0)$ program (See Table 6.4).

(2) Level of Allowance Payment: Once again, new construction was unaffected by the housing allowances and upward movement does not occur. More expensive programs than the 40 percent rent subsidy with 22 percent eligible might induce new construction. In fact, a percent-of-rent formula is more likely to stimulate filtering than a comparable income gap formula — a conclusion to be developed later in Section 8.3. However, the most expensive program erased less than 1/4 of the difference between the new house price and the average price for the bottom third of the housing market. Percentage-of-rent allowances in Pittsburgh of sufficient size to enable the market to profit from upward movement as well as from upgrading would most likely cause too much disparity between recipients and non-recipients or would require too large an eligibility proportion to be practical. Variations of the basic percentage-of-rent formula that might avoid some of these problems will be discussed in Section 8.3.

(3) Earmarking Minimum Housing Standards: Just as in the income gap case, earmarking can dramatically effect the results of percent of rent formulas. However, the effect was somewhat diminished for the percent-of-rent programs since the minimum standard constraint was now least important to those recipients who received the largest subsidies since they are the richest recipients and often meet the standard voluntarily. Obtaining solutions for several percent-of-rent allowances was complicated by convergence problems with the solution algorithm (see Technical Appendix).
Nevertheless, examination of the (30, L) and (L, L) cases with six households eligible indicate that the same earmarking resulted in significantly larger Q changes for the (L, L) income gap program. The most effective level for the housing standard (in terms of efficiency with at least 2/3 participation) appears to be the same in both cases.

The fact that earmarking continues to have an important effect for the price-subsidy payment form (where the price elasticity is assumed to be unity) illustrates the difficulty associated with trying to extrapolate market equilibrium results from reasoning how individual households behave.

(4) Efficiency: Expenditure ratios for percent-of-rent allowances without earmarking were around 0.23, only slightly higher than for income gaps. Housing improvement ratios were also slightly higher (10 instead of 9) for the bottom third of the market but about half as large for the bottom sixth (8 instead of about 16) -- again a result of the emphasis on the wealthier recipients. The dramatic difference in distributional effects is the primary difference between the two payment formulas. Earmarking again improved the expenditure and improvement ratios but could raise them only to 40 and 25 compared with 53 and 51 for the (L, L) program. This time expenditure ratios dropped slightly for the more expensive programs.

(5) Distribution of Benefits: Whereas income gap formulas paid the most to the poorest households, percent of rent formulas pay the most to those with the highest rents. Although the (M*, 0) and (40, 0) programs yield similar average Q values for the bottom third of the market (55.47 and 55.55 respectively), the (40, 0) program does much worse for the poorest sixth (42.47 compared with 44.17 for (M*, 0) and
41.56 without any allowance). Also, percentage of rent forms produce more disparity between recipients and non-recipients. The largest subsidies and, hence, the largest stimulus for competition occur at quality (quantity) levels desired by the wealthiest recipients and the poorest non-recipients. The equilibrium Price vs. Quantity curve is less smooth. The poorest non-eligible households face larger price increases.

(See Figure 8.3)

Locational Effects

Since the model divides Pittsburgh into only four zones (of existing housing) and has only three black households, it is presumptuous to expect conclusive predictions regarding geographic patterns. A few results, though, appear noteworthy.

The racial composition of the zones was fairly stable. In virtually all cases at least one black household remained in zones one and three. The third family either remained in zone one, moved to zone four or, least often, moved to zone two. No pattern emerged. Since the small integer situation meant that one black household switching zones could dramatically affect the racial composition, we hesitated to attribute too much significance to such small changes.

A systematic pattern did emerge for the average income of residents in the poorest zone -- zone 1. Allowances caused this average income to drop. In fact, the largest drops ($421 per month to $354 per month -- not counting the allowance) resulted from the most expensive allowances. This result puzzled us at first but appears explainable. Several factors are involved. Income gap allowances narrow the income distribution at the bottom making households more similar and hence more competitive for some housing. Percent-of-rent allowances spread the income distribution of recipients further apart and,
hence, promotes less competition -- except near the cutoff. In both cases, the two
effects are offsetting -- competition tending to segregate households by income (by
driving up prices and restricting the range of housing that individuals can afford )
while closer incomes tend to reduce inequality. However, for all the allowance
programs, prices increased substantially in the poorer zones and particularly in zone
one. Evidently, in Pittsburgh, this increase was enough to cause further segregation
of the poor -- with those who could afford it controlling the slightly better housing
that had more elastic supply curves and could offer proportionally more increase in Q
per additional dollar of rent.

