file copy

MARYLAND INTERINDUSTRY FORECASTING PROJECT

Research Memorandum No. 25

November 19, 1969

CROSS SECTION ANALYSIS OF APPLIANCE DEMAND

by

Dwight Porter

One of the major inputs to the estimation of the consumption functions of the Maryland Model is an income elasticity for each category calculated from the 1960-61 Bureau of Labor Statistics cross-section consumer survey. Naturally, we have some reservations about continuing use of these income elasticities in 1969.

The most obvious draw-back is that nearly ten years have passed since the survey was conducted, and it is probable that some of the elasticities have changed since then. A second limitation of the BLS study is that it was possible to hold only family size constant when calculating the income elasticities. Clearly many other demographic variables may vary with income and bias such estimates of income elasticity.

Another limitation of the present elasticities is that they tell us only the response as everyone's income rises. But, if income were to rise more than proportionately for certain groups, say single person households, or if the income tax structure were altered in their favor, the elasticities might be invalid. In hope of correcting these deficiencies, we sought a source from which new income elasticities of a more varied nature could be estimated.

We have been partially successful. For the Appliances, Television sets, and Phonograph categories, we have been able to derive elasticities holding a great many more variables constant than in the estimation of the present income elasticities. We have made only limited progress in ascertaining different elasticities for different groups; we found our 11,000 observation sample too small for this purpose. Finally, we have

identified a number of demographic influences on the consumption of these items which should be of substantial value to their manufacturer.

Our data for this study comes from a national survey of the appliance market, begun by General Electric in 1967. The survey called for three thousand families to be interviewed each quarter and then to be re-interviewed twice, once six months following the original interview and once again another six months later.

Each household was queried on the make-up of the household, the income levels of its occupants and a range of other questions so that the demographic characteristics of the household could be clearly identified. Then the interview continued with a series of questions designed to identify the household's stock of forty appliances. If any item in that stock had been acquired within twelve months, then its purchase price was obtained by the interviewer. The combination of information on appliance consumption and the identification of a wide range of demographic variables for each household lends itself ideally to an econometric study of consumer buying patterns. This paper is an account of the progress that has been made in such a study on a selected group of eleven (Table I) from the total of forty appliances for which information was collected in the survey.

From the information made available by the survey, a set of independent, explanatory variables was selected (Table II). The dependent variable selected was the purchase price of the appliance; a zero indicated that there had been no purchase of the appliance within the past twelve months. The independent variables were created largely as dummy variables.

TABLE 1

	APPLIANCE	NUMBER PURCHASED IN 1967	AVERAGE PRICE PAID ($)
1	Television Sets	1,821	315
2	Portable Phonographs	559	56
3	Console Phonographs	336	305
4	Blenders	269	30
5	Tape Recorders	348	64
6	Refrigerators	808	276
7	Disposals	87	56
8	Room Air Conditioners	349	196
9	Freezers	208	215
10	Dryers	448	171
11	Washing Machines	917	198

Variable Number	Variable Name	State of Variable 1	0
1	Intercept	- 1	-
2	Household Income	less than 5,000	other
3	Household Income	5-5999	other
4	Household Income	6-6999	other
5	Household Income	7-7999	other
6	Household Income	10-1499	other
7	Household Income	15-19999	other
8	Household Income	20-24999	other
9	Household Income	25+	other
10	Race	Non-White	White
11	Family Size	1 person	other
12	Family Size	5 or more persons	other
13	Number of Adults	1-9	
14	Number of Children	1-9	
15	Age of Household Head	less than 25	
16	Age of Household Head	25 - 35	
17	Age of Household Head	36 and UP	
18	Educ. of Household Head	No. High School	other
19	Educ. of Household Head	Some High School	other
20	Educ. of Household Head	Some College	other
21	Educ. of Household Head	College Grad.	other

Variable Number	Variable Name	State of Variable 1	0
22	Rent or Own Dwelling	Rent	Own
23	Type of Dwelling	2 Family	other
24	Type of Dwelling	Multiple Row	other
25	Type of Dwelling	Apartment	other
26	Type of Dwelling	Other	other
27	Time at Current Address	less than 1 year	other
28	Time at Current Address	over 5 years	other
29	Region of Residence	South	other
30	Region of Residence	Midwest	other
31	Region of Residence	West	other
32	Area of Residence	large met. c. c.	other
33	Area of Residence	small met. c. c.	other
34	Area of Residence	small met. outside c. c.	other
35	Area of Residence	Non-Met.	
36	Length of Marriage	Not Married	other
37	Length of Marriage	Less than 1 year	other
38	Length of Marriage	1 to 5 years	other
39	Length of Marriage	11 to 15 years	other
40	Length of Marriage	16 to 20 years	other
41	Length of Marriage	21 to 25 years	other
42	Length of Marriage	26 and up years	other
43	Number of Autos	1 to 6	
44	Expected 12 month change in Household Income	Higher	other

Each variable was assigned a certain characteristic, and for every household (observation) that also had that characteristic the variable was given the value of one; otherwise it was set equal to zero. Thus, for example, the variable "race" (variable 12) was assigned the value one for all non-white households. Since information existed to define only two states of this variable, a zero for variable 12 implied a white household. For some of the other household descriptions, more than two states could be defined, requiring a group of dummy variables. The household income level was an example of such a variable; nine different dummies were used to describe it. In this case, a zero entry for any variable told only in what income range it did not fall. The existence of these groups of dummy variables created some difficulties; these will be discussed below. A few of the variables such as the number of automobiles owned by a household could be numerically represented rather than requiring a group of dummy variables.

The data from the survey for 1967 was provided by General Electric in a form which required considerable editing to put it into the form of the variables described above. In addition, approximately one thousand of the twelve thousand observations for the year had to be discarded as they were missing an important piece of information such as the household income range. However, a sample of approximately eleven thousand observations was adequate and cross-section regression equations were estimated for each appliance.

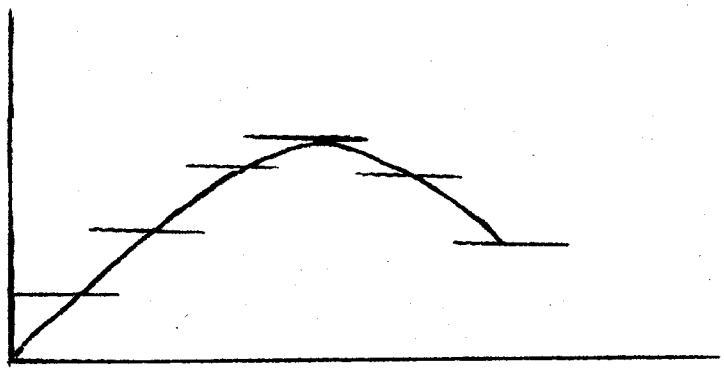
In order to be able to make any regression analysis possible, the structure of the variables included as independent variables had to be altered somewhat.

One dummy variable had to be omitted from each separate group of

dummy variables that defined one particular characteristic of the household, so that each group of dummy variables would be linearly independent from every other group and from the intercept column as well. This resulted in the definition of a standard household which is represented by the constant term coefficient. The other estimated regression coefficients were then deviations from this standard household. The other estimated regression coefficients were then deviations from this standard case. The standard household was defined as a two-to-four person, white family with an income of eight to ten thousand, its head being thirty years old, married six to ten years with a high school diploma and living in a single family home in the Northeast, outside of a large metropolitan area.

The regression program was written so that the first nine variables in each regression were the constant term and the eight income classes (remembering that one income class, eight to ten thousand, was omitted and represented in the constant term). At this stage, each coefficient represented the average amount spent by each household in the different income classes. This was represented graphically by a step-like function as shown in diagram I.

DIAGRAM I



To put this graph into a more familiar form, the midpoint of each income class was connected by a straight line so that the step graph was converted to a continuous graph with eight linear segments. These Engel curves, unadjusted for any demographic detail, appear as the solid lines in the graphs that are attached below.

The next step in the regression program was to proceed to introduce variables in the order in which they made the greatest reduction in the sum of the squared residuals. Generally the next variable introduced was the stock of the appliance being studied. This process was continued until the $\bar{R}^2$ of the regression reached its peak, that is, until the newly entering variable added less explanatory power than was lost by the reduction in degrees of freedom. At this point a second, adjusted, Engel curve was calculated and plotted as the dashed line on the graphs.

While the dollar figures represented by original Engel curves have significance as average expenditures by income classes, the adjusted curves have no such significance. Rather, each of the adjusted curves is one of a family of parallel curves where any one curve in the family represents a different combination of demographic characteristics for any household. This is a result of the fact that the coefficients for dummy variable effect only the intercept and not the slope of the curve. The one curve of each family that was selected was the one that intersected the unadjusted curve in the eight to ten thousand income range. The contrast of the two curves showed the change in spending patterns over income as certain demographic factors were taken into consideration.

The bar graphs in the right hand portion of the charts give an indication of the magnitude of the shifts that the dashed curves would make for certain demographic characteristics. The stalagmite bars denote positive changes and the stalactite bars are negative ones. The thickness of the bars is proportional to the statistical significance of the coefficients (two small blocks correspond to a value of 1.0.) Only large and significant coefficients are included in the bar graphs. The coefficients on "number of appliances owned" is not graphed because it is often too big for the graph. Its value is shown separately.

