Appendix C

THE PURPOSE OF this appendix is to report on some of the modifications in the analysis of the 1960 BLS data discussed in Chapter 4. These cover (1) disaggregating the "goods" into two items for a total of three components of expenditures: services, perishables, and nonperishable goods; (2) dealing with zero expenditures for various market goods; (3) including a race variable in the Engel curves; (4) replacing the education variable with three education dummies; (5) using total current consumption expenditures as the income variable; and (6) considering the relationship between the estimated income elasticities and education coefficients.

1. PERISHABLES-NONPERISHABLES

The Engel curve was fitted to total expenditures for perishables (tobacco, food at home, and alcohol) and for nonperishables (housing, housefurnishings and equipment, clothing, reading and automobiles) with the double-log equation discussed in Chapter 4. The results are shown below (with t values in parentheses):

Item	ln *Consumption*	ln *Education*	*Age*	*Family Size*	*Region*	$\bar{R}^2$
Perishables	0.711 (20.44)	−0.237 (−5.69)	−0.002 (−1.33)	0.101 (4.13)	−0.136 (−8.00)	.97
Nonperishables	1.03 (49.68)	0.011 (0.45)	−0.004 (−4.15)	−0.004 (−0.27)	−0.011 (−1.06)	.99

Evidently, when perishable goods, nonperishable goods, and services are considered as three separate items, the luxury item (services) has a positive education effect, the necessity item (perishables) has a negative education effect, while the third item's income elasticity is not significantly different from unity (a t value of 1.45) and its education elasticity is not statistically different from zero. Thus, this grouping of total expenditures is qualitatively consistent with the neutrality model

of Chapter 2. Notice that the income elasticity of this more homogeneous goods item is approximately unity, and that neither family size nor the region dummy effect is significant. Notice, too, that age is again observed to affect the durables category.

2. ZERO VALUES

The problem caused by an observation containing zero expenditure for an item is a perplexing one. A zero expenditure for an observation is an extreme value in any of the scatter diagrams. Clearly the fit is improved if the observation is deleted, but in the process some information is lost. On the other hand, allowing a few extreme points to dictate a linear form for the regression seems an excessive adjustment. One alternative is to use some statistical technique other than least squares to estimate a relationship.[1] Another procedure is to replace the zeros by a small value, which is the procedure followed in this study—the zero expenditure per year is replaced in Chapter 4 by an expenditure of one dollar per year. The frequency with which the zero values occur for the dozen or so items discussed in Chapter 4 ranges from eleven observations (7 per cent) to none, as shown in the following table:

Item	*Number of Zero Values*	*Item*	*Number of Zero Values*
Education	11		
Tobacco	6	Automobiles	1
Alcohol	5	Utilities	1
Travel (not auto)	4	Housefurnishings	1
Food (away)	2	Recreation	1
Food (home)	1	All others	0

However, it is clearly arbitrary what value is used to replace the zeros. To determine how sensitive the results were to the value used, regressions were rerun using 0.001 (one tenth of a cent per year). There were no appreciable effects on composite items that did not

[1] For example, probit analysis is one alternative. A comparison of results of Engel curve estimates from the 1960 BLS data using probit and least squares is the subject of Mei-chu Wang's dissertation "Problems in the Estimation of Indifference Surfaces Due to Non-Negativity Constraints," University of Rochester, 1967.

have a 0.001 as a final dependent variable.[2] The greatest effects, of course, were on the three items with the most zero values and the smallest average expenditures:

Item	ln *Consumption*	ln *Education*	*Age*	*Family Size*	*Region*	$\bar{R}^2$
Education	1.766	0.542	−0.039	0.529	0.530	.68
	(3.64)	(0.93)	(−1.60)	(1.55)	(2.24)	
Tobacco	0.723	−0.852	−0.034	0.035	−0.044	.70
	(5.32)	(−5.24)	(−5.08)	(0.36)	(−0.66)	
Alcohol	1.666	−0.832	−0.033	−0.248	−0.629	.81
	(9.02)	(−3.76)	(−3.59)	(−1.91)	(−6.97)	

As expected, the use of a more extreme value reduced the $\bar{R}^2$, increased some of the standard errors, and altered some of the coefficients.

The regressions were also run after the observations with zero values were removed. The number of observations for each of these regressions was 157 minus the corresponding number listed in the table above. The regression results using double logs are given in Table C.1 (A). These results more closely resemble those in the text (which set zeros equal to 1.0) than do the results that use zero $\cong$0.001. Since the zeros are extreme values, the $\bar{R}^2$ is improved by omitting those values, as seen by comparing the $\bar{R}^2$'s in Table C.1 with those in Table 1.

