6-1-1987

Effects of the Changing U.S. Age Distribution on Macroeconomic Equations

Ray C. Fair

Kathryn M. Dominguez

COWLES FOUNDATION FOR RESEARCH IN ECONOMICS

AT YALE UNIVERSITY

Box 2125, Yale University
New Haven, Connecticut 06520

COWLES FOUNDATION DISCUSSION PAPER NO. 839

EFFECTS OF THE CHANGING U.S. AGE DISTRIBUTION ON MACROECONOMIC EQUATIONS

by

Ray C. Fair and Kathryn M. Dominguez

June 1987

Effects of the Changing U.S. Age Distribution on Macroeconomic Equations

ABSTRACT

The effects of the changing U.S. age distribution on various macroeconomic equations are examined in this paper. The equations include consumption, money demand, housing investment, and labor force participation equations. Seven age groups are analyzed: 16-19, 20-24, 25-29, 30-39, 40-54, 55-64, and 65+. There seems to be enough variance in the age distribution data to allow reasonably precise estimates of the effects of a number of age categories on the macro variables. The results show that, other things being equal, age groups 30-39 and 40-54 consume less than average, invest less in housing than average, and demand more money than average. Age group 55-64 consumes more and demands more money. If these estimates are right, they imply, other things being equal, that consumption and housing investment will be negatively affected in the future as more and more baby boomers enter the 30-54 age group. The demand for money will be positively affected.

If, as Easterlin argues, the average wage that an age group faces is negatively affected by the percent of the population in that group, then the labor force participation rate of a group should depend on the relative size of the group. If the substitution effect dominates, people in a large group should work less than average, and if the income effect dominates, they should work more than average. The results indicate that the substitution effect dominates for women 25-54 and that the income effect dominates for men 25-54.

Ray C. Fair
Cowles Foundation
Box 2125, Yale Station
New Haven, CT 06520

Kathryn M. Dominguez
Kennedy School
Harvard University
Cambridge, MA 02138

EFFECTS OF THE CHANGING U.S. AGE DISTRIBUTION ON MACROECONOMIC EQUATIONS

by

Ray C. Fair and Kathryn M. Dominguez

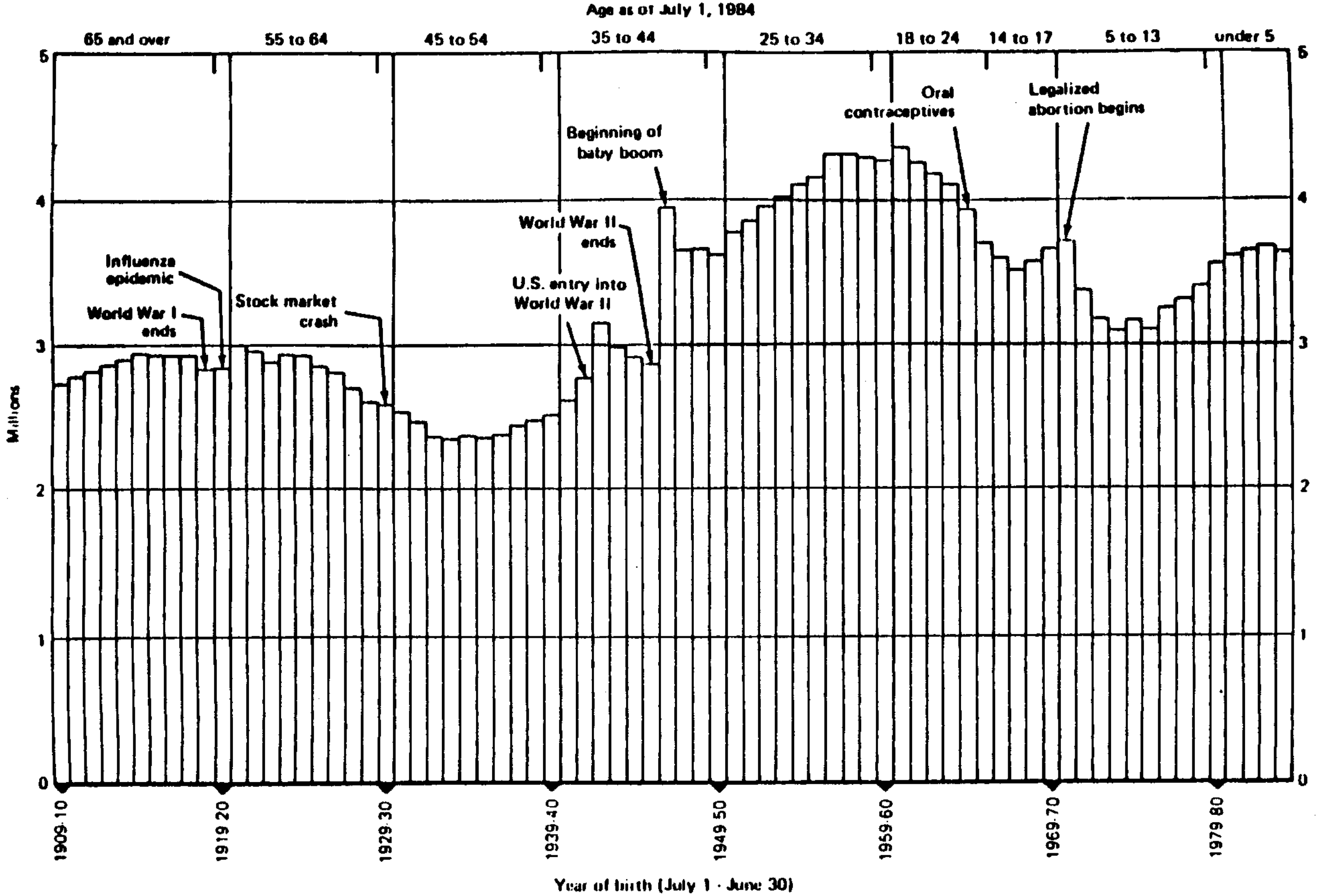
I. Introduction

A common assumption in empirical macroeconomics is that behavior can be summarized in terms of a representative agent. To the extent that this assumption is violated, aggregate equations are misspecified. If the income and age distributions of the population are fairly constant across time, the misspecification due to the representative agent assumption may be small and not of much concern. A striking feature of the post-war U.S. society, however, has been the baby boom of the late 1940's and the 1950's and the subsequent falling off of the birth rate in the 1960's. This can be seen in Figure 1, where the number of births by year is plotted for the period 1910-1984. The number of births rose from 2.8 million in 1945 to 4.3 million in 1961 and then fell back to 3.1 million in 1974. The consequences of this birth pattern for the percentage of middle age people in the working age population can be seen in Figure 2, where the ratio of the population 30-54 to the population 16+ is plotted for the years 1952-1986. This ratio fell from .47 in 1952 to .38 in 1976 as the baby boomers accounted for more and more of the population 16+. The ratio has risen sharply since 1981 as the baby boomers have begun to pass the age of 30. This rapidly changing age distribution clearly casts doubt on the reasonableness of the representative agent assumption.