In examining the graphs, one can see that the adjusted and unadjusted Engel curves are remarkably similar. One can also see the same variables for certain appliances have extremely strong effects on the level of average consumption. Examples of these are the strong positive effect of living in the South on freezer and room air conditioner consumption and the strong negative effect of being non-white on the consumption of dryers, room air conditioners and washing machines.

This last example is striking and merits some detailed discussion. A non-white households spends sixteen dollars a year less on washing machines than does a comparable white household. There are several possible explanations. It is probable that in 1967 the bulk of advertising directed toward selling washing machines was aimed at the white buyer. Similarly, the white buyer probably had much easier access to retail outlet than did the non-white buyer. A third factor may have been the greater unavailability of installment credit for non-whites for reasons

other than an objective credit rating. All of these factors, if they are in fact valid, indicate a large potential non-white market for washing machines. If the average expenditures for non-white homes were brought up to the level for comparable white homes the sales revenue from washing machines could be increased roughly seventy million dollars. Of course, the same factors would apply to, say, console phonographs, of which the non-family buys more than the comparable white family. Another possibility is that the washing machine does not bring the prestige which the phonograph does. It is also, of course, possible that non-white women enjoy the company at the launderette more than the white women do.

In general, most of the results of this study have advertising implications. It is difficult to say without knowing past advertising patterns what specific conclusions might be drawn, but particular demand patterns are clearly identified and quantified as well, showing clearly where advertising could be profitably directed. It is interesting to note that a study of this nature can uncover relationships that would be hidden by a profile study from the same data. For example a profile would probably indicate that residents of the central city of large metropolitan areas buy less than their share of room air conditioners. Such an appearance is the result of low income in the central city. In fact, when all other factors are accounted for, living in the central city area has a substantial positive effect on air conditioner purchases.

The configuration of the Engel curves is of some interest. For the Maryland Interindustry Forecasting Project, the elasticities that can be obtained from them will have applications for the calculations

of some of our consumption functions. From the point of view of a manufacturer, they have a bearing on product planning. Curves with a flat configuration indicate small income effects as income rises over time. Conversely, a steep slope shows a rapid increase in demand as incomes generally rise. But apart from these general income responses, we may desire to anticipate demand changes for particular income changes. For example, if a guaranteed annual income becomes a reality, income changes are going to occur mostly in the lower ranges. By looking at the graphs, one can anticipate rapid increases in demand for room air conditioners, for example, while increases for blenders would be small.

In addition to the overall regressions discussed above, much important additional information can be obtained by splitting the observation matrix on the basis of selected variables. For example the observation matrix for console phonographs was divided on the basis of the race variable and two separate regression equations were estimated. The explanatory power of the independent variables was greatly increased by this split. The sample for non-white homes was too thin to continue with this branch of the split, but the other side could easily be further broken down. This split provides Engel curves for different household patterns that are not forced to be parallel by the nature of the available data. A new procedure for splitting is being tried by forming a group of composite variables such as non-white - South, non-white - West, non-white - large metropolitan, non-white - rent dwelling etc. and using these variables in addition to the regular list in a regression. Then these variables will be introduced when they make the greatest contribution to the reduction in the sum of the squared residuals and will cause only those splits of major significance to be considered.

One variable that is made interesting by its rate and usually insignificant appearance is the expected increase in income in the next 6 months. Although consumers may splurge on appliances after a change in income, they do not do so in anticipation of a raise.

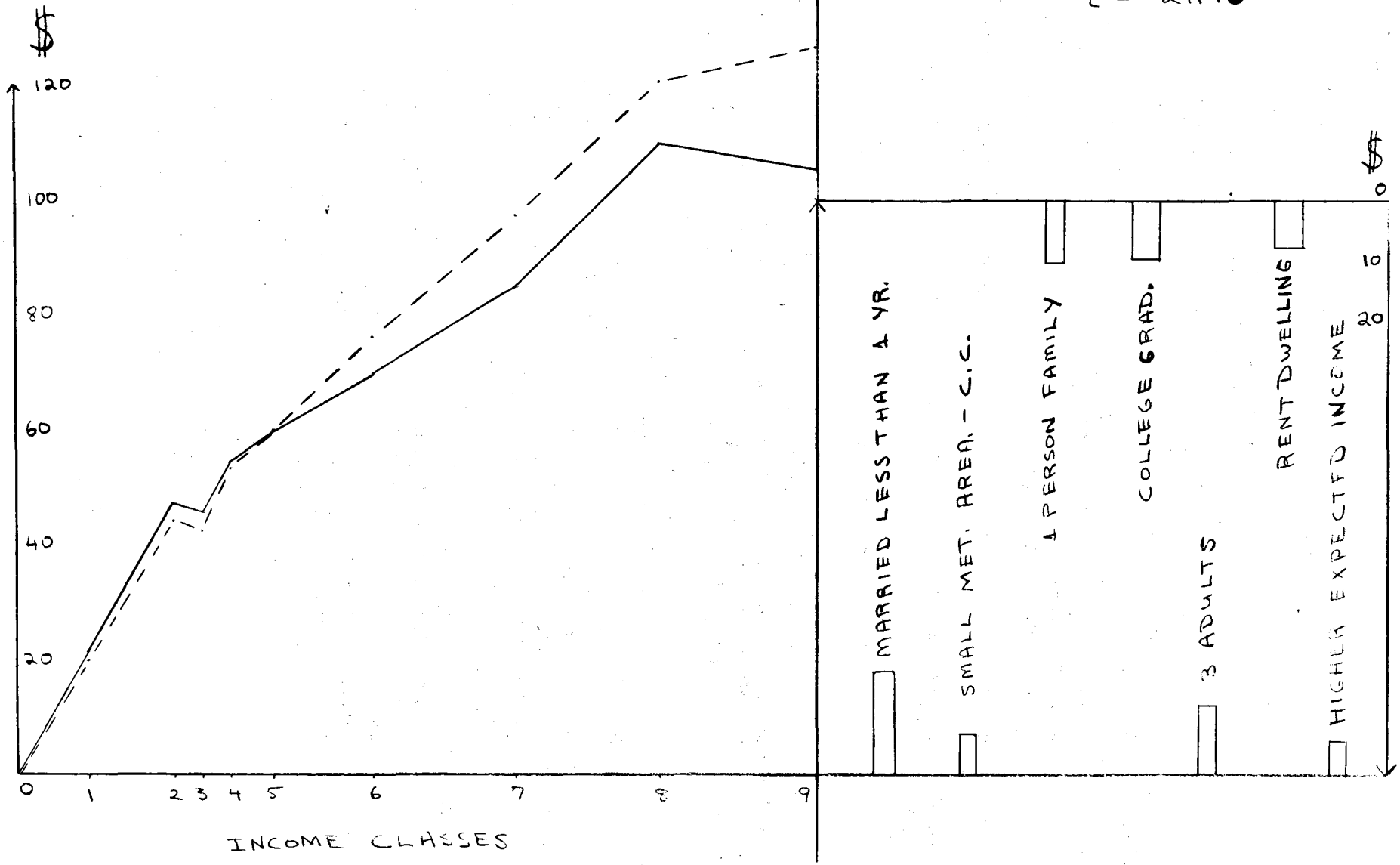
The R^2 levels attained in most of the regressions were quite low, in the range of .03 to .08, though some appliances did get as high as .30. For a cross-section study of this nature these R^2's are in line with the results of similar studies. The low level of the R^2 is also not particularly surprising when considered. If an observer were to know all of the detail included in the list of variables about any particular household and were asked to predict the amount the household was going to spend on a particular appliance, the observer would be hard pressed to give an accurate prediction. Appliances are generally large, indivisible commodities and a consumer must buy all or none of it in a given time period. So, unlike a smaller-price tag commodity such as clothes, it is very difficult to predict how much a household will spend on them in 1967. On the other hand, it is possible to determine the average amount of any particular group of households will spend. That is why the coefficients on the variables are so significant while the R^2's are low. Of course, the average expenditures are what matter for marketing.

