Parsimonious Estimation of Age Structure Effects

David Bloom and David Canning*

September 2001

Abstract: Estimating age structure effects is often difficult due to the large number of variables involved and the high degree of multicollinearity between them. We suggest a set of criteria for selecting between different parsimonious specifications. Applying our criteria to a simple savings model leads to the selection of a specification that restricts age effects to lie on a quadratic curve rather than the use of aggregate age shares, such as the dependency ratio. The specification selected by our criteria, however, depends on the issue being addressed and we recommend a selection methodology rather than a specific parsimonious representation.

JEL Number: J1
Keywords: Demographic effects, dependency rates, polynomial.

Professor David Bloom
Harvard School of Public Health
677 Huntington Avenue
Boston, MA 02115
USA
email:dbloom@hsph.harvard.edu
fax: 1-617-5660365

Professor David Canning*
Queen's University, Belfast
University Road
Belfast, BT71NN
UK
email: d.canning@qub.ac.uk
fax: 44-28-90335156

*author for correspondence

1. Introduction

Social and economic behavior varies by age. For example, in numerous studies using individual or household data, labor supply, saving rates, unemployment rates, and criminal activity have all been found to be age specific. This suggests that when we look at aggregate data on these variables the age structure of the underlying population should be taken into account. Many such studies have been undertaken at the aggregate level, trying to identify how age structure affects behavior.

A recurrent problem in such studies is that a population's age structure is of high dimension. If we have data on age structure by age in years we have 101 variables (truncating at age 100). The United Nations Population Data (1998) gives age structure by 5 year intervals up to age 80 (aggregating those over 80 into a single category) which generates 17 age groups. Using such a large number of variables eats into the small number of degrees of freedom that are usually available in aggregate data. More importantly, the high degree of co-linearity between the populations shares in the different age groups means that the resulting estimated age structure coefficients are usually badly determined, generating very large standard errors that make inference difficult.

The high degree of co-linearity between the proportion of people in different age groups suggests that age structure can be captured by a small number of variables, and most studies that investigate age structure effects try to do this. The approach usually taken to compress the age structure variables is somewhat ad hoc. A particular representation which compresses the age structure effect is put forward and used without a great deal of testing. Two broad approaches can, however, be discerned. The first is to aggregate the data into a small number of age groups (e.g. Bloom and Sachs (1998), Lindh and Malmberg (1999) among many). This assumes that the effect is the same for people within the age group but different between age groups. The second is to assume that each age group has a different effect but that these effects lie on a low order polynomial (e.g. Fair and Dominguez (1991) and Higgins (1998)).

Two problems arise with such approaches. One is that a particular representation may be too parsimonious, and may lose interesting age structure behavior. A second is that the use of different approaches to representing age structure effects may give rise to somewhat different results and we may want to decide which approach is better. We will address both of these issues in this paper.

The problem of finding a parsimonious representation of a large set of co-linear variables is, of course, not new. The problem is similar to that of finding a parsimonious representation of a lag structure; indeed, the use of a low order polynomial is identical to Almon's (1965) use of polynomial distributed lags. Another common approach when faced by multicollinearity is to use principal components (see, for example, Maddala (2001) for a criticial discussion).

It is easy to show that the "best" parsimonious representation of age structure effects depends on the problem being investigated. The aim of this paper is to put forward a set of criteria of evaluating a proposed parsimonious representation of age structure effects. Our criteria follow closely Hendry's (1993) general to specific methodology. Using a parsimonious representation is equivalent to imposing set of restrictions on the coefficients on the age shares. We can test each proposed parsimonious representation against the obvious encompassing model that imposes no restrictions. When an approach gives rise to a sequence of nested models (e.g. polynomials of different orders) we can test each model against others in the sequence. This type of testing essentially ensures that for each approach we use a sufficient number of summary variables to capture the underlying age structure effects. We can then test between different approaches using non-nested methods.