When these items were run linearly, the implied elasticities (analogous to those in Table 3) were as follows:

Mean Elasticities, Zero Values Omitted

Item	*Income*	*Education*	*Family Size*	$\bar{R}^2$
Education (expenditure)	2.721	1.389	−1.102	.695
Tobacco	0.527	−0.576	0.198	.818
Alcohol	1.457	−0.357	−0.401	.837
Food (away)	1.451	−0.012	−0.591	.899
Food (home)	0.526	−0.117	0.548	.950
Housefurnishings	1.031	0.038	0.233	.924

These mean elasticities are again similar to those shown in the text which include the zero values (as zeros). Table C.1 (B) gives the

[2] For example, a household which spent zero for utilities and $500.00 for shelter had as the dependent variable 501.00 versus 500.001, with no appreciable effect on the coefficients. This was also true for the goods-services dichotomy.

TABLE C.1
Regression Equations With Zero Values Removed

(A)
Constant Income Elasticity (with ln E)[a]

	ln *Consumption*	ln *Education*	*Age*	*Family Size*	*Region*	$\bar{R}^2$
Education	1.557	1.693	0.031	0.520	0.434	.878
	(7.08)	(6.30)	(2.75)	(3.36)	(4.09)	
Tobacco	0.756	−0.803	−0.034	−0.012	−0.009	.870
	(9.45)	(−8.33)	(−8.39)	(−0.21)	(−0.22)	
Alcohol	1.586	−0.463	−0.019	−0.198	−0.550	.910
	(13.78)	(−3.31)	(−3.26)	(−2.44)	(−9.74)	
Food (away)	1.196	0.250	0.004	−0.074	0.072	.941
	(17.49)	(3.05)	(1.18)	(−1.54)	(2.17)	
Food (home)	0.676	−0.213	−0.001	0.108	−0.122	.970
	(21.23)	(−5.58)	(−0.43)	(4.78)	(−7.84)	
Housefurnishings	1.011	−0.092	−0.009	0.139	0.127	.946
	(14.65)[b]	(−1.12)	(−2.67)	(2.84)	(3.78)	

(B)
Constant Income Elasticity (with E)[a]

	ln *Consumption*	*Education*	*Age*	*Family Size*	*Region*	($\bar{\epsilon}_E$)	$\bar{R}^2$
Education	1.579	0.169	0.025	0.495	0.376	(1.705)	.884
	(7.72)	(7.06)	(2.56)	(3.50)	(3.60)		
Tobacco	0.741	−0.080	−0.031	0.005	0.019	(−0.797)	.879
	(10.01)	(−9.24)	(−8.76)	(0.09)	(0.49)		
Alcohol	1.609	−0.052	−0.019	−0.213	−0.530	(−0.518)	.914
	(14.82)	(−4.06)	(−3.66)	(−2.83)	(−9.46)		
Food (away)	1.241	0.017	0.001	−0.111	0.070	(0.174)	.940
	(18.64)	(2.23)	(0.30)	(−2.39)	(2.04)		
Food (home)	0.670	−0.021	0.000	0.113	−0.115	(−0.209)	.970
	(22.20)	(−5.94)	(0.15)	(5.39)	(−7.40)		
Housefurnishings	0.994	−0.006	−0.008	0.153	0.128	(−0.062)	.946
	(14.96)[b]	(−0.80)	(−2.57)	(3.32)	(3.76)		

[a] t values are in parentheses.

[b] Coefficient not statistically different from one (the t values for testing the difference from unity are 0.16 for housefurnishings in (A) and −0.10 in (B)).

results obtained from entering the education variable linearly while excluding the zero value observations.

The results shown here suggest that relatively few important changes occur in the coefficients, their statistical significance, or the coefficient of determination as a consequence of removing the zero values rather than replacing them by a value of 1 (ln (1) = 0) in the log regressions, or leaving them in as zeros in the linear regressions. There are greater differences, however, if the zeros are replaced by some small value such as 0.001.

3. THE EFFECT OF RACE

Another environmental variable which could have an independent effect on consumption is race. Differences in the quality of schooling may appreciably affect the stock of human capital, and if such differences are systematically related to race, the race variable may act as a proxy for school quality and thereby affect expenditure patterns. Further, if different ethnic groups within the economy have significantly different environments, this could also have an effect on behavior. To determine whether race has a discernible independent and systematic effect on expenditure patterns, the percentage of nonwhite households in each cell was included in the Engel curves. This variable was included both as a replacement for the education variable and in addition to it. The results for the goods-services dichotomy are given below (with t values in parentheses).