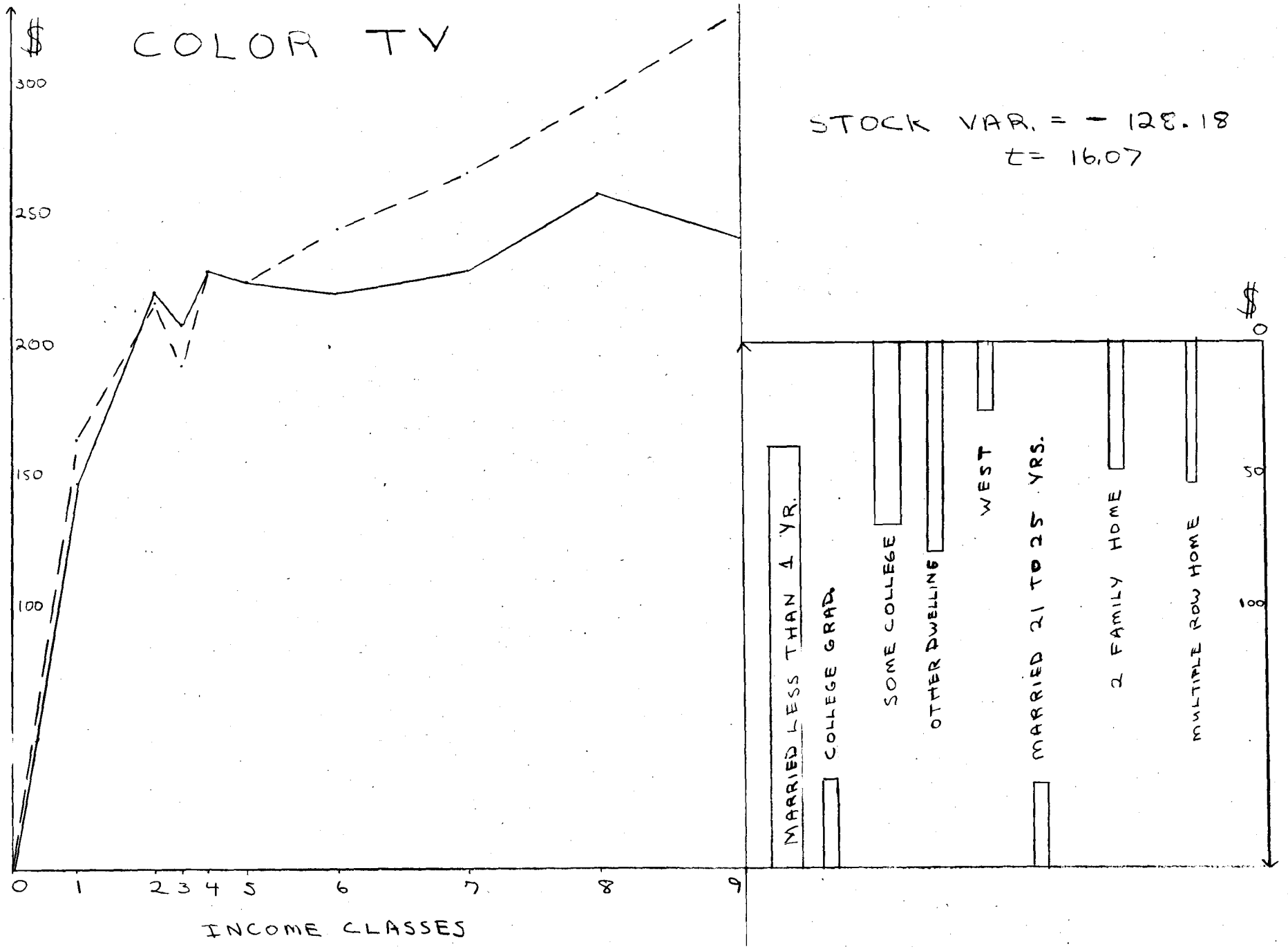
Many appliances are purchased as labor-saving devices for the housewife. Therefore, whether the wife is employed or not promises to be a very interesting variable in future studies. There are also a great number of additional appliances that should be looked at in order to get a complete picture of the appliance market, since it is clear that there is some product interaction.

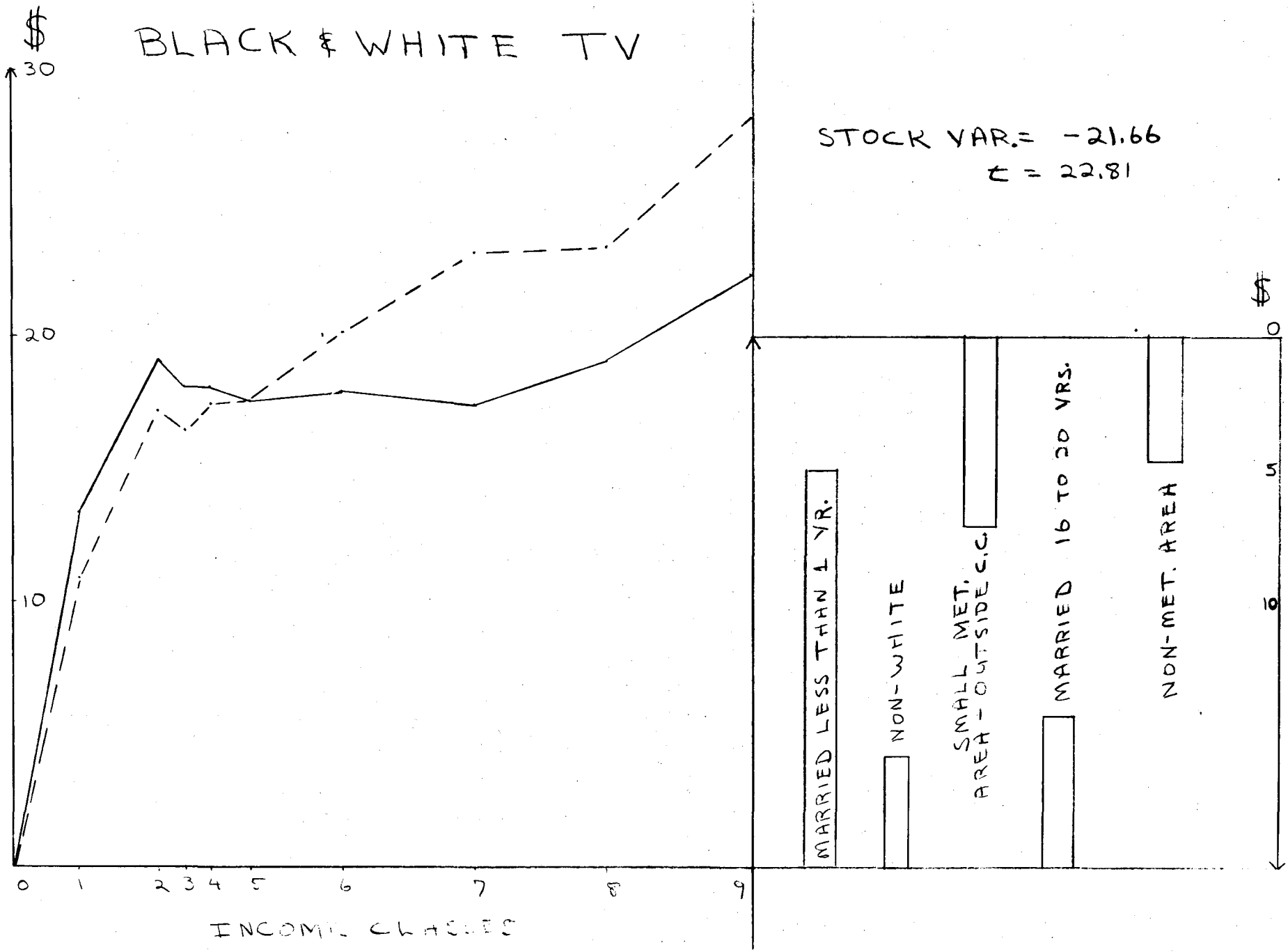
INCOME CLASS	INCOME RANGE
1	1 - 4,999
2	5,000 - 5,999
3	6,000 - 6,999
4	7,000 - 7,999
5	8,000 - 9,999
6	10,000 - 14,999
7	15,000 - 19,999
8	20,000 - 24,999
9	25,000 and up.