In this paper we use U.S. Census Bureau age distribution data to examine the effects of the changing U.S. age distribution on several macroeconometric relationships, including consumption, money demand, housing investment, and labor force participation. Consumption equations are

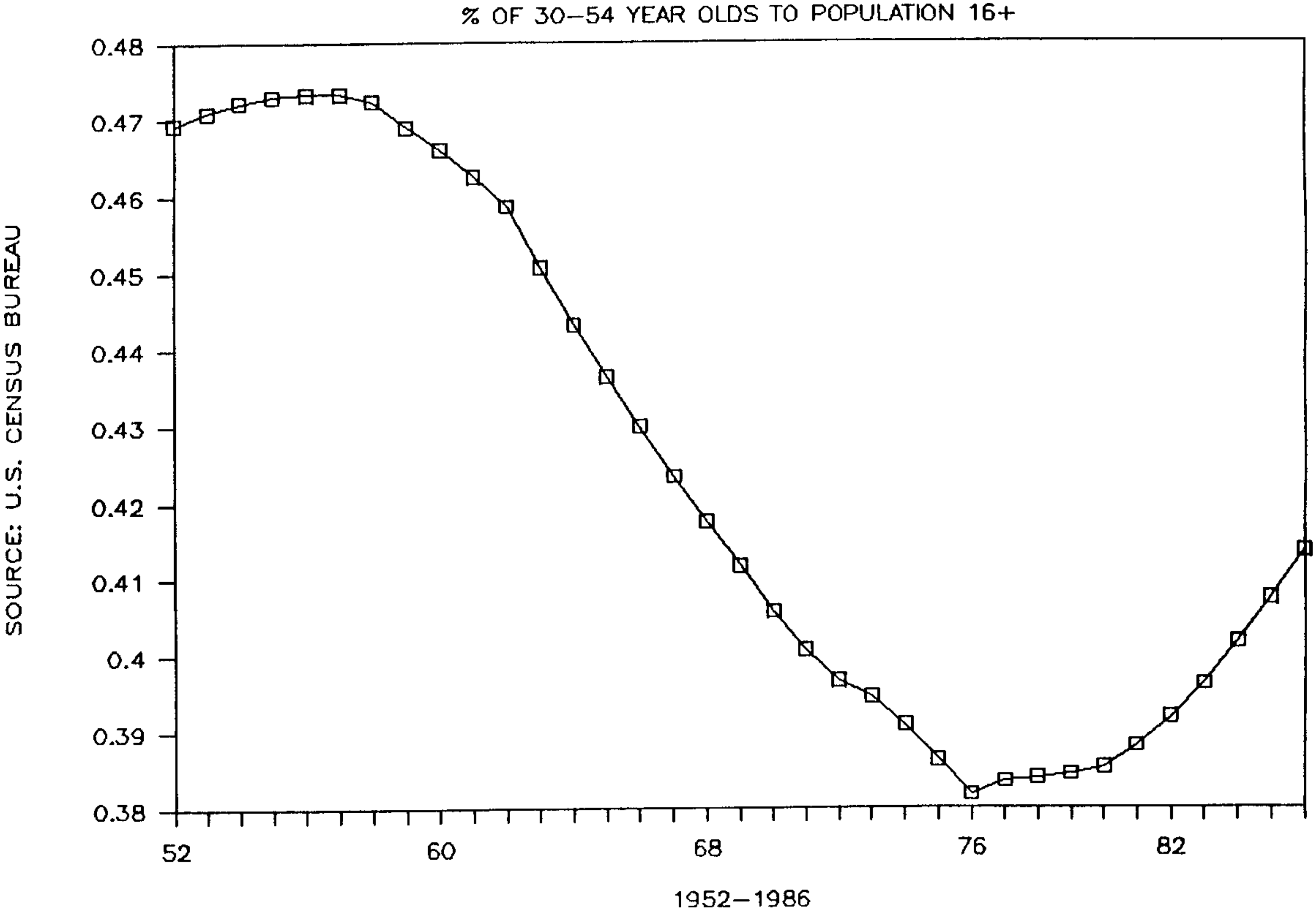
FIGURE 1

NUMBER OF BIRTHS, BY YEAR, 1910–1984



SOURCE: Current Population Reports, Series P-25, No. 965. U.S. Department of Commerce, Bureau of the Census, p.3.

FIGURE 2


% OF 30-54 YEAR OLDS TO POPULATION 16+
0.48
0.47
0.46
0.45
0.44
0.43
0.42
0.41
0.4
0.39
0.38
SOURCE: U.S. CENSUS BUREAU
52
60
68
76
82
1952-1986

considered first. The life cycle model of Ando and Modigliani (1963) predicts that people in their prime working years consume a smaller fraction of their income than do people younger and older. As will be seen, this can be tested using the age distribution data.[1]

Demand for money equations are considered next. In the Baumol (1952) and Tobin (1956) model of the demand for money, there is a positive relationship between the transactions costs associated with obtaining money and the optimal amount of money held by individuals. If the opportunity cost of bank visits is higher for prime-age people, which seems likely, then prime-age people will hold more money relative to their transactions than will people younger and older. This hypothesis can also be tested using the age distribution data.

Housing investment is considered next. If housing consumption is roughly proportional to the stock of housing, then the life cycle model implies that the housing stock (and thus housing investment) relative to

---

[1]As far as we know, no previous study has used the census population data to examine the effects of age on consumption. A number of studies of saving behavior have incorporated demographic factors using household survey data rather than census population data. Examples are Lieberman and Wachtel (1980) and Kane (1984). These studies have not provided reliable evidence on how changes in the age structure of the population affect personal saving behavior. In large part, the problem appears to be due to the limitations inherent in the survey data employed in the tests. Kane, for example, reports that "Survey respondents often provide erroneous information that is inconsistent ... these problems make it difficult to determine whether variability from survey to survey is due to actual changes in behavior or due to measurement difficulties." (p. 316)

Blinder (1975) and Stoker (1986) have tested for the effect of the income distribution on saving behavior and report no systematic relationship. They voice suspicion, however, that their results do not stem from the lack of a relationship, but rather from the relative stability of the U.S. income distribution since World War II. The advantage of the age distribution data used in this study is that there is considerable variation in the data over the sample period. There may thus be a good chance of picking up the effects of the changing age distribution on coefficients in macroeconometric equations if the effects are actually there.

income should vary with age. Prime-age people should consume less housing relative to their income than do older and younger people.