Comparing Income-Gap and Percentage-of-Rent Forms

In Pittsburgh, the choice between income-gap and percent-of-rent allowances depends
heavily on value judgment and administrative considerations. Both programs cause
comparable improvements in housing consumption and are responsive to reasonable
minimum standard earmarking. The major difference is in the distribution of benefits.
Income gap formulas help the poorest the most -- though at a slight loss in overall
efficiency -- because helping the poor focuses on inelastic supply curves and costs
more. Percent-of-rent allowances help the "wealthier" recipients most, have more
uneven price effects and cause greater disparity between recipients and non-recipients.
For Pittsburgh the disadvantages to the percent-of-rent form would appear to outweigh
the small difference of about 1 percent in price increases. However, administrative
ease, work incentives or short-term considerations might tip the choice in the other
direction.* Variations of the percent-of-rent approach offer another alternative and

*Indeed other differences might arise in cities where more upward movement might
occur. Such possibilities will be discussed in Section 8.3.
will be discussed in Section 8.3.

Unitary Income Elasticities

One of the primary model modifications was to modify the unitary household elasticities of demand in keeping with the results of Chapter 4. Income elasticities were changed to 0.47. The unitary price elasticity was consistent with empirical results and was retained. The above simulation runs for Pittsburgh assumed that recipients treat housing allowances as if they were added permanent income (except for earmarking constraints). Conceivably, housing demand of recipients might be more responsive to housing allowances than to other increases in permanent income. While we regard this outcome as unlikely, there is sufficient interest in unit income elasticity so as to warrant a comparison of the above results with those for the higher elasticity case.

As explained in Part 3, the expectations in the utility functions of allowance recipients were altered so that their elasticity of demand with respect to the added income from the allowance was equal to 1.0. The new results for the same set of housing allowance simulations considered earlier are given in Table 8.2. The increased responsiveness of housing demand results in dramatic increases in the effects of the programs. Each dollar buys 2.5 to 3.5 times as much Q improvement (see the housing improvement ratios). Expenditure ratios more than double and the 1 percent tax, 15 percent eligible cases -- e.g., the (M, M) plan -- brings the average price in the lower third of the market up to 1.28. Allowance levels a little higher might induce significant upward movement as well as upgrading.

Another important change is that income-gap programs now appear to dominate the percentage-of-rent programs. Not only do the income-gap plans help the poorest
Table 8.2 Comparing the Pittsburgh Results when $E_{HA} = 1.0$

<table>
<thead>
<tr>
<th>Payant Formula</th>
<th>Allowance Level Eligibility Cost $^1$ (per month x $10^5$)</th>
<th>Price Changes $^5$ Overall Bottom Third</th>
<th>Housing Changes $^6$ Bottom Sixth</th>
<th>Efficiency Measures Participation</th>
<th>Expenditure Ratio $^3$ Bottom Third Bottom Sixth</th>
<th>Housing Change per $100$ $^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Income Gap, Earning (High, O)</td>
<td>22% $54.88$</td>
<td>7.6% 23.9</td>
<td>10.8 21.8</td>
<td>100% 64%</td>
<td>25.5 40.9</td>
<td></td>
</tr>
<tr>
<td>(Med., O)</td>
<td>22 3.20</td>
<td>4.2 15.1</td>
<td>6.5 13.4</td>
<td>100 58</td>
<td>24.8 39.3</td>
<td></td>
</tr>
<tr>
<td>(Med$, O)</td>
<td>15 3.30</td>
<td>5.9 19.0</td>
<td>8.1 19.3</td>
<td>100 65</td>
<td>30.7 56.2</td>
<td></td>
</tr>
<tr>
<td>(Med., Med.)</td>
<td>22 1.80</td>
<td>7.6 24.9</td>
<td>11.2 23.0</td>
<td>89 110</td>
<td>81.5 128.1</td>
<td></td>
</tr>
<tr>
<td>(Low, O)</td>
<td>15 1.25</td>
<td>1.7 4.4</td>
<td>1.7 5.5</td>
<td>100 45</td>
<td>20.9 51.4</td>
<td></td>
</tr>
<tr>
<td>(Low, Low)</td>
<td>15 0.57</td>
<td>2.5 7.3</td>
<td>3.4 9.6</td>
<td>67 107</td>
<td>91.1 321.8</td>
<td></td>
</tr>
<tr>
<td>% Of Rent, Earning (40%, O)</td>
<td>22 3.32</td>
<td>3.4 10.2</td>
<td>6.4 5.8</td>
<td>100 41</td>
<td>23.7 26.3</td>
<td></td>
</tr>
<tr>
<td>(30%, O)</td>
<td>22 2.43</td>
<td>3.4 9.3</td>
<td>5.2 4.7</td>
<td>100 50</td>
<td>26.7 18.7</td>
<td></td>
</tr>
<tr>
<td>(30%, Low)</td>
<td>15 1.10</td>
<td>3.4 9.3</td>
<td>5.0 4.3</td>
<td>67 68</td>
<td>71.4 47.4</td>
<td></td>
</tr>
</tbody>
</table>