TELEVISION SETS

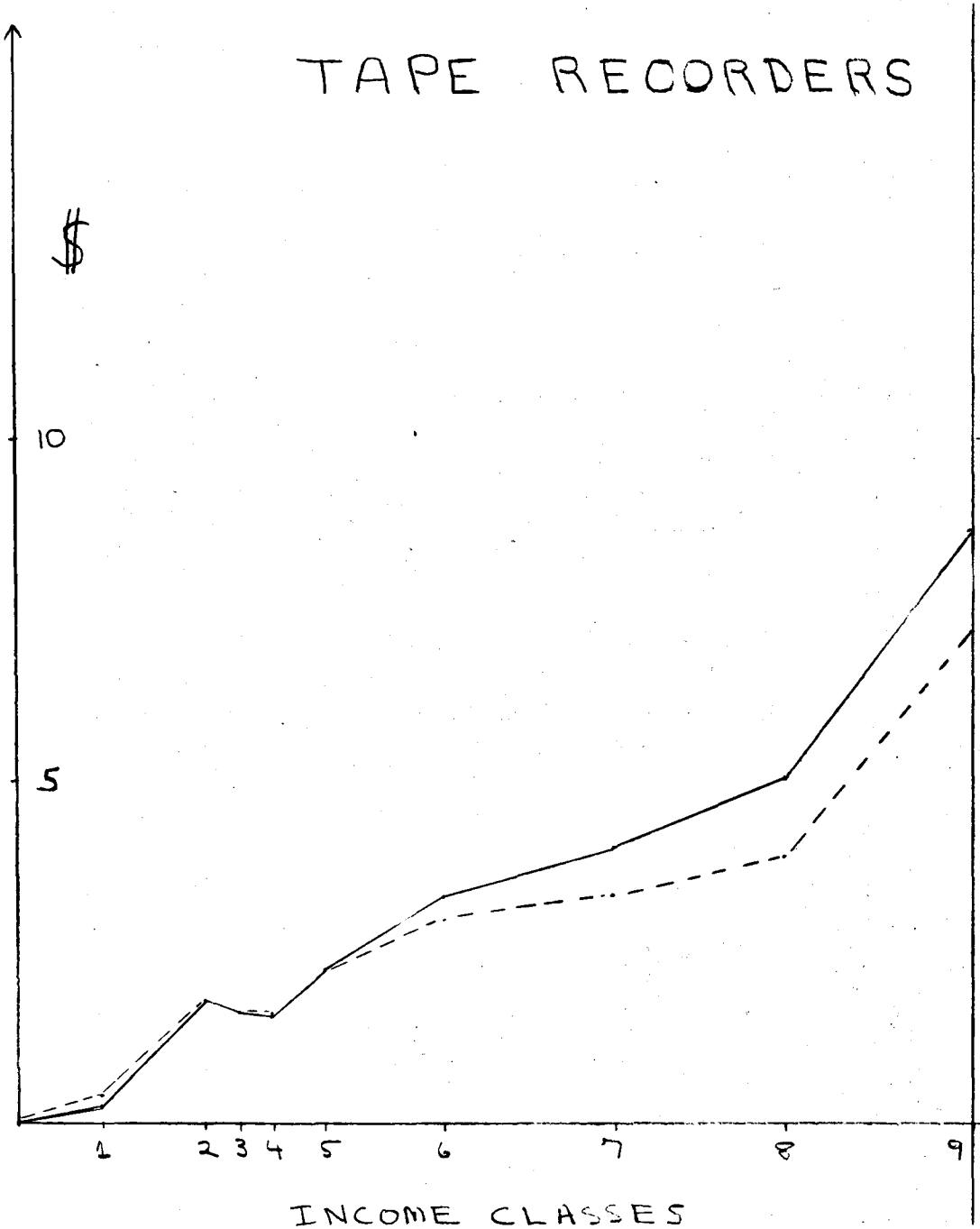
STOCK VAR. = -48.46
t = 21.10



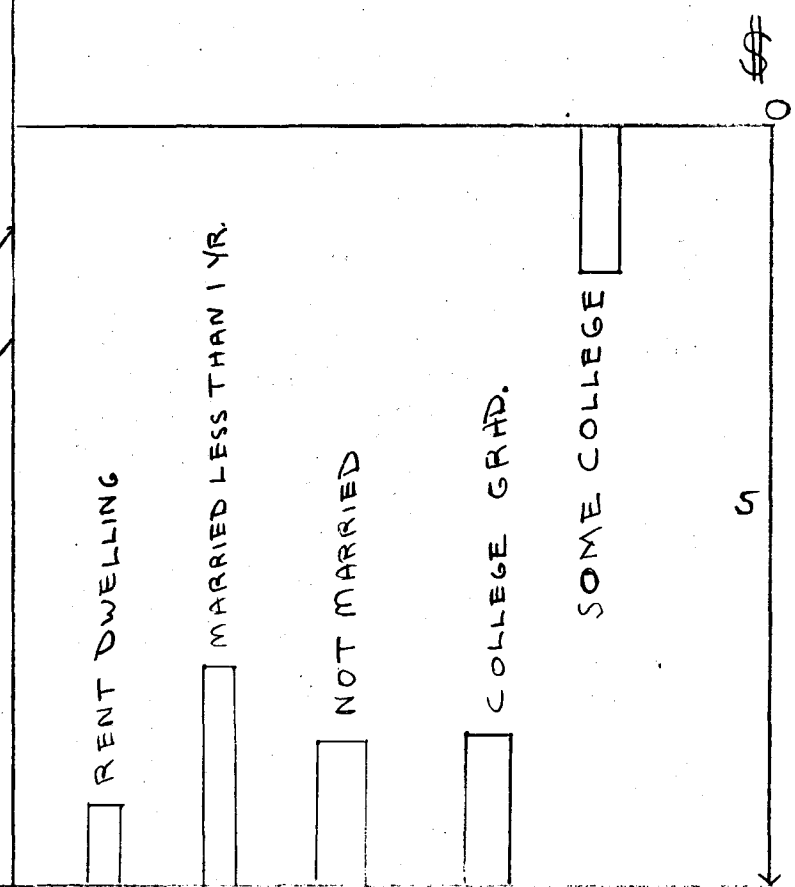

COLOR TV
$
300
250
200
150
100
0 1 2 3 4 5 6 7 8 9
INCOME CLASSES
STOCK VAR. = − 128.18
t = 16.07
$
0
50
100
MARRIED LESS THAN 4 YR.
COLLEGE GRAD.
SOME COLLEGE
OTHER DWELLING
WEST
MARRIED 21 TO 25 YRS.
2 FAMILY HOME
MULTIPLE ROW HOME


BLACK & WHITE TV
$
30
20
10
0 1 2 3 4 5 6 7 8 9
INCOME CLASSES
STOCK VAR.= -21.66
t = 22.81
$
0
5
10
MARRIED LESS THAN 1 YR.
NON-WHITE
SMALL MET. AREA - OUTSIDE C.C.
MARRIED 16 TO 20 YRS.
NON-MET. AREA