Finally, labor force participation equations are examined. Easterlin (1987) and Berger (1985) argue that larger cohorts on average face a lower wage rate because there are more people their own age to compete with. If this is the case, then the size of the cohort should affect the labor force participation of individuals in the cohort. They will work less if the substitution effect dominates, and they will work more if the income effect dominates.

II. Consumption and Money Demand

The Methodology

Divide the population into J age groups. Let $D1_{it}$ be 1 if individual i is in age group 1 in period t and 0 otherwise; let $D2_{it}$ be 1 if individual i is in age group 2 in period t and 0 otherwise; and so on through DJ_{it}. Consider the following equation:

$$(1)\quad C_{it} = Z_{it}\beta + \gamma + \alpha_1 D1_{it} + \ldots + \alpha_J DJ_{it} + U_{it}\ , \qquad \begin{array}{l} i = 1,\ldots,N_t\ , \\ t = 1,\ldots,T\ , \end{array}$$

where C_{it} is the dependent variable (say consumption or money demand of individual i in period t), Z_{it} is a 1 x k vector of explanatory variables not including the constant, β is a k x 1 vector of coefficients, and U_{it} is the error term. The constant term in the equation is $\gamma + \alpha_j$ for an individual in age group j in period t. N_t is the number of people in the population in period t.

Equation (1) is restrictive because it assumes that β is the same across all individuals, but it is less restrictive than a typical macroeconomic equation, which also assumes that the constant term is the same across individuals. Given Z_{it}, C_{it} is allowed to vary across age groups in equation (1). It would, of course, be useful to test whether some of the β coefficients vary across age groups. This is not in general possible, however, since most macroeconomic variables are not disaggregated by age groups, as would be necessary to test for age-sensitive β's. For example, suppose that one of the variables in Z_{it} is Y_{it}, the income of individual i, and let its coefficient be β_1. Assume that β_1 varies across age groups: $\beta_1 = \beta_{11}N_{1t} + \dots + \beta_{1J}N_{Jt}$, which introduces variables like $Y_{it}N_{1t}$ into the equation. The sum of this variable across i is the income of individuals in age group 1, for which data are not generally available. One is thus restricted to assuming that age-group differences are reflected in different constant terms in equation (1).

To aggregate equation (1) across individuals, let N_{jt} be the number of people in age group j in period t; let C_t equal the sum of C_{it}, where i runs from 1 through N_t; let Z_t be the 1 x k vector whose elements are the sums of the corresponding elements in Z_{it}, where i runs from 1 through N_t; and let U_t equal the sum of U_{it}, where i runs from 1 through N_t. Given this notation, summing equation (1) for a given t from 1 though N_t yields:

$$(2) \quad C_t = Z_t\beta + \gamma N_t + \alpha_1 N_{1t} \dots + \alpha_J N_{Jt} + U_t \quad , \quad t = 1,\dots,T \ .$$

It will be useful to rewrite equation (2) in per capita terms. Let $c_t = C_t/N_t$; let z_t be the vector Z_t with each element divided by N_t; let $u_t = U_t/N_t$; and let $p_{jt} = N_{jt}/N_t$, which is the proportion of people in age group j in the population at time t. Dividing equation (2) by N_t yields

$$(3) \quad c_t = z_t\beta + \gamma + \alpha_1 p_{1t} \dots + \alpha_J p_{Jt} + u_t \ , \quad t = 1,\dots,T \ .$$

A test of whether the age distribution matters is simply a test of whether the α_j coefficients in equation (3) are different from zero.[2] If the coefficients are zero, one is back to a standard macroeconomic equation. Otherwise, given z_t, c_t varies as the age distribution varies. Since the sum of p_{jt} across j is one and there is a constant in the equation, a restriction on the α_j coefficients must be imposed for estimation. The obvious restriction is that $\Sigma_{j=1}^{J}\alpha_j = 0$, and this was imposed in the estimation work reported in this paper.

The Data

The age distribution data are from the U.S. Bureau of the Census, Current Population Reports, Series P-25. The data from the census surveys, which are taken every ten years, are updated yearly using data provided by the National Center for Health Statistics, the Department of Defense, and the Immigration and Naturalization Service. The data are estimates of the total population of the United States, including armed forces overseas, in each of 86 age groups. Age group 1 consists of individuals less than 1 year old, age group 2 consists of individuals between 1 and 2 years of age, and so on through age group 86, which consists of individuals 85 years old and over. The published data are annual (July 1 of each year), and we have constructed quarterly data by linearly interpolating between the yearly

---

[2] Stoker (1986) characterizes this test, that all proportion coefficients are zero, as a test of microeconomic linearity or homogeneity (that all marginal reactions of individual agents are identical). He shows that individual differences or more general behavioral nonlinearities will coincide with the presence of distributional effects in macroeconomic equations.

points.[3]

We have considered seven age groups in this study: 16-19, 20-24, 25-29, 30-39, 40-54, 55-64, and 65+. The "total" population, N_t, is taken to be the population 16+. In terms of the above notation, we have created seven p_{jt} variables ($j = 1,\ldots,7$), where the seven variables sum to one for a given t.

The consumption and money demand equations are taken from the Fair (1984) model. For present purposes the equations are estimated for the period 1954 I - 1987 I.[4] The estimation technique is two stage least squares, with account sometimes taken of serial correlation of the error term. The first stage regressors for each equation are reported in Fair (1984).

Results for Consumption

The results for three categories of consumption -- service, nondurable, and durable -- are presented in Table 1. The possible set of explanatory variables for each consumption category consists of the real value of wealth (A), the after-tax nominal wage rate (W), the price level (P), the after-tax interest rate (R), after-tax nonlabor income (YN), a labor constraint variable (Q), and the lagged dependent variable. The theory behind the consumption equations is that households choose consumption and labor supply to maximize a multiperiod utility function. The variables that affect this

[3]The quarterly data for the 86 age groups from 1952 through 1986 are available on diskette from the authors upon request.

[4]The sample period for durable consumption ended in 1986 IV. The data for 1987 I were preliminary, and the observation for durable consumption for this quarter seemed extreme. Given the preliminary nature of the data, it seemed best to ignore this observation in the estimation work.

decision are the initial value of A and the current and expected future values of W, P, R, and YN. The current or lagged values of W, P, R, and YN and the lagged dependent variable are meant in part to be proxying for the unobserved expected future values.

The labor constraint variable Q is designed to pick up possible constraints on the household sector in how much it can work during a period at the current set of wage rates. Let C_t^* denote the amount that the household sector would choose to consume in period t if it could work as much as it likes at the current set of wage rates. Let C_t denote the observed amount of consumption. If households as a group are not constrained, then $C_t = C_t^*$; otherwise one would expect C_t to be less than C_t^*. The labor constraint variable is a function of labor market tightness. It is zero or close to zero when labor markets are tight, and its gets more and more negative as labor markets get looser and looser. The postulated equation for C_t is then

$$(4) \quad C_t = C_t^* + \eta Q_t \quad , \quad \eta > 0.$$

This specification thus introduces Q_t into the equation along with the variables that affect C_t^*.