We demonstrate the use of these criteria by means of an example, estimating a simple model explaining national savings by life expectancy, age structure and country fixed effects. We then check the adequacy of four different parsimonious representations of age structure; age groupings, a low dimensional polynomial, principal components and the results of a mechanical "parsing down" of age share variables.

The result of our criteria is to select the representation in which the age structure effects lie on a quadratic. If we eliminate this representation, the next representation selected would be age groupings (young (0-14) and old (64+)). It is clear in this example that the use of a more parsimonious representation (e.g. using only the dependency ratio) over constrains the estimation and is rejected by the data.

The relative failure of the principal components methodology is somewhat surprising at first site, given its theoretical appeal. The reason for its failure, and that of mechanical parsing, appears to be that they do not take into account the natural ordering of the age structure variables. Individual and household data (e.g. see Deaton and Paxson (1997), Paxson (1996)) suggests that savings rates vary reasonably continuously with age, so that people of similar ages act in similar ways. The low order polynomial imposes such continuity directly. The age groupings approach imposes that the effect of people within the same age group are identical while effects differ across groups, implying continuity of response at least within groups. Neither the principal components or parsing down methodologies exploit any continuity of the response by age. Indeed, the parsing down methodology tends to generate a very discontinuous function, with large effects in the few age categories that remain and zero effects being imposed elsewhere.

The result that a polynomial in age structure is selected in our example is, of course, not a general result[1]. Our criteria need to be applied anew in each application. What is clear is that using application specific information (such as continuity of behavior with age, or known discontinuities, for example at mandatory retirement age) can improve the fit of a parsimonious model relative to alternative approaches. We such information is used, the methodology presented here can still be used to test the resulting parsimonious representation against alternatives.

2. Parsimonious Representations.

We take the information from the United Nations Population Division (United Nations (1998)) as our basic age structure data. This gives population by five year age groups for 184 countries at five year intervals over the period 1950-1995 giving us 1840 data points in total. This data is mainly estimated and interpolated from census data. We construct age shares in each group by dividing through each cohort size by the total population.

Table 1 shows the correlation matrix between the different cohort shares in this data. As can be seen from this figure many of the age shares are highly correlated. This leads us to expect that their variation can be captured by a small number of summary statistics. Let us label the age share variables as x_0 to x_n is order of increasing age and suppose we wish to explain the variable y with a model of the form

$$y = c + \sum_{i=1}^{n} \beta_i x_i + \delta z + \varepsilon \quad (1)$$

where c is a constant and z represents another variable (or variables) in the relationship and ε is an error term. Note that we can drop the dependence of y on the first age group, $x_{0,}$ since the age shares add to one and including them all leads to co-linearity with the constant.

A parsimonious representation of this model is equivalent to imposing restrictions on the coefficients β_i. A set of n minus m restrictions on the coefficients define a null space, a n minus m dimensional subspace of the n dimensional space spanned by the share variables x_1 to $x_{n,}$ that should not matter for the determination of y. This is equivalent to saying that using m variables that span the m dimensional subspace orthogonal to the null space, instead of all n share variables, losses no explanatory power in equation (1). It follows that using m variables that are linear combinations of the n share variables in place of the actual share variables in equation (1) is exactly the same as imposing n-m restrictions on the coefficients in (1), and we can define a parsimonious model by either approach. We investigate four types of parsimonious representations.