TAPE RECORDERS

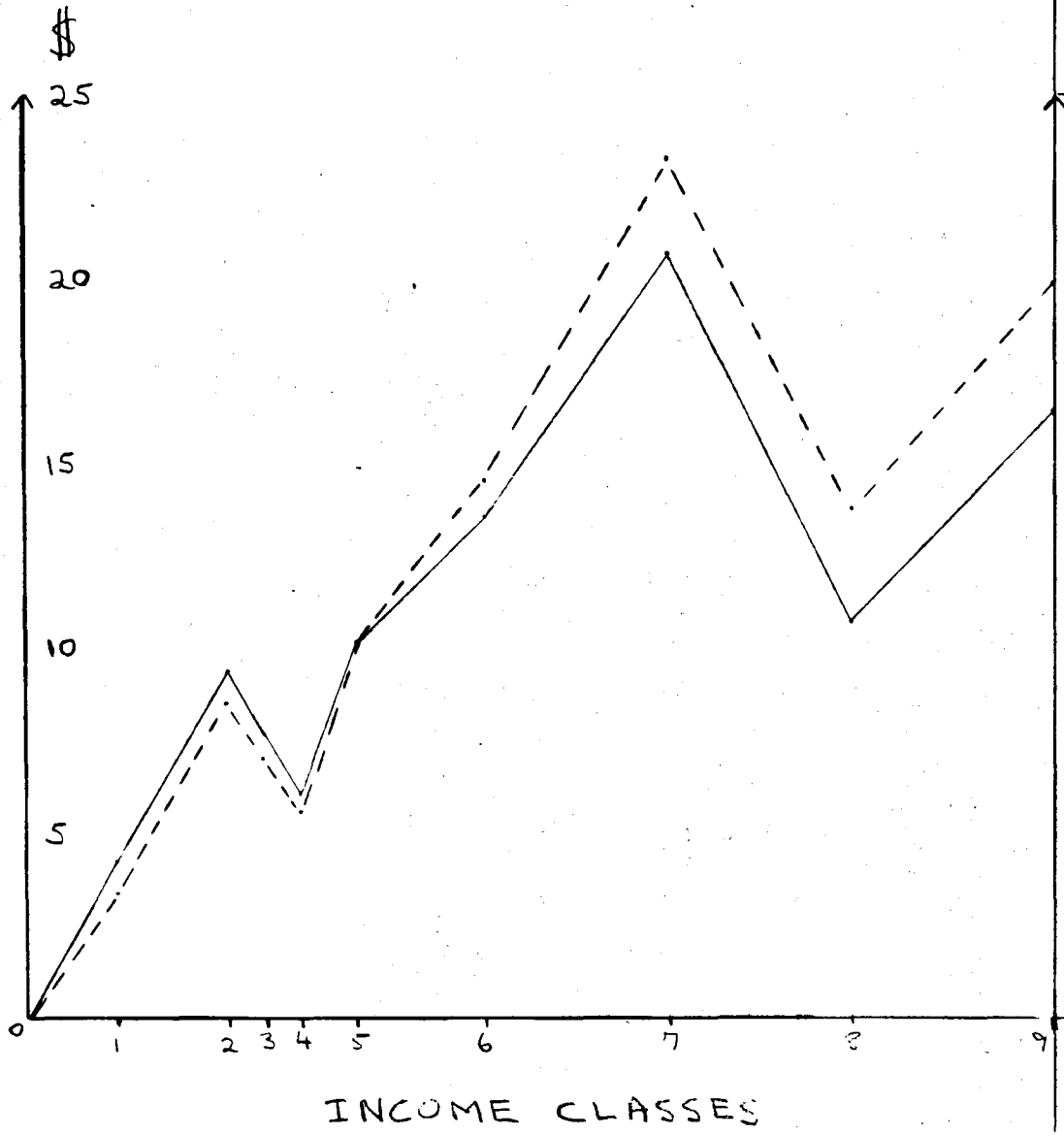


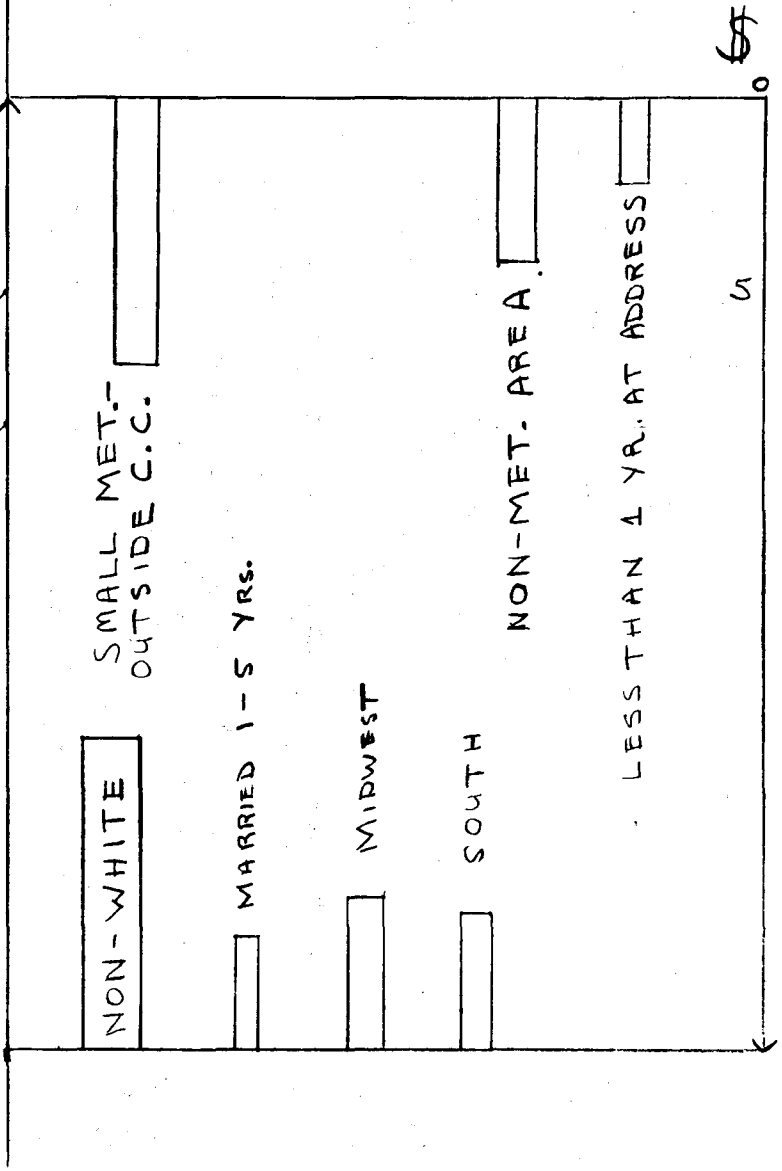
STOCK VAR. - INSIGNIFICANT



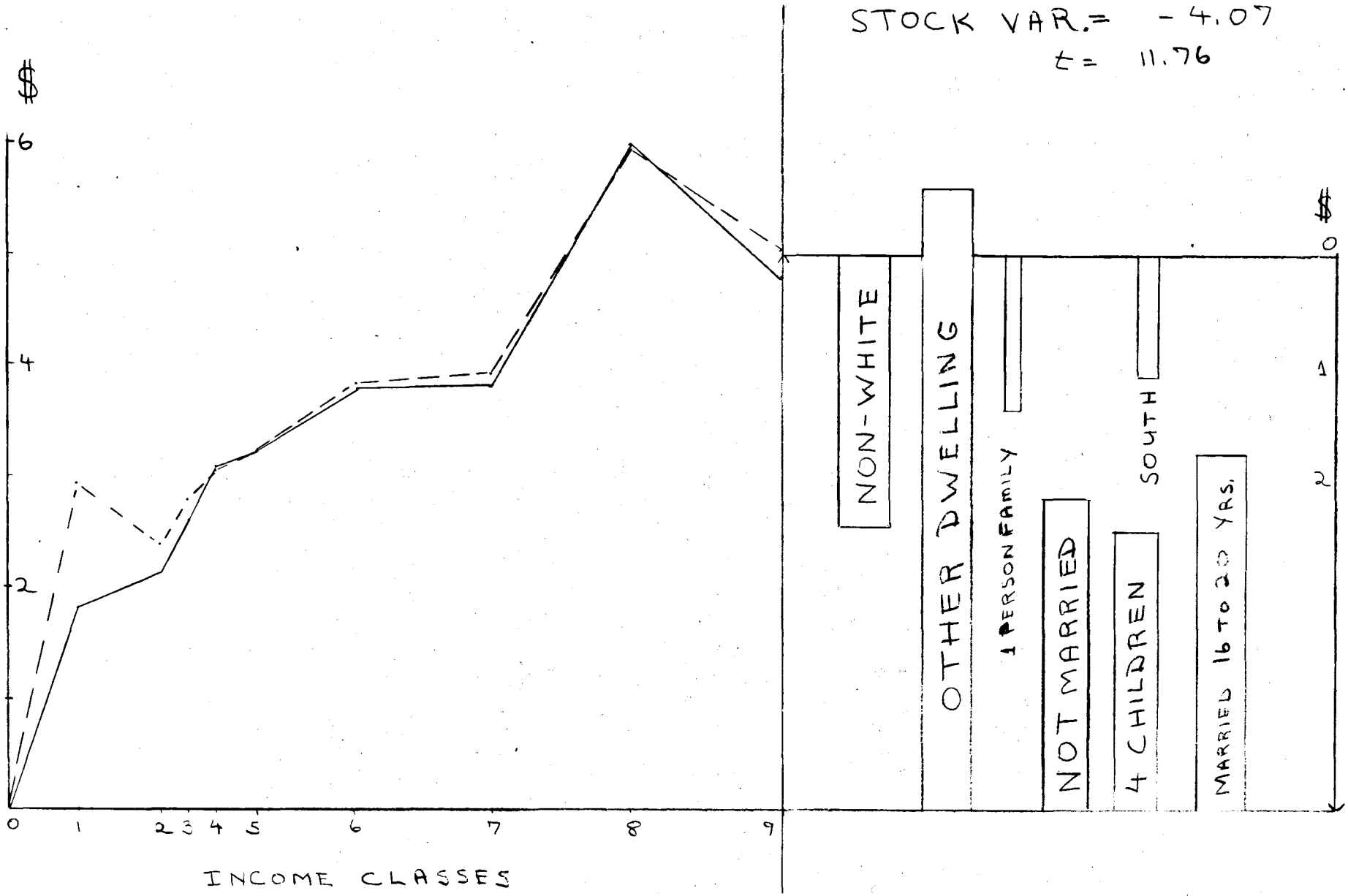
CONSOLE PHONOGRAPHS

STOCK VAR. = -13.71
t = 10.20

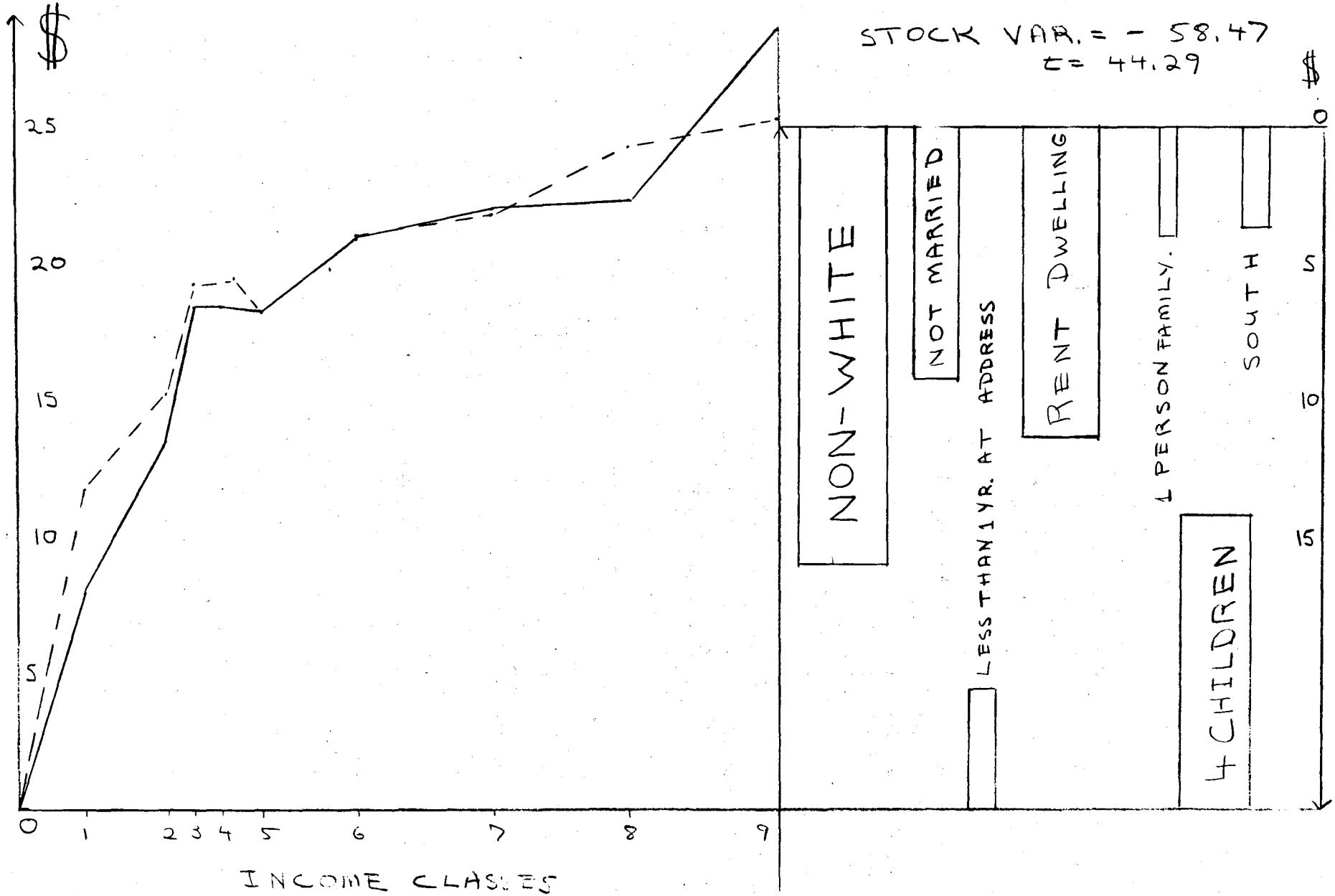




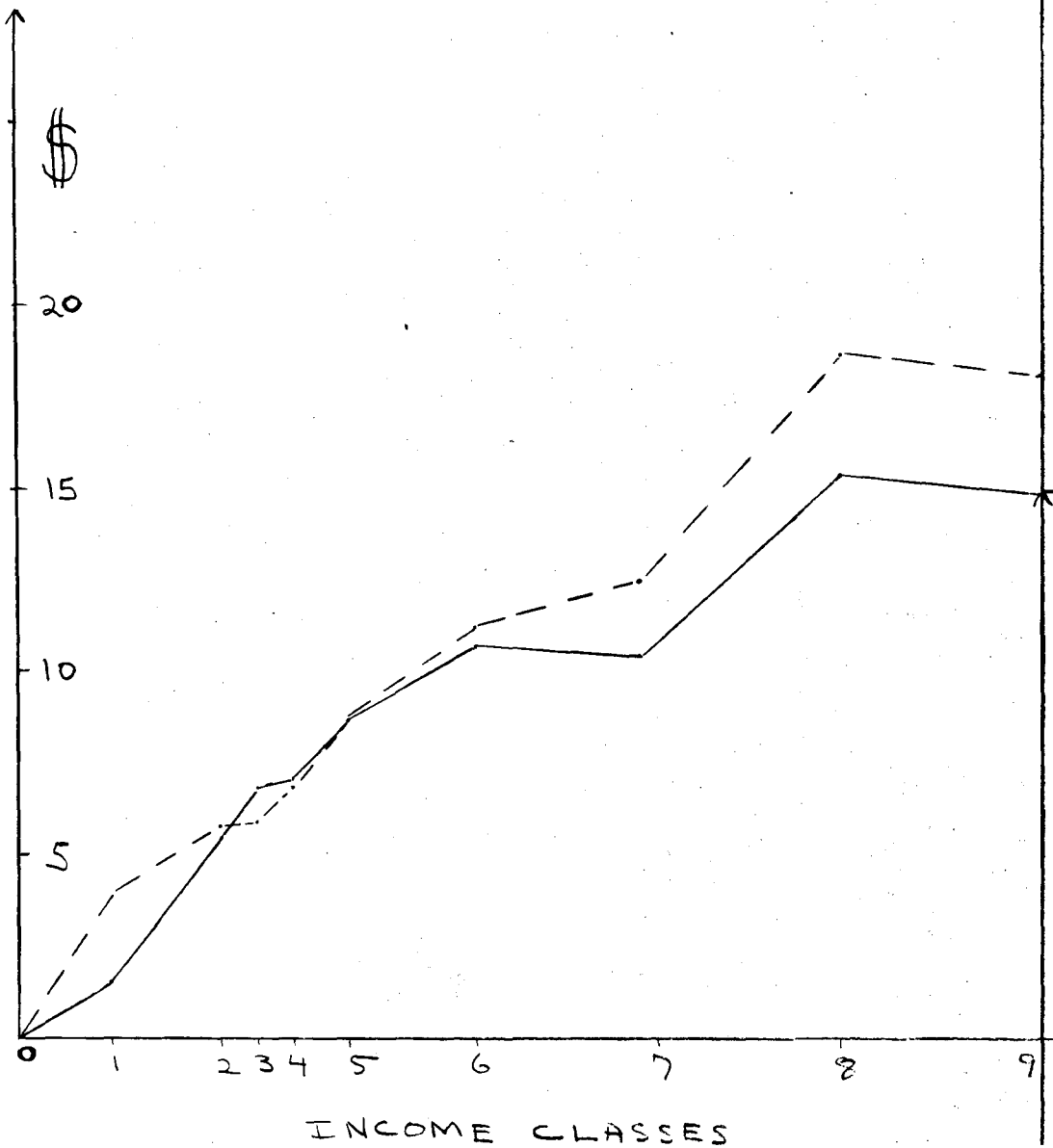
PORTABLE PHONOGRAPHS