Q is constructed as follows. First, the ratio of the total number of hours worked in the economy to the total population 16+ (JJ) is plotted for the 1952 I - 1987 I period. Another variable (JJP) is constructed from peak to peak interpolations of the JJ series. Q is then taken to be 1 - JJP/JJ. Q is thus a nonlinear function of the number of hours worked in the economy. When there is a lot of slack in the economy and JJ is considerably smaller than JJP, a one unit increase in JJ has more of an effect on Q than it does

TABLE 1. Estimates of the Consumption Equations

Explanatory Variable	CS/POP (1)	CS/POP (2)	CN/POP (3)	CN/POP (4)	CD/POP (5)	CD/POP (6)
Constant	.0139	.0770	.350	.460	.0660	.251
	(1.97)	(1.84)	(7.97)	(4.50)	(2.23)	(2.73)
LDV	.955	.926	.226	.233	.556	.510
	(49.75)	(31.57)	(2.60)	(2.34)	(7.68)	(5.89)
$(A/POP)_{-1}$	.00111	.00099	.00171	.00272	.00186	.00342
	(3.01)	(2.24)	(3.20)	(3.83)	(5.36)	(3.98)
W	.187[b]	.155[b]	.609[c]	.550[c]	.269[a]	.452[a]
	(2.45)	(1.36)	(7.75)	(7.13)	(2.48)	(2.90)
P	-.0959[b]	-.0374[b]			-.132	-.244
	(2.34)	(0.65)			(1.51)	(1.95)
R	-.00124	-.00182			-.00608	-.00772
	(4.48)	(4.98)			(5.33)	(5.79)
YN/POP			.0719	.0933	.155	.007
			(5.38)	(2.55)	(4.98)	(0.11)
Q			.401	.387	.208	.042
			(5.64)	(4.97)	(3.22)	(0.53)
P_1 (16-19)		.10		-.33		-.39
P_2 (20-24)		-.16		-2.08		-.79
		(0.42)		(2.65)		(1.02)
P_3 (25-29)		-.24		.34		.16
		(1.33)		(0.84)		(0.29)
P_4 (30-39)		-.19		-.85		-.46
		(0.98)		(1.57)		(1.06)
P_5 (40-54)		-.28		-.99		-1.33
		(1.21)		(2.78)		(2.38)
P_6 (55-64)		1.15		4.28		3.81
		(1.35)		(2.90)		(1.64)
P_7 (65+)		-.38		-.37		-1.00
		(0.72)		(0.41)		(0.79)
$\hat{\rho}$	-.178	-.216	.587	.492		
	(2.10)	(2.51)	(6.61)	(4.29)		
SE	.00513	.00496	.00573	.00546	.00905	.00854
DW	2.08	2.12	2.07	2.06	1.83	1.92
χ^2		12.97		13.24		25.72

TABLE 1 (cont.)

Notes:

The estimation technique is two stage least squares.
Sample period = 1954 I - 1987 I for CS and CN equations.
= 1954 I - 1986 IV for CD equation.
t-statistics in absolute value are in parentheses.
$\hat{\rho}$ = estimate of first order serial correlation coefficient of the error term.

The χ^2 test is a test of the hypothesis that the coefficients of the age variables are zero. The critical values (6 degrees of freedom) are 12.59 at the 5 percent level and 16.81 at the 1 percent level.

a = variable is lagged one quarter.
b = variable is lagged two quarters.
c = variable is W/P.

Notation:

A = Real value of total net worth of the household sector.
CD = Real value of the consumption of durables.
CN = Real value of the consumption of nondurables.
CS = Real value of the consumption of services.
LDV = Lagged dependent variable.
P_j = Percentage of the population 16+ in age group j.
P = Price deflator. Price deflator for CS for the CS equation; price deflator for CN for the CN equation; price deflator for CD for the CD equation.
POP = Population 16+.
Q = Labor constraint variable.
R = After-tax interest rate. Short term interest rate for the CS equation. Long term interest rate for the CD equation.
W = After-tax nominal wage rate.
YN = After-tax nonlabor income. Total nonlabor income for the CN equation. Transfer payment income for the CD equation.

when JJ is close to JJP. In other words, Q is more affected by the number of hours worked in slack times than in tight times.

Note that labor income is not an explanatory variable in the consumption equations. If households jointly determine consumption and labor supply and are not constrained in their labor supply choice, labor income is simply a byproduct of this decision, given the wage rate. It is not appropriate in this case to include labor income, which is the labor supply times the wage rate, as an explanatory variable in the consumption equations. If, on the other hand, households are constrained in their labor supply, then labor supply is no longer a decision variable, and it is appropriate to consider labor income as a determinant of consumption. In the present specification, Q and W are jointly highly correlated with labor income in slack times, but less so in tight times. The specification is thus like the Keynesian story in low employment periods, but it differs from this story more and more as the economy comes closer and closer to full employment.

As a final point about the equations, not all the potential explanatory variables had coefficient estimates that were significant and/or of the right sign in all the equations. If after some experimentation with lagged values a variable always had a coefficient estimate of the wrong sign, it was dropped from the equation. The equations reported in Table 1 are the "final" estimated equations.

Consider first the results for the consumption of services in Table 1. Column (1) presents the results without the age variables, and column (2) presents the results with the age variables.[5] (The age variables were always taken to be exogenous in the estimation work.) The age variables as a group are significant at the 5 percent level according to the chi-square test.[6] The key question regarding the signs of the age variables is whether the coefficient estimates are negative for people 30-39 and 40-54, i.e., whether, other things being equal, people in their prime working years consume less than do those younger and older. This is the case in Table 1. Also, there is a large positive coefficient for people 55-64, which is as expected.

The results for the consumption of nondurables in columns (3) and (4) also show that the age variables are significant at the 5 percent level and that the estimates are negative for people 30-39 and 40-54 and positive for people 55-64. The results are even stronger for the consumption of durables in columns (5) and (6), where the age variables are significant at the 1

---

[5]The population variable POP in Table 1 that is used to put the equations in per capita terms differs slightly from the population variable N that is used to construct the p_j variables. POP is constructed from monthly data from the Bureau of Labor Statistics, and N is constructed from yearly data from the Current Population Reports. N is used in the construction of the p_j variables so that they will sum to one across j.