$^1$ $EHA$ is the Elasticity of demand with respect to Housing Allowance Subsidies.
$^3$ Terms in parenthesis refer to allowance level and minimum standard earmarking (see text and Table 6.1).
(continued on next page)
2. Terms in parenthesis refer to percent to rent subsidized and minimum standard earmarking (see text and Table 6.4).
3. Percent of all Pittsburgh households.
4. Approximate cost in millions of 1970 dollars per month (exclusive of administrative costs).
5. The average percent increase in the price per unit of service. "Bottom third" refers to the average over that third of the households demanding the lowest quantity of housing.
6. The average percent increase in the quantity of housing provided by those dwellings occupied by the "poorest" third and sixth of all households. The overall change is small since the better quality housing changes little (see Tables 6.1 and 6.4).
7. The increase in rental expenditure of eligible households as a percent of the total housing allowance subsidy.
8. The change in the average number of housing units consumed by the "poorest" third and sixth of all households. This change is measured per $100, per month transferred to the "poorest" third of all households.
the most but all the expenditure and improvement ratios are larger for income-gap
plans than for percent-of-rent plans with comparable government costs and earmarking
requirements. If only income elasticities were equal to one, then we could be optimistic
and confident in our choice of payment formula!
8.3 Recommendations for Housing Allowance Payment Formulas

In this section, we generalize the Pittsburgh conclusions of Section 8.2 and develop recommendations about the effects on various types of cities of the two "pure" housing allowance payments formulas that were studied. Much of the discussion refers to the comparison of Pittsburgh results as presented in Table 8.1.

- **Motivating the poor to improve their housing voluntarily through income transfer** is quite costly. Without some form of earmarking*, expenditure ratios are around 20 percent (see Table 8.1). Moreover, much of the increased expenditures goes toward increased prices — about two-thirds. **

- **Earmarking minimum housing standards** can double the increase in housing consumption that would otherwise occur and can substantially improve a program's efficiency. For example, in Table 8.1, medium earmarking applied to the medium level ($3.3 million dollars per month) income gap formula (plans $M, O$ and $M, M$) increased the amount of improvement in housing from 4.8 percent to 9.6 percent and more than doubled the expenditure ratio (rental income per dollar of allowance). Raising earmarking standards higher resulted in less housing improvement due to substantially lower participation rates.

- **Substantial tradeoffs are possible between efficiency and equity** depending upon

---

*That is, minimum housing standards, rent certificates, expenditure minimums and the like.

**To the extent that these higher prices are needed to maintain upgraded housing, such price increases should not be viewed as inflationary and inefficient. The two-thirds figure comes from comparing the third and fourth columns in Table 8.1.
the level of earmarking that is used. For the \((M, M)\) income gap case just mentioned, the housing changes were substantial and the efficiency appeared high. However, one-third of the eligible households choose not to participate. Moreover, this one-third comprised the poorest households. In such a case, the non-participants and the poorer non-eligibles find themselves occupying better housing -- but at higher prices -- than before the allowance went into effect even though they experience no increase in income.

Though earmarking can induce households to improve their housing conditions high earmarking levels force the poorest off the allowance. They have too little income to accept rental levels required to meet the standard. Low levels of earmarking have an effect only on the poorest households since the others will meet the standard voluntarily (but, since it is voluntarily, with low expenditure ratios).

One must choose between efficient programs with expenditure ratios that can rise to around 50% (but with participation rates of about two-thirds) and more equitable programs that keep the poorest households on the allowance but result in expenditure ratios of around 20%. A possible way around this dilemma would be to use a set of minimum housing standards that differed by household income (as well as size) thus inducing all eligible households to increase their housing consumption as much as possible.

- Initial market conditions can substantially alter the effect of housing allowances. In cities such as Pittsburgh the (pre-allowance) housing market is such that the low-quality housing is offered at price discounts (i.e., prices per unit of service substantially less than the new construction price). In such situations housing improvement caused by housing allowance comes more from the upgrading of existing dwellings than from the process of new construction, abandonment, and the movement of households into
better quality housing. As a result, housing improvement is costly (the price and quantity changes are tied to the slopes of the supply curves in the low-quality sub-markets). In cities where substantial upward movement might occur, improving the housing might be much less costly and expenditure ratios might be higher.*

- Benefit levels increase roughly in proportion to total allowance levels (for a given payment formula) in cities such as Pittsburgh.** As long as housing improvement comes primarily through upgrading, this is likely to be the case. If one could induce significant upward involvement, program efficiencies would improve. Hence (in other cities) one might find "critical" cost levels whereby programs spending more than this amount could induce substantial upward movement and, as a result, higher efficiencies.

If such upward movement is a possibility the next allowance dollar should be spent on increasing the number of recipients. Though such action could appear to help the poor less, it would (by stimulating upward movement) induce a market response that in fact helped the poor more. Where upgrading is likely, the next dollar is best spent on the poorest recipients.