Item	ln *Consumption*	ln *Education*	*Per Cent Nonwhite*	*Age*	*Family Size*	*Region*	$\bar{R}^2$
Goods	0.884		−0.001	−0.001	0.061	−0.045	.995
	(59.94)		(−0.98)	(−2.06)	(7.48)	(−5.71)	
Goods	0.920	−0.078	−0.001	−0.003	0.030	−0.040	.996
	(57.66)	(−4.56)	(−1.82)	(−4.87)	(2.90)	(−5.25)	
Services	1.235		0.001	+0.000	−0.134	0.105	.986
	(42.48)		(0.70)	(0.18)	(−8.43)	(6.71)	
Services	1.143	0.200	0.002	0.006	−0.056	0.091	.989
	(38.18)	(6.20)	(1.85)	(4.80)	(−2.90)	(6.38)	

The substitution of the race variable for the education level results in a poorer fit, reduced significance for several of the variables, and a movement away from one in income elasticity.[3] Including the race variable in addition to education, on the other hand, improves the fit slightly and increases the significance of the education, age, and family-size coefficients. The race variable itself is also significant, and suggests that the larger the fraction of nonwhites, ceteris paribus, the larger the share of services in total expenditures.

The results for the thirteen separate items are given in Tables C.2 and C.3. Replacing the education variable with the race variable generally reduces the $\bar{R}^2$ and shifts the income elasticity away from one. Including both education and the nonwhite percentage generally

[3] This is as one would expect with the new variable as a poorer proxy for nonmarket efficiency, since the income variable picks up some of the consumption income effect.

TABLE C.2
Expenditure Functions With Race Variable Replacing Education, 1960 BLS Data[a]

Dependent Variable	ln *Consumption*	*Per Cent Nonwhite*	*Age*	*Family Size*	*Region*	$\bar{R}^2$
Food (home)	0.5957	0.0032	0.0062	0.1952	−0.1571	.937
	(13.05)	(1.76)	(4.06)	(7.80)	(−6.39)	
Food (away)	1.3121	−0.0016	−0.0039	−0.1775	0.0951	.924
	(17.26)	(−0.54)	(−1.52)	(−4.26)	(2.32)	
Tobacco	0.5019	0.0132	−0.0093	0.2810	−0.1284	.788
	(4.95)	(3.28)	(−2.75)	(5.05)	(−2.35)	
Alcohol	1.5384	0.0153	−0.0069	−0.0588	−0.6822	.895
	(12.47)	(3.13)	(−1.68)	(−0.87)	(−10.27)	
Shelter	0.8128	−0.0044	−0.0017	−0.1251	−0.1265	.960
	(22.06)	(−2.99)	(−1.39)	(−6.19)	(−6.38)	
Household operations	1.3158	0.0018	−0.0025	−0.2329	0.1866	.971
	(31.84)	(1.11)	(−1.83)	(−10.27)	(8.39)	
Housefurnishings and equipment	1.0472	0.0061	−0.0065	0.1537	0.0830	.939
	(14.64)	(2.14)	(−2.70)	(3.92)	(2.16)	
Clothing	1.2782	0.0046	−0.0048	0.0595	0.0821	.983
	(31.44)	(2.89)	(−3.50)	(2.67)	(3.75)	
Personal care	0.9379	0.0042	−0.0066	0.0276	0.1313	.974
	(25.38)	(2.86)	(−5.35)	(1.36)	(6.60)	
Medical care	0.8224	−0.0016	0.0072	0.0075	0.0427	.888
	(13.14)	(−0.63)	(3.42)	(0.22)	(1.27)	
Leisure	1.3375	−0.0026	−0.0114	−0.0766	−0.0416	.976
	(26.44)	(−1.30)	(−6.74)	(−2.76)	(−1.53)	
Education	2.4120	−0.0001	−0.0254	−0.1348	0.4939	.851
	(10.10)	(−0.01)	(−3.19)	(−1.03)	(3.86)	
Travel	1.1607	0.0003	−0.0157	0.1497	0.0560	.951
	(14.83)	(0.09)	(−6.00)	(3.49)	(1.33)	

[a] *t* values are in parentheses.

TABLE C.3

Expenditure Functions With Race Variable and Education, 1960 BLS Data[a]