[6]The chi-square test is a follows. The 2SLS objective function is $u'Z(Z'Z)^{-1}Z'u = S$, where u is a Tx1 vector of error terms and Z is a TxK vector of first stage regressors. u is a function of the coefficients and the endogenous and predetermined variables in the equation. When the equation is estimated under the assumption of first order serial correlation of the error terms, u is a nonlinear function of the coefficients if the serial correlation coefficient is counted as a structural coefficient, which is the treatment here. Now, assume that there are r restrictions on the coefficients. (In the present case there are 6 zero restrictions.) Let S^* be the value of S when the restrictions are not imposed, and let S^{**} be the value of S when the restrictions are imposed. Let $\hat{\sigma}^2$ be the estimate of the variance of the error term in the unrestricted case. Then $(S^{**} - S^*)/\hat{\sigma}^2$ is asymptotically distributed as chi-square with r degrees of freedom. A general proof of this is in Andrews and Fair (1987).

percent level. The sign pattern of the age variables is the same for all three categories of consumption except for the 16-19 and 25-29 groups.

The overall results for consumption are thus quite strong. Even though seven age categories are used (six unrestricted coefficient estimates), the expected signs for the main age categories are obtained. The age variables as a group are always significant at the 5 percent level, and even some of the individual estimates are significant.

As a final test of the effects of the age variables on consumption, the demand for imports equation in the Fair model was estimated with the age variables added. Since imports include more than consumption items, the import results are probably less reliable than the other consumption results. The results are presented in Table 2. The age variables are significant at the 1 percent level, and the signs for the 30-39 and 40-54 groups are negative as expected. The addition of the age variables did, however, make the domestic price variable highly insignificant (and of the wrong sign), which indicates some collinearity problems.

Results for Money Demand

A typical demand for money model begins by postulating that the long-run desired level of real money balances (M_t^*/P_t) is a function of real income (Y_t) and a short-term interest rate (R_t):

$$(5) \quad M_t^*/P_t = \alpha + \beta Y_t + \gamma R_t .$$

An adjustment equation is then postulated, where the adjustment may either be in real terms (M_t/P_t adjusting to M_t^*/P_t) or in nominal terms (M_t adjustment to M_t^*). The results in Fair (1987) strongly support the nominal adjustment hypothesis, and so this will be used here. The hypothesis is

$$(6) \quad M_t - M_{t-1} = \lambda(M_t^* - M_{t-1}) + \mu_t \quad .$$

Combining (5) and (6) yields:

$$(7) \quad M_t/P_t = \lambda\alpha + \lambda\beta Y_t + \lambda\gamma R_t + (1-\lambda)(M_{t-1}/P_t) + \mu_t/P_t \quad .$$

The results of estimating the demand for money equations are presented in Table 3. The first set of results concerns the demand for money (demand deposits and currency) of the household sector. The equation estimated in column (1) is the same as equation (7) except that it is in per capita terms.[7] The equation estimated in column (2) has the age variables added. The age variables are significant at the 1 percent level, and some of the individual coefficient estimates are quite significant. The coefficient estimates are positive for the 30-39 and 40-54 groups and negative for the 55-64 and 65+, which is as expected. The estimates indicate that people in their prime working years demand more money relative to their transactions than otherwise because the opportunity cost of their time is higher. The other noticeable result is that the addition of the age variables has considerably lessened the size of the coefficient of the lagged dependent variable (the estimate of $1-\lambda$).

The second set of results in Table 3 concerns the demand for currency. Sectors other than the household sector demand currency, and so the results

---

[7]When there are lagged values in per capita equations, the issue always arises as to whether the lagged values should be divided by the current population value or the relevant lagged population value. In practice this turns out to make little difference. In the present case the relevant lagged population values have been used. Note also that the error term in the per capita version of equation (7) is $\mu_t/(P_t POP_t)$. For purposes of the estimation work this error term has been assumed to be homoscedastic.

TABLE 2. Estimates of the Import Equation

Explanatory Variable	IM/POP (1)	IM/POP (2)
Constant	-.129	-.030
	(3.91)	(0.33)
$(IM/POP)_{-1}$	.676	.561
	(9.22)	(6.44)
X/POP	.0603	.0990
	(4.05)	(3.28)
PM	-.0916	-.185
	(2.89)	(3.30)
PX	.215	-.0114
	(4.62)	(0.09)
R	-.00435	-.00194
	(2.58)	(0.85)
p_1 (16-19)		-2.63
p_2 (20-24)		.07
		(0.06)
p_3 (25-29)		.89
		(1.52)
p_4 (30-39)		-.43
		(0.63)
p_5 (40-54)		-1.41
		(2.64)
p_6 (55-64)		.16
		(0.06)
p_7 (65+)		3.35
		(2.09)
$\hat{\rho}$	.236	.137
	(1.84)	(1.08)
SE	.0124	.0115
DW	2.01	2.00
χ^2		26.41

TABLE 2 (cont.)

Notes:

The estimation technique is two stage least squares.
Sample period = 1954 I - 1987.
t-statistics in absolute value are in parentheses.
$\hat{\rho}$ = estimate of first order serial correlation coefficient of the error term.

The χ^2 test is a test of the hypothesis that the coefficients of the age variables are zero. The critical values (6 degrees of freedom) are 12.59 at the 5 percent level and 16.81 at the 1 percent level.

Notation:

IM = Real value of imports.
p_j = Percentage of the population 16+ in age group j.
PM = Price deflator for IM.
POP = Population 16+.
PX = Price deflator for domestic output.
R = After-tax long-term interest rate.
X = Real value of domestic sales.

TABLE 3

Estimates of the Money Demand Equations

Explanatory Variable	MH/(POP*P) (1)	(2)	CUR/(POP*P) (3)	(4)
Constant	-.0722	-.820	.0116	.00291
	(1.08)	(4.31)	(0.84)	(0.09)
LDV	.986	.802	.944	.873
	(39.14)	(14.15)	(59.17)	(15.37)
YD/POP	.0601	.309	.0144	.0296
	(5.42)	(4.57)	(8.41)	(2.19)
R	-.00630	-.00807	-.00113	-.00205
	(3.13)	(3.45)	(2.75)	(3.11)
P_1 (16-19)		3.27		.09
P_2 (20-24)		5.15		.60
		(2.57)		(1.11)
P_3 (25-29)		-1.37		.03
		(1.13)		(0.12)
P_4 (30-39)		5.50		.25
		(4.05)		(0.83)
P_5 (40-54)		4.18		.27
		(2.82)		(0.77)
P_6 (55-64)		-8.43		-.28
		(1.82)		(0.28)
P_7 (65+)		-8.31		-.86
		(4.03)		(1.60)
$\hat{\rho}$			-.300	-.314
			(3.51)	(3.41)
SE	.0301	.0277	.00781	.00738
DW	2.21	2.07	2.02	2.07
χ^2		23.68		14.83

TABLE 3 (cont.)

Notes:

The estimation technique is two stage least squares.
Sample period = 1954 I - 1987 I.
t-statistics in absolute value are in parentheses.
$\hat{\rho}$ = estimate of first order serial correlation coefficient of the error term.