The first is to use age groupings. We split the age shares into m groups and impose the restrictions that the coefficients β_i are the same within each grouping (n-m linear restrictions). This can be carried out in practice by forming m aggregate variables given by

$$g_j = \sum_{i=j-1}^{j} x_i \quad (2)$$

and estimating

$$y = c + \sum_{j=1}^{m} \alpha_j g_j + \delta z + \varepsilon \quad (3)$$

Again, the first group can be dropped because of co-linearity with the constant term. A difficulty with this approach is that it is not clear how to define the boundaries of each age group. One approach is to define the boundaries so as to make each group of roughly equal size. Another is to try to impose boundaries at points where behavior is thought to change (for example, at age 16 and age 65 which have legal significance for school leaving and retirement respectively in many countries). We use a compromise approach. For two groups we use the shares in the range 0-14 plus those 65+ versus those 25-64 as an approximation to the dependency ratio. With three groups we use 0-14, 15-64 and 65+, effectively separating out the effects of the young and the old. For four groups we use 0-14, 15-39, 40-64, and 65+. In general the choice of age groups may be problem specific, with breaks being imposed where discontinuities are suspected and with group sizes being smaller where the effect is most sensitive to changes in age. One advantage our the age grouping we have constructed is that they form a nested sequence from more to less restrictive.

The second approach is to impose the restriction that the coefficients β_i lie on a polynomial of order m. That is, for each i

$$\beta_i = \alpha_0 + \alpha_1 i + \alpha_2 i^2 + \alpha_3 i^3 + ... + \alpha_m i^m \quad (4)$$

If we take $\beta_i = f(i)$, with f continuous, this polynomial approximation of f(i) follows from Taylor's theorem. This approximation imposes n-m restrictions on the coefficients by imposing zero coefficients on the n-m higher order terms in the polynomial approximation that would be needed to completely span the age share space . Again, rather than impose these restrictions directly on the parameters in (1) a simple way to estimate this model is to construct the moment variables

$$m_j = \sum_{i=0}^{n} i^j x_i \quad (5)$$

and then estimate

$$y = c + \sum_{j=1}^{m} \alpha_j m_j + \delta z + \varepsilon \quad (6)$$

Note that m_0 is just the sum of the age shares, which is one, and can be subsumed into the constant term. The second, m_1, represents the mean age of the population.[2] It follows that this approach encompasses using the mean age of the population as a summary statistic for its age structure as advocated by Behrman, Duryea and Szekely (1999). Using m_1 and m_2 is equivalent to using mean age and the variance of age. Once we have estimated (6) the original implied restricted coefficients in (1) can be recovered via equation (4). We use only this polynomial form; of course alternatives, such as a polynomial spline or a harmonic distributed lag, could be investigated.

A third possible model is to construct the principal components of x_0 to x_n. The principal components p_0 to p_m can be written as orthogonal linear combinations of the share variables[3]

$$p_j = \sum_{i=1}^{n} \psi_{ij} x_i \quad (7)$$

and we can use the first m of these (those with the largest associated eigenvalues) to estimate

$$y = c + \sum_{j=1}^{m} \sigma_j p_j + \delta z + \varepsilon \quad (8)$$

Since each principle component is a linear combination of the share variables x_0 to x_n we can again recover the implied restricted coefficients in (1) by setting

$$\beta_i = \sum_{j=1}^{m} \sigma_j \psi_{ij} \quad (9)$$

This approach implies that the space spanned by the remaining n-m principal components does not matter for the determination of y and these are constrained to have zero coefficients in the estimation.

In the case of the mechanical "parsing down" methodology the variables selected (i.e. the restrictions imposed) can not be decided on ex ante. We remove the age share with the lowest t-ratio and re-estimate the regression, continuing to remove variables one at a time until all the remaining age share coefficients are statistically significant. This parsing results in an estimation including only m of the share variables implicitly imposing n-m zero restrictions on the others. This approach is not recommended, but is often followed in practice in an ad hoc manner when researchers are faced with multicollinearity; we include it for completeness.