- The choice between income gap and percent of rent payment formulas depends to a large extent on value judgments and initial market conditions. Again refer to Table 8.1 and compare income gap and percent of rent formulas with comparable costs (total allowance levels). See, for example, the \((M,O)\) income gap with medium cost

---

* Since the same improvement in housing now comes with less of a price increase, households will be willing to channel more of the allowance into additional housing consumption.

** Of course, we have not considered administrative costs here. They may experience economies of scale or diminishing returns.
level (3.3 million per month) and zero earmarking, and the 40% rent subsidy --
the (40%, O) plan -- with zero earmarking.\textsuperscript{+} Percent of rent formulas provide some-
what larger overall increase in housing consumption with less inflation but income
gap formulas do much more to help the very poorest (the "bottom sixth" of the house-
holds who live in the poorest housing). The 40% of rent subsidy provides a 22%
expenditure ratio and 10.3 additional units of housing service per $100 transferred
(to the "bottom third" of the households). The (M, O) income gap provides only
20% and 9.0 respectively but has almost three times the housing change in the "bottom
sixth" of the market. Earmarking is also more effective for income gaps since ear-
marking matters more for the poorest and they have much higher allowances under
income gaps.\textsuperscript{**}

In cities where substantial upward movement might be triggered by housing allowances
percent of rent forms might perform much better. In such cities without the initial price
discount for low-quality housing, allowances might be able to induce upward movement
by pushing some households (presumably in the middle income range) into new housing.
Since percent-of-rent forms increase demand (and subsidies) the most for the "wealthier"
recipients, they are more likely to stimulate filtering than are income gap formulas.
Hence, percentage-of-rent forms are likely to be preferable in markets without initial
price discounts, whereas income-gap forms look a little better in markets such as
Pittsburgh.

\textsuperscript{+}For a fair comparison formulas should have similar costs and the same eligible households
since efficiency measures and housing changes will depend on which households receive
subsidies.

\textsuperscript{**}One might have expected the opposite since percent of rent formulas act as a price
subsidy and the price elasticity is assumed equal to one. The results indicate the difficulty
associated with trying to predict market equilibrium results based on reasoning about
each individual's behavior.
with price discounts. They are not necessarily those with high vacancy rates. Vacancy rates tend to reflect shorter-term behavior not included in the model. This price discounting results from the interaction of a host of market forces such as the concentration (or dispersal) of income and the level of housing services in the population and the housing stock.

- Using an income elasticity of 0.47 instead of 1.0 significantly altered the model calibration and the housing allowance simulations. The modified model appeared to produce more realistic parameter estimates and resulted in substantially lower expectations for housing allowances and in a more difficult choice between income gap and percentage-of-rent programs.

- Housing allowances is tended to increase the concentration of the poor in the poorest zones. This outcome results from the interaction of several market forces that depend on the extent to which allowances narrowed the distribution of income and increased competition among the poor and on the distribution of services of the housing stock. For other cities it might not occur. Even for Pittsburgh, the result is tentative since the aggregation of the model and the requirement that households assignments be discrete limit the model's ability to identify such geographic effects.

- Finally, we should comment on what the model left out. The comparison of housing allowance payment formulas was essentially a comparison of the income gap and percent of rent alternatives in their purest form. Other "hybrid" forms might combine the better features of each plan.

For example, variations of percent-of-rent forms might reduce their undesirable distributional effects. Subsidy percentages that decreased with income are an example. Of course the trade-off is now less incentive for recipients to increase income by working.
A brief study of pure rent gap forms (allowance = $R^* - xR$) indicated that the incentive to purchase less housing was too large and an income gap formula could always be found that was preferable. Study of several triangle formulas, suggested that they would be more promising. However, to stimulate significant housing improvement comparable to those using income gaps one had to either make the sides of the triangle too steep or permit a large number of eligible households (25%-30%).

Selection and analysis of desirable percentage-of-rent forms that depended an income was beyond the scope of this study.

No consideration was given to administrative costs or to short-run effects. As in all models only a limited number of factors could meaningfully addressed at once.

Our focus was on the long-run dynamics that housing markets would take under housing allowances. Other aspects of the Center study focus on administrative considerations and short run responses.

The results for Pittsburgh probably look worse (in terms of efficiency) than they would for other cities since the improvement comes from the relatively costly upgrading rather than upward mobility. Indeed, the efficiency predictions are likely to be a lower bound on housing allowance effects since the discreteness problem mentioned earlier and other questions regarding the realism of the model are likely to have resulted in less predicted housing change per added dollar rather than more.

---

*That is, an allowance that provided a fixed percent of rent subsidy below some rent, $R^*$, cutoff and looked like a rent gap above $R^*$. 