The χ^2 test is a test of the hypothesis that the coefficients of the age variables are zero. The critical values (6 degrees of freedom) are 12.59 at the 5 percent level and 16.81 at the 1 percent level.

Notation:

CUR = Currency held outside banks, current dollars.
LDV = Lagged dependent variable. $MH_{-1}/(POP_{-1}P)$ for MH equation.

$CUR_{-1}/(POP_{-1}P)$ for CUR equation.

MH = Demand deposits and currency of the household sector, current dollars.
p_j = Percentage of the population 16+ in age group j.

P = Price deflator for domestic sales.
POP = Population 16+.
R = After-tax short-term interest rate.
YD = Real value of disposable income for the MH equation. Real value of domestic sales for the CUR equation.

for the currency equation are not expected to be as good. The results show that the age variables as a group are significant at the 5 percent level, although none of the individual coefficient estimates of the age variables is significant. As expected, the sign for the groups 30-39 and 40-54 is positive and the sign for the groups 55-64 and 65+ is negative.

III. Housing Investment

The following model in Fair (1984) lies behind the housing investment equation. Let KH^{**} denote the desired stock of housing. If housing consumption is proportional to the housing stock, then the factors discussed in Section II that affect consumption, including the age variables, also affect KH^{**}:

$$(8) \quad KH^{**} = f(\ldots) \ ,$$

where the arguments in f are A, W, P, R, YN, and the age variables. Two types of lagged adjustment are then postulated. The first is an adjustment of the housing stock to its desired value:

$$(9) \quad KH^{*} - KH_{-1} = \lambda(KH^{**} - KH_{-1}).$$

The physical depreciation of the housing stock is assumed to be proportional to the size of the stock, with depreciation rate δ. Gross investment in housing (IH) is thus equal to $KH - (1-\delta)KH_{-1}$. Given KH^{*} from (9), desired gross investment is thus

$$(10) \quad IH^{*} = KH^{*} - (1-\delta)KH_{-1} \ .$$

The second type of adjustment is an adjustment of gross investment to its desired value:

$$(11)\quad IH - IH_{-1} = \gamma(IH^{*} - IH_{-1}).$$

Combining (8)-(11) yields

$$(12)\quad IH = (1-\gamma)IH_{-1} + \gamma(\delta-\lambda)KH_{-1} + \gamma\lambda f(\ldots)\ .$$

This treatment thus adds the lagged dependent variable and the lagged stock of housing to the housing investment equation, both of which seem to be important explanatory variables in practice.

According to (12), the age variables affect housing investment to the extent that they are arguments in f, i.e., to the extent that they affect housing consumption. One problem with this is that the age variables may affect the adjustment parameters λ and γ. In particular, adjustment may be faster for the young than for the old.[8] Unfortunately, as discussed above, sufficient data are not available to allow parameters other than the constant term to be a function of the age variables, and so it must be assumed that λ and γ do not vary by age. The housing results are thus likely to be less reliable than the consumption results in Table 1.

The results of estimating equation (12) are presented in Table 4. Adding the age variables to the equation makes a considerable difference. The original equation was estimated under the assumption of third order serial correlation of the error term, and adding the age variables made all

---

[8]Hanushek and Quigley (1979) hypothesize that households' consumption of housing in any given period will deviate significantly from their desired level due to the substantial transactions and search costs associated with the housing market. They find (based upon reinterview data gathered on low-income renter households) that young households initially consuming "too little" housing close the gap between actual and desired consumption more rapidly than do older households who are consuming "too much" housing.

TABLE 4. Estimates of the Housing Investment Equation

Explanatory Variable	IH/POP (1)	IH/POP (2)
Constant	.0858	.516
	(2.55)	(5.64)
$(IH/POP)_{-1}$	.728	.863
	(10.20)	(30.48)
$(KH/POP)_{-1}$	-.00916	-.0371
	(2.50)	(7.92)
$(A/POP)_{-1}$	.00244	.00275
	(3.38)	(4.59)
W_{-2}	.0379	.0086
	(1.27)	(0.13)
$(YN/POP)_{-1}$	.163	.161
	(3.86)	(4.10)
R_{-1}	-.00577	-.00511
	(3.76)	(6.48)
P_1 (16-19)		-.87
P_2 (20-24)		1.84
		(3.56)
P_3 (25-29)		-1.38
		(4.87)
P_4 (30-39)		-.41
		(1.38)
P_5 (40-54)		-1.24
		(4.27)
P_6 (55-64)		-1.91
		(1.33)
P_7 (65+)		3.97
		(3.59)
$\hat{\rho}_1$	.342	-.080
	(3.03)	(0.86)
$\hat{\rho}_2$	.219	-.066
	(2.22)	(0.73)
$\hat{\rho}_3$	.161	.027
	(1.77)	(0.30)
SE	.00728	.00607
DW	1.94	2.01
χ^2		57.74

TABLE 4 (cont.)

Notes:

The estimation technique is ordinary least squares.
Sample period = 1954 I - 1987 I.
t-statistics in absolute value are in parentheses.
$\hat{\rho}_i$ = estimate of the i^{th} order serial correlation coefficient of the error term.

The χ^2 test is a test of the hypothesis that the coefficients of the age variables are zero. The critical values (6 degrees of freedom) are 12.59 at the 5 percent level and 16.81 at the 1 percent level.

Notation:

A = Real value of total net worth of the household sector.
IH = Real value of housing investment of the household sector.
KH = Real value of the stock of housing of the household sector.
p_j = Percentage of the population 16+ in age group j.
POP = Population 16+.
R = After-tax long-term interest rate.
W = After-tax nominal wage rate.
YN = After-tax nonlabor income.

three estimates of the serial correlation coefficients small and insignificant. The age variables as a group are highly significant, and a number of the individual estimates are significant. The coefficient estimate of the lagged dependent variable went up considerably, as did the absolute value of the coefficient estimate of the housing stock variable. The signs for the age groups 30-34 and 40-54 are negative as expected: prime-age households consume less housing, other things being equal, relative to the old and the young. The signs for the age groups 25-29 and 55-64 are also negative, however, which is contrary to the consumption results in Table 1. The signs for age groups 20-24 and 65+ are positive, as expected, and the size of the coefficient on the latter age group is quite large.[9]

The results for the housing investment equation thus seem quite good. The fit of the equation is considerably improved by the addition of the age variables, and serial correlation of the error terms is eliminated. However, the negative coefficient estimates for age groups 25-29 and 54-65 are not necessarily as expected, which may mean that some of the estimates are spurious or that something else is going on aside from life cycle considerations.