Dependent Variable	ln *Consumption*	ln *Education*	*Per Cent Nonwhite*	*Age*	*Family Size*	*Region*	$\bar{R}^2$
Food (home)	0.6695	−0.1607	0.0023	0.0013	0.1317	−0.1455	.940
	(13.06)	(−2.91)	(1.27)	(0.56)	(4.02)	(−5.99)	
Food (away)	1.2189	0.2027	−0.0005	0.0024	−0.0974	0.0805	.926
	(14.11)	(2.18)	(−0.15)	(0.62)	(−1.77)	(1.96)	
Tobacco	0.8517	−0.7611	0.0089	−0.0328	−0.0197	−0.0735	.839
	(8.36)	(−6.94)	(2.49)	(−7.29)	(−0.30)	(−1.52)	
Alcohol	1.7762	−0.5172	0.0124	−0.0228	−0.2632	−0.6449	.902
	(12.97)	(−3.51)	(2.59)	(−3.78)	(−3.01)	(−9.93)	
Shelter	0.7154	0.2120	−0.0032	0.0048	−0.0413	−0.1418	.966
	(18.16)	(5.00)	(−2.30)	(2.77)	(−1.64)	(−7.59)	
Household operations	1.1623	0.3339	0.0037	0.0077	−0.1009	0.1625	.979
	(28.76)	(7.68)	(2.62)	(4.34)	(−3.91)	(8.48)	
Housefurnishings and equipment	1.0599	−0.0275	0.0059	−0.0073	0.1429	0.0850	.939
	(12.84)	(−0.31)	(2.05)	(−2.01)	(2.71)	(2.17)	
Clothing	1.2778	0.0007	0.0046	−0.0047	0.0598	0.0820	.983
	(27.23)	(0.01)	(2.84)	(−2.30)	(1.99)	(3.69)	
Personal care	0.9861	−0.1048	0.0036	−0.0099	−0.0138	0.1389	.974
	(23.52)	(−2.32)	(2.45)	(−5.33)	(−0.52)	(6.99)	
Medical care	0.8122	0.0223	−0.0014	0.0079	0.0163	0.0411	.887
	(11.24)	(0.29)	(−0.57)	(2.46)	(0.35)	(1.20)	
Leisure	1.2746	0.1368	−0.0018	−0.0072	−0.0225	−0.0515	.977
	(22.18)	(2.21)	(−0.91)	(−2.85)	(−0.61)	(−1.89)	
Education	1.7082	1.5314	0.0086	0.0217	0.4703	0.3835	.877
	(6.88)	(5.73)	(0.99)	(1.98)	(2.96)	(3.26)	
Travel	1.3574	−0.4279	−0.0021	−0.0289	−0.0193	0.0869	.957
	(16.10)	(−4.72)	(−0.72)	(−7.78)	(−0.36)	(2.17)	

[a] t values are in parentheses.

improves the $\bar{R}^2$, has little effect on the magnitude of the income or education coefficient, but tends to reduce their significance somewhat. There was no appreciable effect upon the relationship between these two elasticities. The education and race effects seem to be in opposite directions for these detailed items—in only one case (household operations) are their signs the same, with both statistically significant; however, the negative correlation is certainly not a strong one and these results are not inconsistent with a hypothesis of no relationship between the two variables. Likewise, there is no apparent relationship between the race variable and the income elasticity (with education held fixed), as seen from the following two-way diagram:

Per Cent Nonwhite	*Income Elasticity*		
	$\eta > 1$	$\eta \simeq 1$	$\eta < 1$
(+)	Alcohol Household operations Clothing	Housefurnishings and equipment Personal care	Tobacco
(0)	Food (away) Leisure Education Travel		Food (home) Medical care
(−)			Housing

From these results, one concludes that there appears to be no systematic relationship between the effect on expenditures of this race variable and either income or education. This tentative conclusion, however, is subject to the qualification in footnote 16 of Chapter 4.

4. EDUCATION DUMMIES

Most of the regressions discussed so far have assumed that the effect of education on productivity and on the derived demand is proportionate at all levels of education—that the education coefficient is a constant elasticity. The procedure of using double logs also implies that small changes in the education level will have a distinguishable and continuous effect. To relax these conditions somewhat the $(\ln E)$ variable was replaced by three education dummy variables designed to examine the effect of broader changes in education—from grade

school to high school, from high school to college, et cetera—and to see if the relative effects diminish or increase as education rises.

The four education classes were grade school (0 through 8 years), high school (8–12), college (13–16), and postgraduate (17+), and the three dummies were defined to show the *marginal* effects of each category. Thus, E_1 shows the effect of having a high school education relative to having only grade school training, E_2 shows the effect of college relative to high school, and E_3 shows the effect of postgraduate work relative to college. The effect of, say, college relative to grade school would be $(E_1 + E_2)$, and so forth. The results for the goods-services dichotomy are as follows (with t values in parentheses):

	Item	
Explanatory Variable	*Goods*	*Services*
ln Consumption	0.930 (75.54)	1.137 (49.04)
E_1	−0.034 (−3.26)	0.079 (4.05)
E_2	−0.034 (−4.05)	0.080 (5.03)
E_3	−0.007 (−0.47)	0.017 (0.60)
Age	−0.003 (−4.53)	0.005 (3.91)
Family size	0.026 (2.99)	−0.058 (−3.48)
Region	−0.040 (−5.76)	0.089 (6.85)
$\bar{R}^2$	.996	.989

The effect of this substitution of three dummies for ln E is seen by comparing these equations with those in Chapter 4. There is no important change in the magnitude, sign, or significance of the coefficients of the four other variables. For the education effect itself, the regressions suggest that the magnitude of the effect is quite similar for high school relative to grade school and college relative to high school, while the effect of additional education is small. Similarly, the significance of the first two dummies is roughly the same. In neither case is the post-graduate dummy significant. In all cases the sign is consistent with the neutrality model.