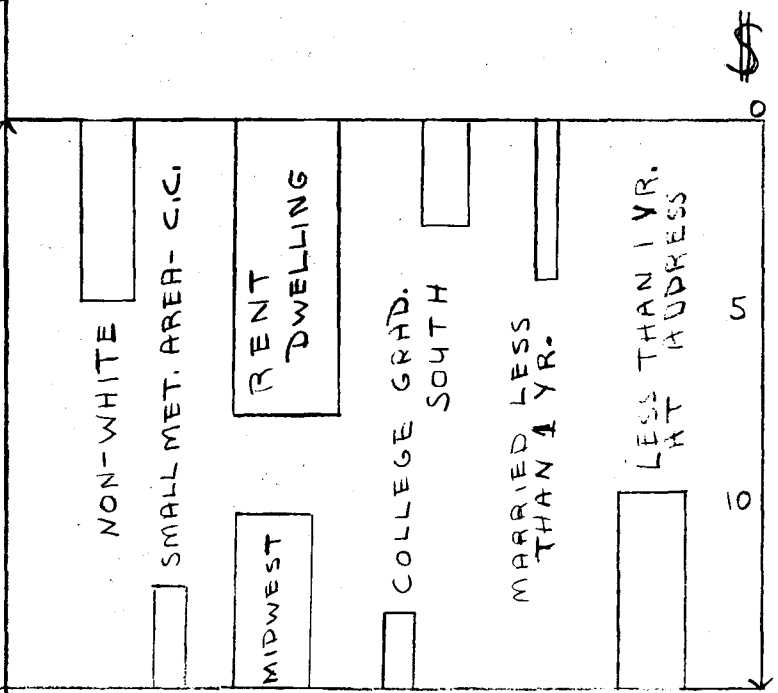
WASHING MACHINES



DRYERS

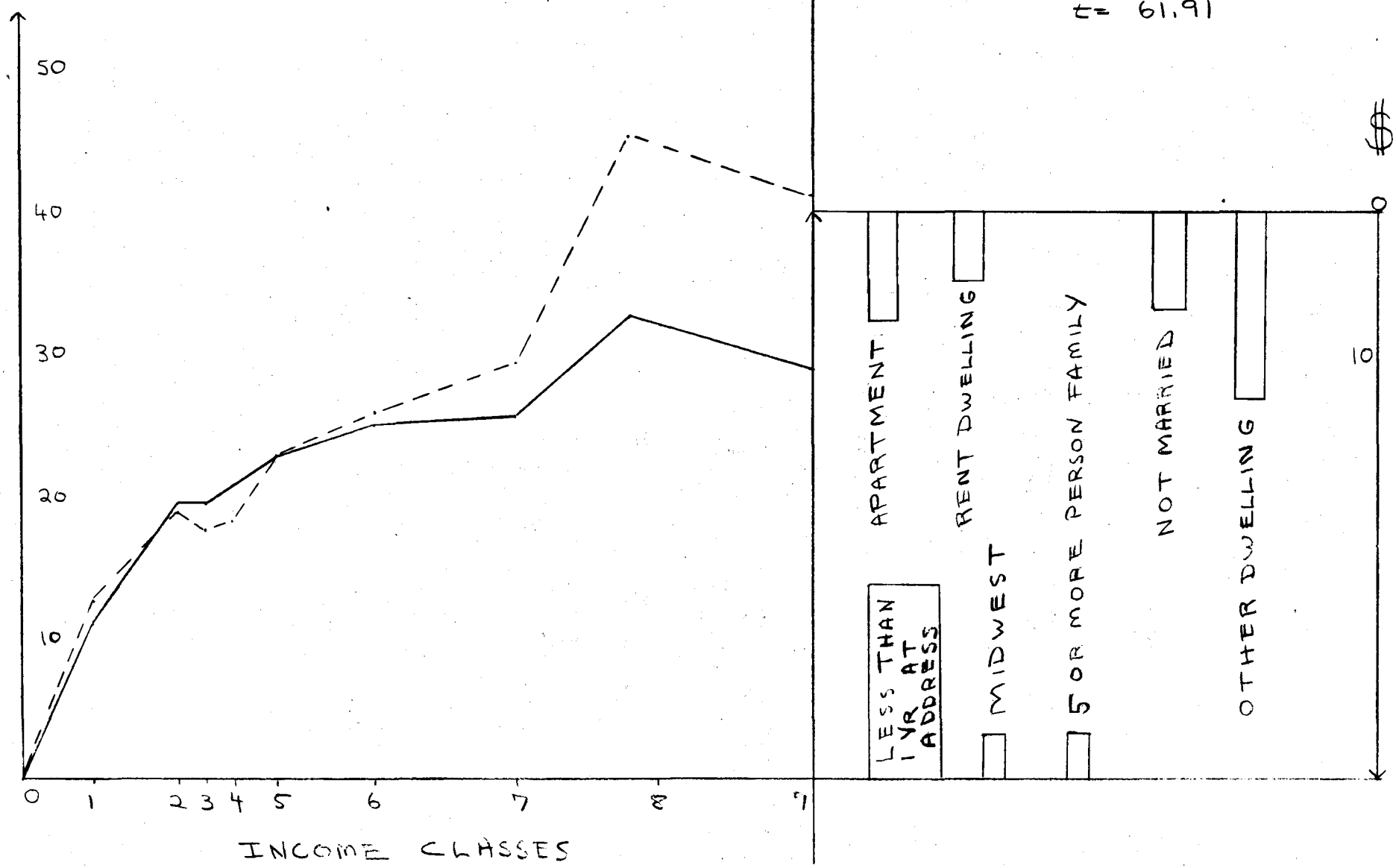


STOCK VAR. = - 19.11
t = 21.82

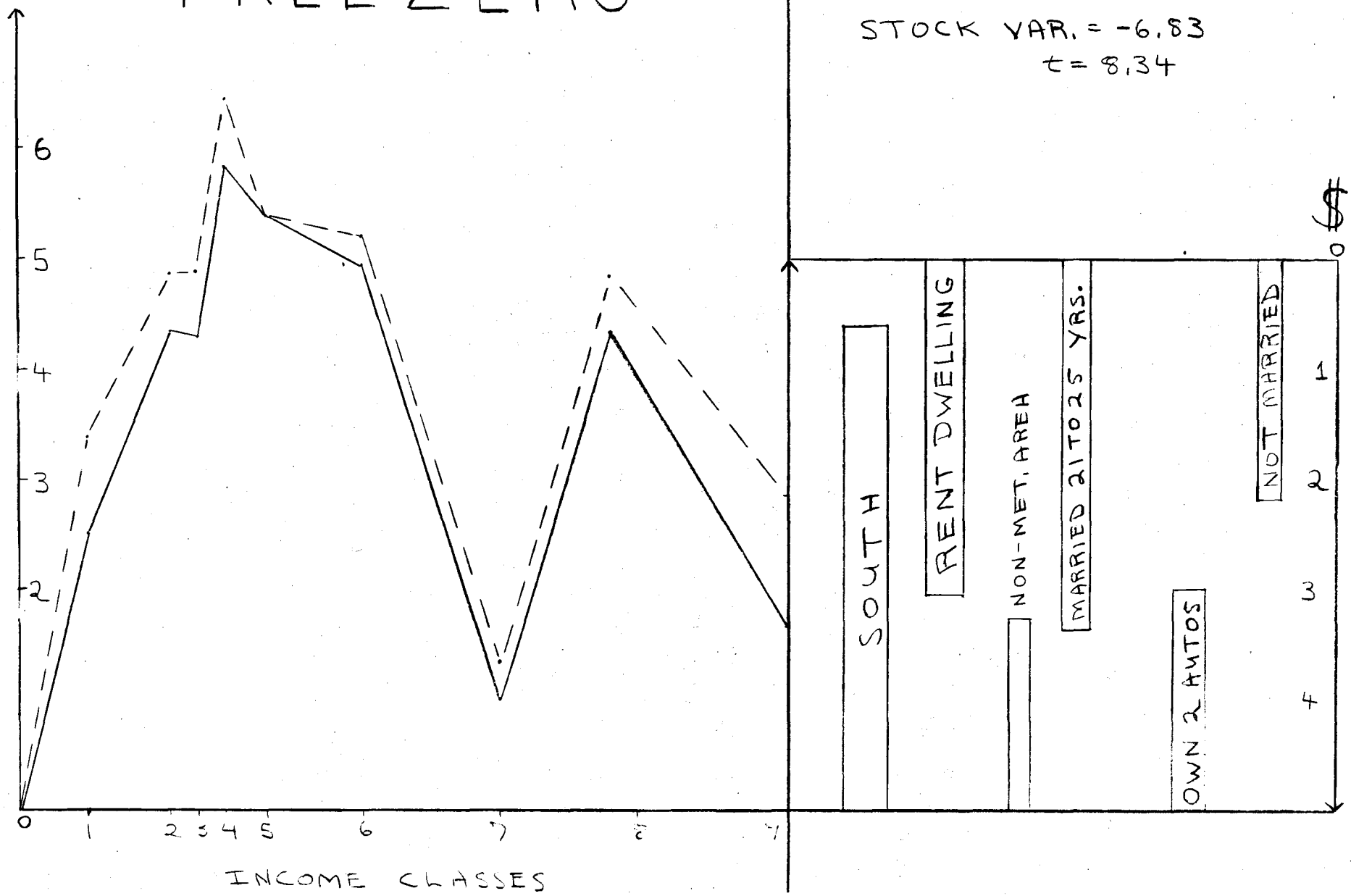


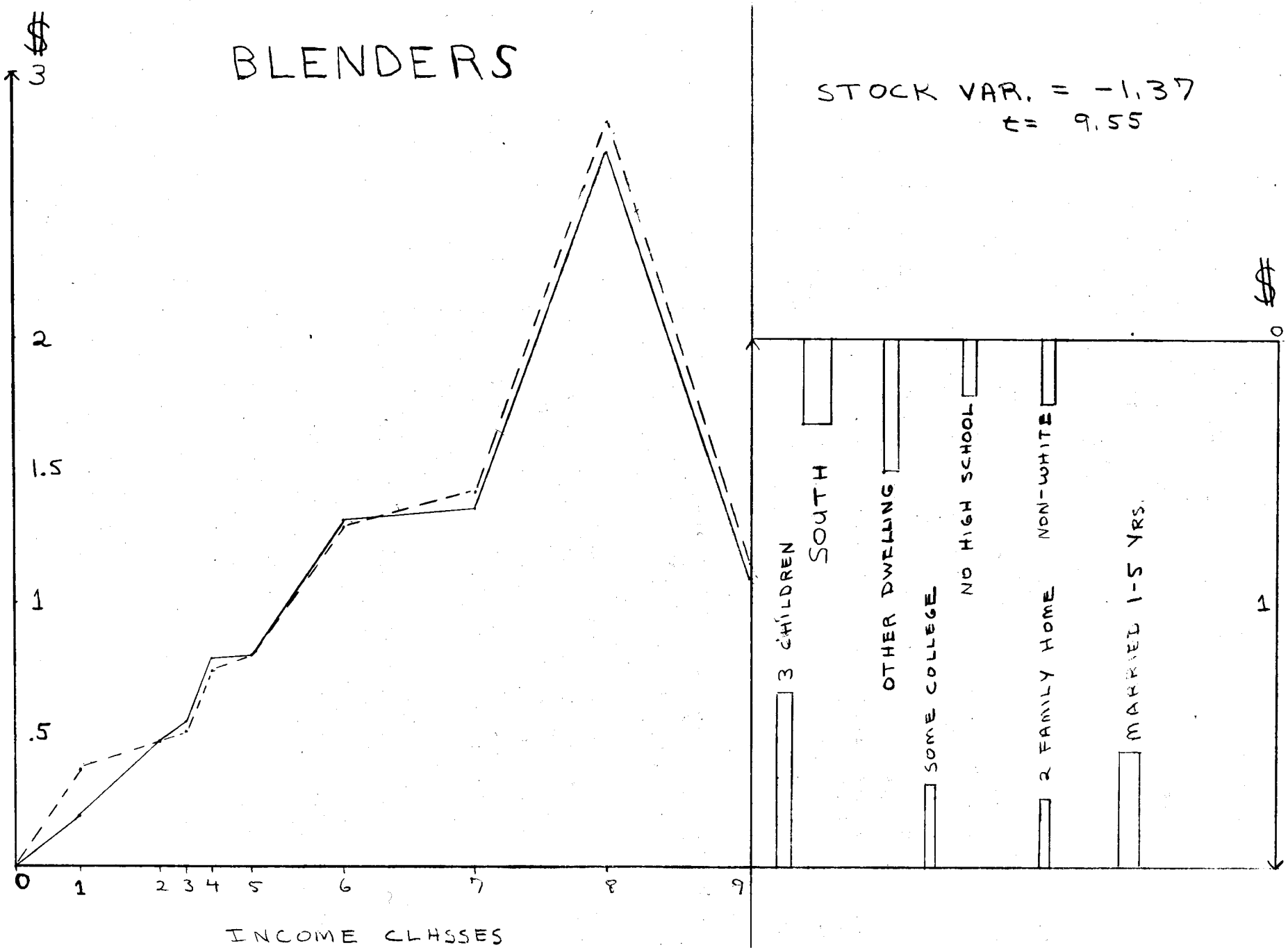
REFRIGERATORS

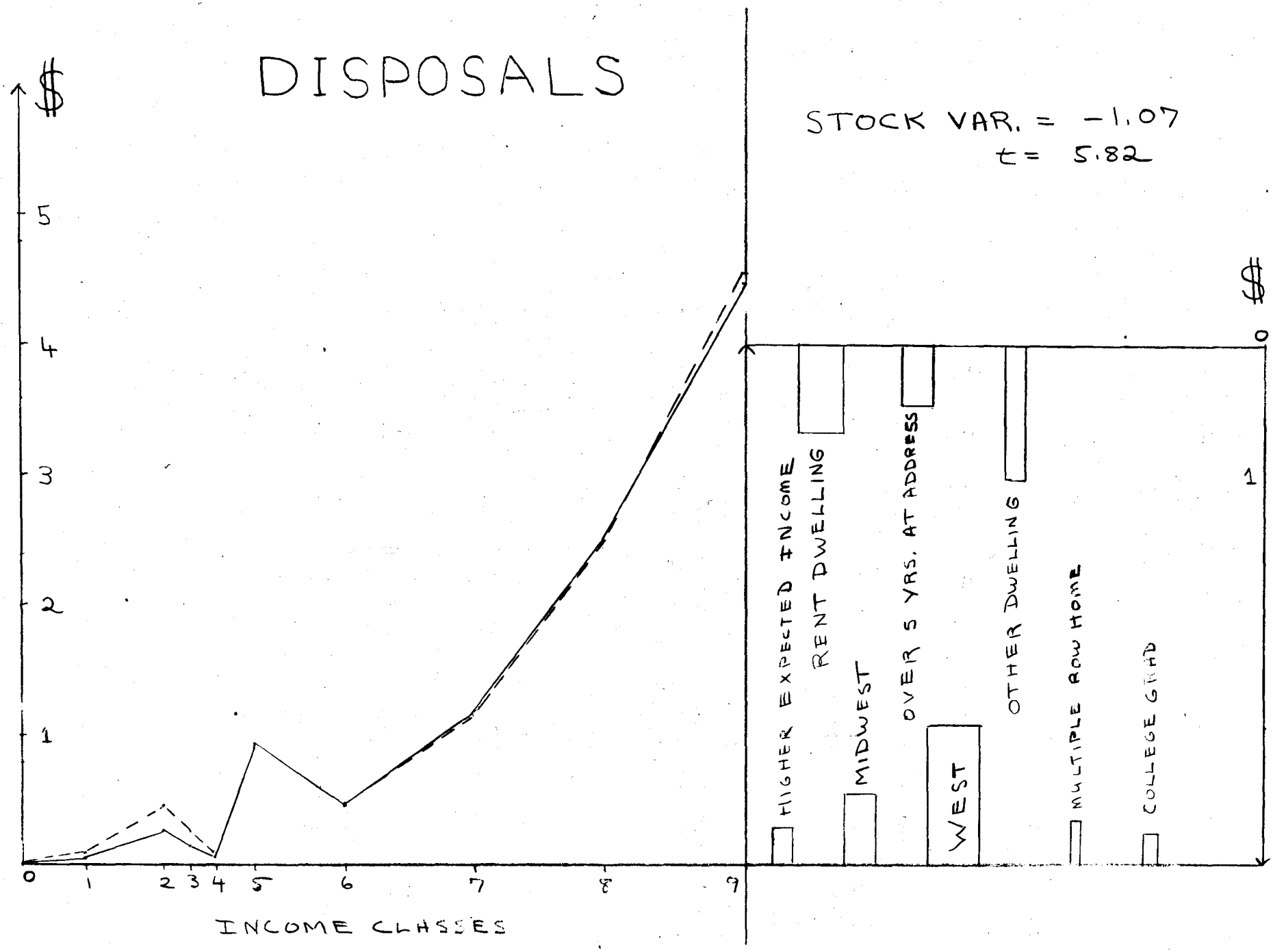
STOCK VAR. = - 124.84
t= 61.91