IV. Labor Force Participation Equations

In most macroeconometric models, including the Fair model, labor force

---

[9] Russell (1982) points out that older people have been major contributors to housing trends because of the rising rate of household (as opposed to individual) incidence in that age group. The decision not to disband existing households when a spouse retires or dies means that older people continue to occupy the existing housing stock, which requires younger households to buy new rather than existing houses.

participation equations are disaggregated by age and sex. A typical equation is of the form

$$(13)\quad L^{s}_{jt} = Z_t\beta + u_t \quad , \quad t = 1,\ldots,T \ , \ s = m, f.$$

where L^{s}_{jt} is the ratio of the number of individuals of sex s in age group j in the labor force to the total population of people of sex s in age group j. The variables in Z_t include variables like the real wage and variables that are designed to pick up possible "discouraged worker effects."

Although the left-hand-side variables in equations like (13) are disaggregated by age and sex, the right-hand-side variables are typically not age-sex specific. The aggregate real wage is used in place of the more appropriate but unobserved real wage of the particular age-sex group. The implicit assumption in this treatment is that the real wage relevant to age group j (say W_{jt}) is proportional to the aggregate wage (W_t): $W_{jt} = \lambda_j W_t$. The Easterlin hypothesis suggests that λ_j varies across time and is a negative function of the percent of people in age group j in the total population. For baby boomers, Berger (1985) finds that λ_j is low relative to the λ_j for other age cohorts because there are so many baby boomers competing with each other.

A way to test the Easterlin hypothesis is to postulate that the ratio of W_{jt} to W_t is a function of p_{jt}:

$$(14)\quad W_{jt}/W_t = \gamma_0 + \gamma_1 p_{jt} \ ,$$

where γ_1 is negative. Assume that in equation (13) W_{jt} is the appropriate explanatory wage variable, so that $\beta_1 W_{jt}$ is one of the terms in the equation. Substituting (14) into (13) then results in the terms $\beta_1\gamma_0 W_t$ and

$\beta_1\gamma_1 p_{jt}W_t$ in the equation. Since γ_1 is negative, one expects the coefficients of W_t and $p_{jt}W_t$ to be of opposite signs. If the substitution effect dominates, β_1 is positive, and so one expects the coefficient of W_t to be positive and the coefficient of $p_{jt}W_t$ to be negative. The opposite is true if the income effect dominates.

The Easterlin hypothesis actually consists of two parts. The first is that there is not perfect substitution across age groups in the labor market and so the more people there are in the age group the smaller is the average wage for that group. In the present case this is represented by equation (14). The second part, termed the relative income hypothesis, says that young peoples' consumption aspirations are shaped by their parents' living standards. In the face of unfavorable labor market conditions a large cohort will adjust demographic and economic behavior in order to maintain its consumption aspirations. The baby boom generation, Easterlin argues, delayed marriage and children and increased labor participation of young women[10] in response to lower average wages.[11] In other words, Easterlin is postulating that the income effect dominates for women. Therefore, the

---

[10]Easterlin notes that since most young men (in the family forming ages) are already committed to the labor force, increased labor force participation will come primarily from young women, but possibly also via moonlighting by the men.

[11]An alternative sociological explanation for the increased labor participation (and drop in fertility) of women, as discussed for example in Perry (1977), is based not on the economic incentives brought about by the decline in the relative earnings of the baby boom generation, but on the changing attitudes about the role of women in society brought about by the "women's movement."

Johnson and Skinner (1986) find support for the hypothesis that future divorce probabilities increase current labor supply for married women. They conclude that the rise in the frequency of divorce since 1960 may account for one-third of the unexplained increase in women's post war labor force participation.

coefficient of $p_{jt}W_t$ in the equation explaining the participation of women in age group j should be positive: women baby boomers should work more.

For sake of the present discussion, "prime age" will be defined to be ages 25-54. The results of estimating labor force participation equations for prime age men and prime age women are presented in Table 5. The first set of results is for prime age men. Column (1) contains the results for the equation without the addition of the $p_{jt}W_t$ term. Labor force participation is a function of the after-tax nominal wage (W), the price level (P), the labor constraint variable (Q), and the lagged dependent variable. The labor constraint variable is meant to pick up discouraged worker effects (discouraged in the disequilibrium sense of not being able to find a job at the current set of wage rates, not discouraged in the sense that the current set of wage rates is low). The lagged dependent variable is meant to pick up expectational and partial adjustment effects. The coefficient estimate for W is negative and the coefficient estimate for P is positive, which means that the real wage has a negative effect on participation. The estimates thus indicate that the income effect dominates. When the variable $p_{jt}W_t$ is added in column (2), its coefficient estimate is positive and significant, which is as expected when the income effect dominates. In column (3) the time trend is added to see if the wage and price variables may be erroneously picking up general trend effects. The results are little affected by the addition of the time trend.

The second set of results in Table 5 is for prime age women. The results without the $p_{jt}W_t$ variable are in column (4). The nominal wage rate has a positive coefficient estimate and the price level has a negative coefficient estimate, which means that the real wage has a positive effect

TABLE 5

Estimates of the Labor Force Participation Equations

Explanatory Variable	L1/POP1 (1)	L1/POP1 (2)	L1/POP1 (3)	L2/POP1 (4)	L2/POP1 (5)	L2/POP1 (6)
Constant	.266	.403	.429	.0357	.0662	.0832
	(4.32)	(5.72)	(5.82)	(3.15)	(4.79)	(4.79)
LDV	.726	.586	.558	.914	.826	.771
	(11.45)	(8.09)	(7.35)	(30.02)	(21.64)	(15.07)
W	-.079	-.149	-.189	.113[a]	.250[a]	.237[a]
	(4.14)	(5.52)	(4.36)	(3.65)	(5.14)	(4.85)
P	.0429	.0356	.0477	-.0469[a]	-.0259[a]	-.0066[a]
	(3.98)	(3.39)	(3.25)	(3.63)	(1.90)	(0.37)
Q	.0085	.0122	.0133	.0241	.0353	.0413
	(1.11)	(1.64)	(1.79)	(2.13)	(3.15)	(3.53)
$W*(p_3+p_4+p_5)$		.143	.170		-.250	-.262
		(3.51)	(3.64)		(3.53)	(3.73)
t			.0000374			.0000943
			(1.17)			(1.58)
SE	.00191	.00182	.00182	.00294	.00279	.00276
DW	2.15	2.05	2.01	2.19	2.20	2.13

Notes:

The estimation technique is two stage least squares.
Sample period - 1954 I - 1987 I.
t-statistics in absolute value are in parentheses.

a - variable is lagged one quarter.