The same regression was run for the thirteen detailed expenditure items. These results are given in Table C.4. Again, the use of the dummies seems to have had little effect on the other variables' coefficients—no significant coefficient changed in sign. The $\bar{R}^2$ is generally increased by the use of the dummies despite the introduction of two additional independent variables; leisure and housefurnishings are the only exceptions.

TABLE C.4
Expenditure Functions With Education Dummy Variables[a]

Dependent Variable	ln *Consumption*	E_1	E_2	E_3	*Age*	*Family Size*	*Region*	$\bar{R}^2$
Household operations	1.1339	0.1178	0.1450	0.0837	0.0054	−0.0983	0.1587	.982
	(38.38)	(4.76)	(7.16)	(2.27)	(3.20)	(−4.62)	(9.56)	
Clothing	1.2334	0.0389	−0.0398	−0.1272	−0.0029	0.0774	0.1181	.983
	(34.22)	(1.29)	(−1.61)	(−2.83)	(−1.39)	(2.98)	(5.83)	
Personal care	0.9480	−0.0242	−0.0731	−0.1325	−0.0084	−0.0011	0.1752	.977
	(30.42)	(−0.93)	(−3.42)	(−3.41)	(−4.72)	(−0.05)	(9.99)	
Medical care	0.8242	−0.0458	0.1007	−0.1219	0.0055	0.0056	0.0255	.892
	(14.85)	(−0.99)	(2.65)	(−1.76)	(1.74)	(0.14)	(0.82)	
Leisure	1.3356	0.0794	0.0104	0.0345	−0.0075	−0.0484	−0.0641	.976
	(29.29)	(2.08)	(0.33)	(0.61)	(−2.86)	(−1.47)	(−2.50)	
Education	1.7749	0.3349	0.7381	0.2502	0.0001	0.3507	0.3278	.884
	(9.36)	(2.11)	(5.68)	(1.06)	(0.01)	(2.57)	(3.08)	
Travel	1.3561	−0.2544	−0.1432	−0.0502	−0.0304	−0.0306	0.1016	.960
	(21.18)	(−4.75)	(−3.26)	(−0.63)	(−8.32)	(−0.66)	(2.86)	
Food (home)	0.6197	−0.0677	−0.0629	−0.0556	0.0023	0.1522	−0.1217	.940
	(15.36)	(−2.01)	(−2.27)	(−1.10)	(1.00)	(5.25)	(−5.37)	
Food (away)	1.2914	0.1670	−0.0089	−0.0833	0.0043	−0.1227	0.0819	.928
	(19.27)	(2.98)	(−0.19)	(−1.00)	(1.13)	(−2.55)	(2.17)	
Tobacco	0.6236	−0.2957	−0.2330	−0.4217	−0.0263	0.0877	0.8238	.845
	(7.94)	(−4.50)	(−4.33)	(−4.31)	(−5.86)	(1.55)	(0.54)	
Alcohol	1.5445	−0.1628	−0.2259	−0.2867	−0.0165	−0.1604	−0.5352	.902
	(14.28)	(−1.80)	(−3.05)	(−2.13)	(−2.67)	(−2.06)	(−8.80)	
Housefurnishings and equipment	0.9715	−0.0676	0.0243	−0.0658	−0.0091	0.1641	0.1187	.938
	(14.79)	(−1.23)	(0.54)	(−0.80)	(−2.44)	(3.47)	(3.21)	
Shelter	0.7916	0.0565	0.0740	0.0563	0.0033	−0.0915	−0.1722	.964
	(24.81)	(3.11)	(3.44)	(1.43)	(0.12)	(−4.38)	(−9.69)	

[a] t values are in parentheses.

Turning to a qualitative comparison of the three education coefficients in relation to the income elasticity and the neutrality model, the findings for the thirteen regressions are summarized as follows:

Education Class	*Number of Items Qualitatively Consistent With the Neutrality Model*
High school dummy (E_1)	10
College dummy (E_2)	6
Postgraduate dummy (E_3)	8
College-grade school ($E_1 + E_2$)	8
Postgraduate-grade school ($E_1 + E_2 + E_3$)	9

The effects of education vis-à-vis the neutrality model appear to be stronger at lower levels of education, but there does not appear to be any systematic change in the magnitude of the education effect from one schooling dummy to another. Comparing these regressions with those in the text, the improvement in the $\bar{R}^2$, which is slight in most cases (with the exception of tobacco), must be weighed against the convenience of the double-log form, from which the constant education elasticities allow us to estimate the elasticity of consumption income.