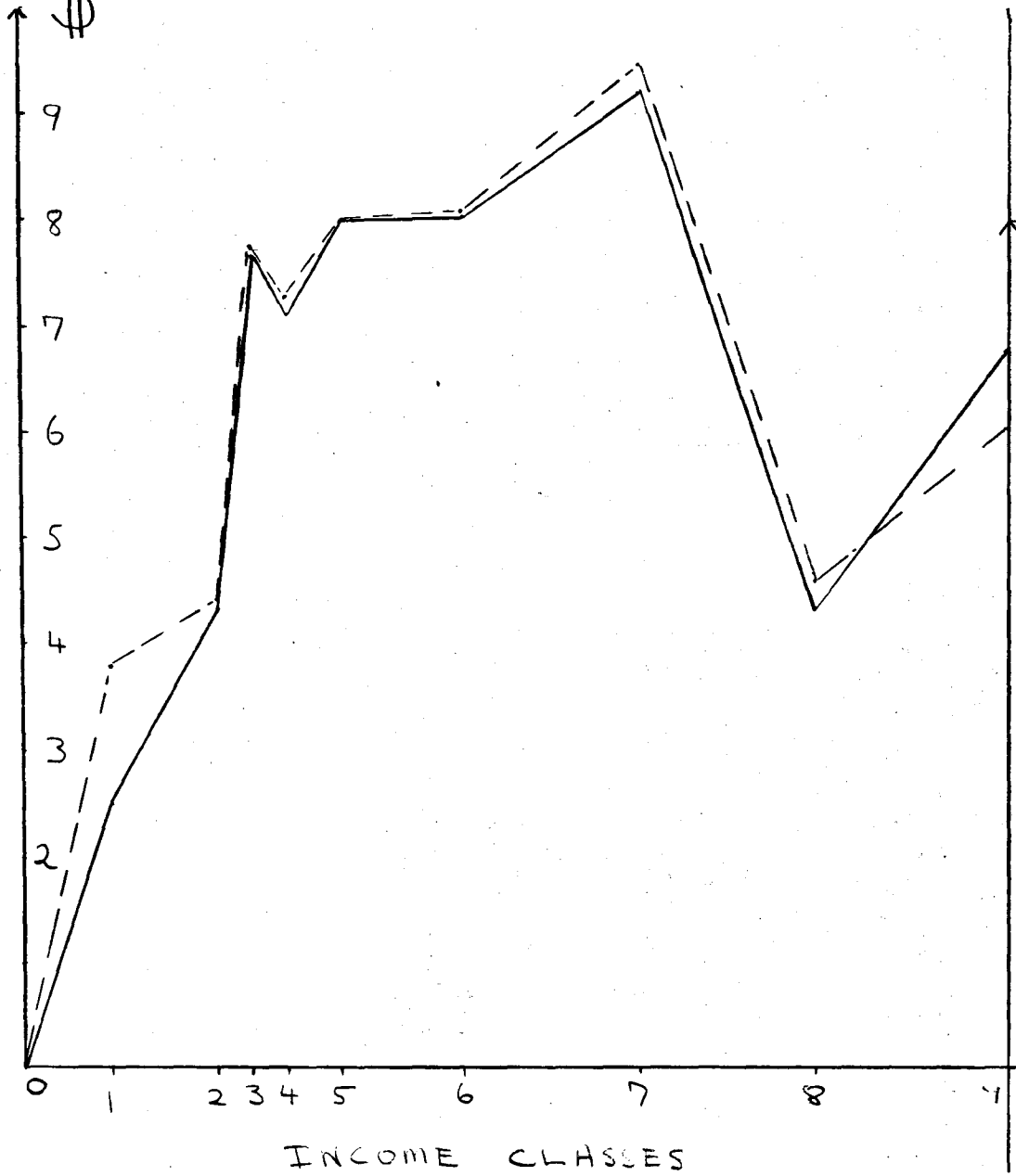
FREEZERS


STOCK VAR. = -6.83
t = 8.34
6
5
4
3
2
0
1 2 3 4 5 6 7 8 9
INCOME CLASSES
$
0
1
2
3
4
SOUTH
RENT DWELLING
NON-MET. AREA
MARRIED 21 TO 25 YRS.
OWN 2 AUTOS
NOT MARRIED


BLENDERS
$
3
2
1.5
1
.5
0
1
2 3 4
5
6
7
8
9
INCOME CLASSES
STOCK VAR. = -1.37
t= 9.55
$
0
1
3 CHILDREN
SOUTH
OTHER DWELLING
SOME COLLEGE
NO HIGH SCHOOL
2 FAMILY HOME
NON-WHITE
MARRIED 1-5 YRS.


DISPOSALS
$
5
4
3
2
1
0
1
2
3
4
5
6
7
8
9
INCOME CLASSES
STOCK VAR. = -1.07
t = 5.82
$
0
1
HIGHER EXPECTED INCOME
RENT DWELLING
MIDWEST
OVER 5 YRS. AT ADDRESS
WEST
OTHER DWELLING
MULTIPLE ROW HOME
COLLEGE GRAD

ROOM AIR CONDITIONERS



STOCK VARIABLE - INSIGNIFICANT