Notation:

L1 - Total labor force of men 25-54.
L2 - Total labor force of women 25-54.
LDV - Lagged dependent variable.
$(p_3+p_4+p_5)$ - percentage of the population 16+ between the ages of 25 and 54.
P - Price deflator.
POP1 - Population of men 25-54.
POP2 - Population of women 25-54.
Q - Labor constraint variable.
t - time trend: 1 in 1952 I, 2 in 1952 II, etc.
W - After-tax nominal wage rate.

on participation. Contrary to the case for men, the estimates indicate that the substitution effect dominates for women. When the variable $p_{jt}W_t$ is added in column (5), its coefficient estimate is negative and significant, which is as expected when the substitution effect dominates. In column (6) the time trend is added, but as was the case for men, the results are little affected by the addition of the time trend.

To summarize, the results in Table 5 show that the age variable when multiplied by the aggregate wage is significant. The coefficient estimates of W_t and $p_{jt}W_t$ are also of opposite signs, as expected from equation (13). The results thus provide support for the part of the Easterlin hypothesis embodied in equation (13). They do not, however, support Easterlin's hypothesis that the income effect dominates for women. The income effect dominates for men, but the substitution effect dominates for women.[12]

It is fairly clear from examining the data why the income effect dominates for men. The after-tax real wage generally grew from the beginning of the data set (1952) to about 1974, after which it flattened out. The participation rate of prime age men fell slightly from 1952 to 1967, fell at a faster rate from 1967 to about 1976, and then flattened out after that. The estimates thus attribute the fall in the participation rate to the rise in the real wage and the flattening out of the participation

---

[12] Wachter (1977) regresses labor force participation by fourteen different age-sex groups on the proportion of the population aged 16-34 in the population 16+, the unemployment rate, a time trend, and lagged labor force participation. He finds that the young-worker variable (16-34) has significantly negative coefficient estimates for men 25-64 and women 45-65+ over the period 1949-76. However, because the regressions do not include the wage rate, it is not possible to interpret the results in terms of income and substitution effects. Further, it is somewhat unclear as to the expected effects of one age group's relative size on other age groups' labor force participation, which makes the regressions difficult to interpret.

rate to the flattening out of the real wage. This thus seems to be the income effect at work. The participation rate of prime age women, on the other hand, has risen fairly steadily over the entire 1952-1987 period, and the estimates are attributing at least some of this rise to the rise in the real wage before 1974. This thus seems to be the substitution effect at work.

V. Conclusion

The results in this paper are to some extent rather striking. The changing age distribution of the U.S. population seems to have a highly significant effect on consumption (including imports), money demand (including currency demand), housing investment, and labor force participation. There seems to be enough variance in the age distribution data to allow reasonably precise estimates of the effects of a number of age categories on the macro variables. The results show that, other things being equal, age groups 30-39 and 40-54 consume less (including less housing) than average and demand more money (including currency) than average. Age group 55-64 consumes more (although not more housing) demands less money. The results also show that the labor force participation rate of both men and women 25-54 is affected by the percent of people of this age in the population 16+. For men the income effect dominates and the effect is positive, and for women the substitution effect dominates and the effect is negative.

Since the use of the age distribution data in the manner done in this study has not been tried before, the results must be interpreted with some caution. Adding six new explanatory variables to any macroeconometric equation is risky, since collinearity problems may lead to very imprecise and possibly ridiculous results. The results may also be sensitive to the specification of the equations in the Fair model, and results using other macroeconomic equations would be of interest in future work.

If it turns out that the estimates obtained here are roughly right, they have important consequences for the future course of the economy. They say that consumption and housing investment will be negatively affected in the future, other things being equal, as more and more baby boomers enter the 30-54 age group. According to the estimates, people in this age group consume less than average and invest in less housing than average. This movement, on the other hand, will have a positive effect on the demand for money, other things being equal, since people in this age group demand more money than average.

Finally, note that if the estimated age effects in this paper are not spurious, they could lead to a considerable increase in forecasting accuracy. Beginning, say, at age 16, the age proportion variables can be very accurately forecast for many years ahead. The age proportion variables are thus easily forecasted exogenous variables that seem to have considerable explanatory power, a forecaster's dream.

REFERENCES

Ando, Albert, and Modigliani, Franco, "The 'Life Cycle Hypothesis' of Saving", American Economic Review, 51(1): 55-84, March 1963.

Andrews, Donald T.M., and Ray C. Fair, "Inference in Econometric Models with Structural Change", Cowles Foundation Discussion Paper No. 832, April 1987.

Baumol W. J., "The Transactions Demand for Cash: An Inventory Theoretic Approach", Quarterly Journal of Economics, 66: 545-56, November 1952.

Berger, Mark C., "The Effect of Cohort Size on Earnings Growth: A Reexamination of the Evidence", Journal of Political Economy, 93(3): 561-573, June 1985.

Blinder, Alan S., "Distribution Effects and the Aggregate Consumption Function", Journal of Political Economy, 83: 447-75, June 1975.

Easterlin, Richard A., Birth and Fortune: The Impact of Numbers on Personal Welfare, 2nd edition, Chicago: University of Chicago Press, forthcoming 1987.

Fair, Ray C., Specification, Estimation, and Analysis of Macroeconometric Models, Harvard University Press, 1984.

Fair, Ray C., "International Evidence on the Demand for Money", Review of Economics and Statistics, forthcoming 1987.

Hanushek, Eric A. and Quigley, John M., "The Dynamics of the Housing Market: A Stock Adjustment Model of Housing Consumption", Journal of Urban Economics, 6: 90-111, 1979.

Johnson, William R. and Skinner, Jonathan, "Labor Supply and Martial Separation", American Economic Review, 76: 455-469, June 1986.

Lieberman, Charles, and Wachtel, Paul, "Age Structure and Personal Saving Behavior", in Social Security Versus Private Saving, edited by George von Furstenberg, pp.315-357. Cambridge, Mass.:Ballinger Publishing Company, 1980.

Kane, Edward J., "Microeconomic Evidence on the Composition of Household Savings in Recent Years", in Level and Composition of Household Savings, edited by Patric H. Hendershott, pp.151-180. Cambridge, Mass.:Ballinger Publishing Company, 1984.

Perry, George L. "Potential Output and Productivity", Brookings Papers on Economic Activity, 1: 11-47, 1977.

Russell, Louise B., The Baby Boom Generation and the Economy, The Brookings Institution, Washington D.C., 1982.

Stoker, Thomas M., "Simple Tests of Distributional Effects on Macroeconomic Equations", Journal of Political Economy, 94: 763-95, August 1986.

Tobin, James, "Liquidity Preference as Behaviour Towards Risk", Review of Economic Studies, 25: 65-86, February 1958.

U.S. Bureau of the Census, "Estimates of the Population of the U.S. by Age, Sex and Race: 1952-1985", Current Population Reports, Series P-25. (Nos. 310, 519, 917, 965, 985.) U.S. Government Printing Office, Washington, D.C..

Wachter, Michael L., "Intermediate Swings in Labor-Force Participation", Brookings Papers on Economic Activity, 2: 545-576, 1977.