5. TOTAL CURRENT CONSUMPTION EXPENDITURE

As was mentioned briefly in Chapter 4 in connection with the goods-services dichotomy, the total consumption expenditure item includes both durable goods and the somewhat unsatisfactory shelter expenditure (which includes rent paid but excludes both mortgage principal repayment and a return on owner's equity). The shelter variable was redefined in terms of rental payments and the independent variable C was replaced by C^*, defined as (C — housefurnishings — automobile expenditures — shelter + rental variable). The simple correlation matrix of the independent variables in the two cases is given in Table C.5 below ((A) contains 157 observations and (B) only 148, since the rental variable is not defined for cells with no renters).

Table C.6 shows the results of substituting C^* for C in regressions on the same set of dependent variables. Since C^* includes the rental item as defined in the text, the percentage of households in the cell which rent their homes was also added as an explanatory variable. In only two cases did that variable have a t value greater than 2.0, and those two are reported here:

Item	ln C^*	ln *Education*	*Age*	*Family Size*	*Region*	*Per Cent Renters*	$\bar{R}^2$
Medical care	0.829	−0.091	−0.004	−0.006	0.035	−0.496	.896
	(12.03)	(−1.10)	(−1.07)	(−1.18)	(1.08)	(−3.25)	
Utilities	0.409	0.011	0.002	0.009	−0.133	−0.555	.868
	(6.69)	(0.14)	(0.46)	(1.99)	(−4.70)	(−4.10)	

A comparison of Table C.6 with Table 1 again reveals no statistically significant coefficients changing signs. Most of the $\bar{R}^2$'s are higher in Table C.6 (notably tobacco), while the $\bar{R}^2$*'s* of the durable items are lower; only the housing regression seems appreciably affected by the use of this new consumption variable, undoubtedly due to the different definitions of the housing variable.

TABLE C.5
Simple Correlation Matrix

(A)
Total Consumption Expenditure

	ln *Education*	*Region*	ln *Consumption*	*Family Size*
Age	−0.730	0.035	−0.698	−0.633
ln Education		−0.234	0.588	0.228
Region			−0.227	0.074
ln Consumption				0.815

(B)
Total Current Consumption Expenditure

	ln *Education*	*Region*	ln C^*	*Family Size*	*Per Cent Renters*
Age	−0.733	0.035	−0.671	−0.638	0.153
ln Education		−0.235	0.596	0.232	−0.099
Region			−0.270	0.072	−0.023
ln C^*				0.790	−0.619
Family size					−0.667

The regressions were also run with the (ln C^*)·(ln E) and (ln C^*)·(A) interaction effects, both with and without the percentage of renters as an additional variable. Table C.7 indicates the implied elasticities for each item from the "best fit" of the four regression forms. Of the constant elasticity estimates in Table C.6, nine of the fourteen education coefficients are consistent with the neutrality

TABLE C.6
Total Current Consumption Expenditure[a]

Dependent Variable	ln C^*	ln *Education*	*Age*	*Family Size*[b]	*Region*	$\bar{R}^2$
Food (home)	0.6884	−0.2114	−0.0021	0.0145	−0.1120	.9430
	(14.05)	(−3.71)	(−0.88)	(4.51)	(−4.90)	
Food (away)	1.2468	0.1857	−0.0032	−0.0059	0.1026	.9287
	(15.05)	(1.93)	(−0.80)	(−1.09)	(2.66)	
Tobacco	0.8225	−0.8600	−0.0370	0.0002	0.0115	.8493
	(8.55)	(−7.70)	(−7.92)	(0.03)	(0.26)	
Alcohol	1.7623	−0.6880	−0.0320	−0.0223	−0.5150	.9069
	(13.42)	(−4.51)	(−5.01)	(−2.59)	(−8.41)	
Housing	0.8735	0.1522	0.0010	−0.0137	−0.1738	.9505
	(19.17)	(2.88)	(0.44)	(−4.60)	(−8.18)	
Household operations	1.1591	0.2739	0.0023	−0.0061	0.2108	.9822
	(31.99)	(6.51)	(1.29)	(−2.56)	(12.47)	
Housefurnishings and equipment	0.9820	−0.0696	−0.0121	0.0209	0.1337	.9317
	(11.44)[c]	(−0.70)	(−2.89)	(3.71)	(3.34)	
Clothing	1.2703	−0.0805	−0.0116	0.0098	0.1383	.9826
	(27.62)	(−1.51)	(−5.21)	(3.25)	(6.45)	
Personal care	0.9748	−0.1661	−0.0153	0.0018	0.1791	.9722
	(22.68)[c]	(−3.33)	(−7.29)	(0.65)	(8.93)	
Medical care	0.8674	−0.0042	0.0038	0.0027	0.0498	.8887
	(12.37)[c]	(−0.05)	(1.11)	(0.58)	(1.52)	
Leisure	1.3444	0.1023	−0.0135	0.0002	−0.0309	.9733
	(22.16)	(1.45)	(−4.57)	(0.06)	(−1.09)	
Education	1.5926	1.4814	0.0146	0.0574	0.4580	.8731
	(6.44)	(5.16)	(1.22)	(3.54)	(3.97)	
Travel	1.3680	−0.4428	−0.0361	0.0036	0.0843	.9451
	(14.62)	(−4.08)	(−7.92)	(0.59)	(1.93)	
Utilities	0.4519	0.1082	0.0109	0.0190	−0.1163	.8534
	(7.12)	(1.47)	(3.53)	(4.58)	(−3.93)	