One approach to deciding if a particular parsimonious representation is likely to be adequate is to try to invert it. That is, given a particular representation, can we recover the underlying share variables x_0 to x_n? The more co-linear the share variables the greater the amount of their variation that falls in a space of low dimension and the easier it is to represent them with a small number of summary variables. Table 2 shows the R^2 from a regression of each share variable on a single summary statistic, corresponding to m=1 in the above definitions. The final row is a standard statistic of the amount of the total variation in the x's (after normalization) captured by the summary variables defined as

$$1-\frac{\sum_{i=1}^{n} v_i' v_i}{\sum_{i}^{n} x_i' x_i} \quad (10)$$

where each v_i is the vector of residuals from a regression on x_i on the summary variable. As we would expect, the first principal component maximizes the overall fit as measured by the last row; by construction, it is the linear combination of age share variables that maximizes this fit. The average age comes close to this limit on the overall fit, while the dependency ratio (those in the ranges 0-14 and 65+, versus those 15-64) performs somewhat worse overall.

None of the single summary statistics do well at explaining the age shares between 20 and 29. The dependency ratio, while is does worse on overall fit, is markedly better than either the average age or the first principal component at fitting the 30-34 age share. This explains why the choice of a summary statistic varies with the problem at hand. For example, if y is very sensitive to the 30-34 age share the dependency ratio will perform much better than the other summary statistics despite its poor fit elsewhere in the age structure. In general, each set of summary variables implies different restrictions on the coefficients in equation (1) and when these restrictions are not satisfied the summary statistics do a poor job of measuring the actual effect of age structure on y.

Tables 3 and 4 repeat the exercise for two summary variables and three summary variables respectively. The first three principal components capture 87% of the variation in age structure across populations and seem to do a reasonably good job at fitting each age share. The age groups now outperform the cubic in age in terms of overall fit, though again this ranking is reversed in terms of fitting some particular age groups.

3. Criteria for Choosing Between Parsimonious Representations

No single parsimonious representation is always best. Each excludes part of the variation in the age shares and so potentially gives rise to a misspecification. In practice, we have to choose between representations in an effort to minimize the resulting misspecification.

Hendry (1993) advocates a general to specific modeling strategy. A general model allows the restrictions implicit in a specific model to be tested and makes model selection less ad hoc. A natural general model in our text is to all unrestricted age effects in (1). This leads to :

Criterion 1: We cannot reject the parsimonious representation against the alternative of allowing unrestricted age structure effects.

This involves a simple test of the restrictions at some conventional level of significance. Not rejecting, however, is not the same as accepting a particular parsimonious representation. Failure to reject may occur simply because of a lack of power in the test. Over restricting the model leads to inconsistency while relaxing the restrictions only leads to a loss of efficiency. This suggests a second minimal criteria

Criterion 2: We cannot reject the parsimonious representation against the alternative of adding an extra term of the same form.

In the case of principals components this involves adding the next component (the one with the highest remaining eigenvalue). For a low order polynomial we add the next term in the expansion. In addition, we have a natural ordering in both these cases; we can order polynomial by their highest indices and can order principal components by their eigenvalues. In both cases the effect of adding an extra term is to remove a restriction in the estimation.

In the case of age groups what is meant by “an extra term” is less clear. If we redraw all the age group boundaries to maintain equal group sizes the new model has fewer restrictions than the old, but these are a new set of restrictions and need not be strictly weaker. A strict relaxation of the restrictions involves breaking up an existing age group into two components. Even here they is some arbitrariness about the choice of which age groups to split. However, adding groupings by breaking an existing grouping into two again allows a natural ordering from more to less restrictive groupings.

The general to specific approach will automatically satisfy criterion 2 when parsing down of the regression ends since a variable is only excluded if this is not rejected by the data. However, criterion 1 still needs to be addressed in this case and may led to variables that appear statistically insignificant on their own being retained.

Criteria 1 and 2 are minimal criteria. For each approach there is always a specification that satisfies these criteria (for example, setting m=n, so that the summary variable have the same dimension as the share variables imposes no restriction and is always acceptable). If parsimony is a goal, than we want to select the most parsimonious specification of a particular approach that meets the minimal criteria given by 1 and 2.

Criterion 3: For a given approach, select the most parsimonious model that satisfies criteria 1 and 2.