[a] t values are in parentheses.
[b] Family size is in units different by a factor of 10 from previous results.
[c] Not statistically different from one (the t values for testing the difference from unity are −0.21 for housefurnishings, −0.59 for personal care, and − 1.89 for medical care.)

model; in Table C.7, ten are consistent. In comparing the best fit here with those in Table 4 (where comparisons are permissible), we find the major differences in those items which include the percentage of renters as an independent variable (notably medical care, personal care, housefurnishings, and housing), or in those items in which the linear form was the best fit in the text (i.e., tobacco). With all these differences taken into account, the use of C^* in place of C still does not appreciably change the overall expenditure patterns discussed in Chapter 4.

TABLE C.7
Expenditure Functions With Income Defined as Total Current Consumption Expenditures
Form With Highest $\bar{R}^2$

Item	*Regression[a] Form*	η^*	ϵ_E	ϵ_F	$\bar{R}^2$
Food (home)	4.	0.6908	−0.1613	+0.5975	.9437
Food (away)	4.	1.2265	0.1822	−0.2249	.9312
Tobacco	3.	0.6846	−0.7290	+0.1381	.8548
Alcohol	2.	1.8010	−0.6007	−0.4337	.9081
Housing	2.	0.8833	0.1743	−0.3695	.9507
Utilities	4.	0.3299	0.0620	+0.3084	.8700
Household operations	3.	1.1626	0.2546	−0.2442	.9833
Housefurnishings	2.	0.9565	−0.1272	+0.4851	.9326
Clothing	4.	1.1399	0.0084	+0.3566	.9855
Personal care	4.	0.7932	−0.0376	+0.1253	.9841
Medical care	4.	0.7896	−0.0884	−0.2474	.8963
Leisure	3.	1.3388	0.0897	−0.0321	.9731
Education	1.	1.5926	1.4814	+1.8440	.8731
Travel	4.	1.0933	−0.2918	+0.1028	.9570
Goods	4.	0.9376	−0.1096	+0.1285	.9941
Services	2.	1.1572	0.1275	−0.1381	.9887

[a] Regression forms were: (1) constant elasticity; (2) constant elasticity with the per cent renters as an additional explanatory variable; (3) interaction form discussed in the text; and (4) interaction form with the per cent renters as an additional explanatory variable.

6. THE ELASTICITY OF CONSUMPTION INCOME

Chapters 4, 5, and 6 present several estimates of the relationship between the income and education elasticities. In this section several additional, comparable estimates are shown for other combinations of the Engel curves, all taken from the data in Chapter 4. The purpose here is to present further evidence on the degree of sensitivity in the estimates presented in the text.

Each Engel curve estimates the income elasticity η_i and the education elasticity ϵ_{iE} for a particular market good. The regression discussed in this section take these pairs of elasticities, (η_i, ϵ_{iE}), as observations and regress ϵ_{iE} on η_i. The regression coefficient, b, obtained from fitting

$$\epsilon_{iE} = a + b\eta_i + u_i, \tag{C.1}$$

or

$$\epsilon_{iE} = b(\eta_i - 1) + u_i, \tag{C.2}$$

is an estimate of the elasticity of consumption income ϵ_{Y_cE} as defined in Chapter 2, since equation (2.11) states that

$$\epsilon_{iE} = \epsilon_{Y_cE}(\eta_i - 1).$$

Equations (C.1) and (C.2) differ only when the relevant means of the variables ϵ_{iE} and $(\eta_i - 1)$ and not zero.[4] Although for most of the sets of observations the means are not far from zero, all sets were run both weighted and unweighted for both equations (C.1) and (C.2).

Four additional sets of regressions across the Engel curves are discussed below. These will be denoted by the letters A, B, C, and D here to simplify the exposition. Set "A" includes fifteen constant elasticity Engel curves. The values of ϵ_{iE} and η_i are those given in Table 1 for the following items: food at home, food away, alcohol, housing, household operations, housefurnishings, clothing, personal care, leisure, education, and utilities. For the items tobacco and medical care, the estimates are slightly changed, the differences resulting from omitting the explanatory variables with t values of less than one; for the travel item, the expenditures were broken down into expenditures on automobiles and expenditures on other travel. The two elasticities, η_i and ϵ_{iE}, for these four items are, respectively, 0.758 and −0.819 for tobacco; 0.844 and 0.017 for medical care; 1.378 and −0.526 for automobiles; and 1.605 and −0.110 for travel other than by auto.

Set "B" for the same fifteen items replaces the constant elasticity form with the interaction form for six of the fifteen; the remaining eight are the same observations as in "A." The interaction effects for housing, automobiles, and other travel are given in the table of "best fits" (Table 4) in the text. For the remaining three—food at home, housefurnishings, and utilities—the interaction mean elasticities ($\bar{\eta}_i$ and $\bar{\epsilon}_{iE}$) are, respectively, 0.614 and −0.145, 0.961 and −0.045, and 0.430 and 0.130.

[4] Weighting by expenditures and summing over all goods, equation (C.1) becomes

$$\sum_i X_i \epsilon_{iE} = a \sum_i X_i + b \sum_i \eta_i X_i + \sum_i u_i X_i.$$

Dividing by the sum of the expenditures,

$$\bar{\epsilon}_E = a + b\bar{\eta} + \bar{u}$$

where the bar denotes a weighted mean. Since $\epsilon_E = 0$, $\bar{\eta} = 1$, and $\bar{u} = 0$, $-a = b$. So the weighted regression is also run as

$$\epsilon_{iE} = b(\eta_i - 1) + u_i.$$

Verbally, when weighted, the point of means of ϵ_{iE} and $(\eta_i - 1)$ is the origin.

TABLE C.8
Relationship Between Income and Education Elasticities Across Items[a]

	Means				*Simple Correlation* ($\epsilon_{iE}\eta_i$)		*Regression*[b] $\epsilon_{iE} = a + b\eta_i + u_i$				$\epsilon_{iE} = b(\eta_i - 1) + u_i$
	Weighted		*Unweighted*				*Weighted*		*Unweighted*		*Weighted*
Set	ϵ_{iE}	η_i	ϵ_{iE}	η_i	*Weighted*	*Unweighted*	a	b	a	b	b
A	−0.03	1.02	+0.02	1.11	−0.052	+0.210	0.070	−0.161	−0.326	0.309	−0.117
							(0.26)	(−0.60)	(−0.70)	(0.77)	(−0.44)
B	+0.01	0.97	+0.05	1.07	+0.142	+0.310	−0.107	0.086	−0.456	0.474	0.112
							(−0.44)	(0.33)	(−1.01)	(1.18)	(0.42)
C	+0.01	0.96	+0.05	1.07	+0.155	+0.326	−0.081	0.070	−0.478	0.492	0.083
							(−0.39)	(0.31)	(−1.08)	(1.25)	(0.42)
D	−0.02	0.92	+0.05	1.11	+0.522	+0.473	−0.571	0.629	−0.929	0.881	0.496
							(−3.59)	(3.10)	(−1.29)	(1.42)	(3.75)

[a] See pp. 127, 129 for definition of sets A–D.
[b] t values are in parentheses.

Set "C" is the same as the fifteen items given in Table 4, except for the substitution of the tobacco and medical care observations from set "A" described above. Set "D" is a subset of set "A," containing only the nine nondurable items: food at home, food away, tobacco, alcohol, household operations, personal care, medical care, leisure, and education. This nondurables set was considered most likely to be free of the durables bias discussed in Appendix B.

Table C.8 summarizes the regressions—weighted and unweighted—for these four sets of items. This evidence suggests again that the positive consumption income effect is considerably higher in the unweighted case, and much higher when estimated from only the nondurable items. Naturally, the results shown in the text most closely resemble those in set "C." If, in fact, the durables bias discussed in Appendix B is an important factor in these Engel curves, then the findings for set "D" may be the least biased, and this would suggest that the elasticity of real income with respect to education is somewhat larger than the value presented in Chapter 4. This finding for nondurables is qualitatively consistent with the conclusion reached at the end of Chapter 5; quantitatively the two sets of nondurables are not entirely comparable, and hence the magnitudes of the estimated coefficients differ.