However, this leads to the issue of how are we to choose between approaches. One seemingly attractive option is simply to select the approach that generates the most parsimonious (i.e. smallest number of summary variables) by criterion 3. However, it is not clear that a smaller model is necessarily a better model. Choosing between approaches poses a different problem from testing with a particular approach. Within an approach the testing structure is nested, in the case of the polynomial curve fitting, and principal components, by the natural ordering of the summary variables, and in the case of testing down from general to specific by the methodology itself. Choosing between approaches, however, involves a non-nested test. A natural way of carrying out such a test is to form the encompassing model and testing the restriction that the competing model does not matter; i.e. it has zero coefficients in the hybrid structure. This gives us:

Criterion 4. We cannot reject the parsimonious representation against the alternative of an encompassing model that adds a competing specification selected by criterion 3.

It is possible to accept two different models using criteria 4. In this case both give a satisfactory parsimonious representation of the age structure effects and we can ague in each case that no extra information is obtained by adding the alternative approach. Equally, it is possible that neither model is acceptable using criteria 4. In this case the encompassing model, or some enlarged variant of one of the competing models, should be investigated as a parsimonious representation.

4. An Example: National Savings

We apply our criteria to a simple model of national savings. The effects of age structure on savings rates have been studied in numerous papers (e.g. Auerbach and Kotlikoff (1992), Fry and Mason (1982), Higgins and Williamson (1997), Kelly and Schmidt (1996), Leff(1969), Mason A. (1988)). All of these papers choose an age structure specification in a fairly ad hoc manner.

For our example we calculate the gross domestic savings rate from the World Bank's *World Development Indicators 1999.* Following Bloom, Canning and Graham (2000) we model the average national savings rate in each five year period as depending on beginning of period life expectancy (calculated from the *World Development Indicators)*, and age structure (from United Nation's *Demographic Indicators 1950-2050).* Time dummies and country fixed effects are also included. Life expectancy is included because Bloom, Canning and Graham (2000) argue that age specific savings rates may rise with life expectancy. Life expectancy is highly correlated with an age structure that is skewed towards older cohorts and omitting it could bias our age structure results. The model is still excludes a large number of variables that may be important (e.g. see Doshi K. (1994), Masson, Bayoumi and Samiei (1998) how estimate more general models), however our purpose is expositional, to show how our criteria can be applied to a particular example, rather than an effort to estimate the correct savings model.

We begin by estimating the unrestricted model. The resulting estimated age structure effects, along with their confidence intervals (taken as plus and minus 1.96 times the associated standard errors) are plotted in figure 1. The high degree of multicollinearity between the different cohort shares makes the confidence intervals very large.

Next we follow each of our four approaches to obtain a parsimonious specification that satisfies criteria 1 to 3. For the fitted polynomial in age and age grouping approaches we find that at least two terms are required. With principal components we need four terms. Parsing down, using the general to specific methodology, is much less parsimonious and leaves us requiring six terms. The results of the unrestricted regression and the regressions selected by criteria 1 to 3 for each of our four approaches are reported in table 5. The final two rows give tests for each of our parsimonious models, testing firstly that parsimonious model against the unrestricted model and then the parsimonious model against the effect of adding one extra term. In each case we cannot reject our parsimonious model, implying that each model satisfies criteria 1 and 2.

Figures 2 to 5 shows the estimates for the age share coefficients obtained under the restrictions implied by each of the models in table 5. While each model is not rejected by the data, it is important to note that the shape of the relationship in the is quite sensitive to which approach we use. Note that, while each parsimonious model gives results that to some extent track the unrestricted estimates, "fitting" the unrestricted coefficient estimates is not one of our required criteria; these unrestricted estimates have very large standard errors (as shown in figure one) and may be quite different from the true parameter values we are trying to find.

We now test to select between approaches. Table 6 shows a test of each model against alternative hybrid models that combine a pair of approaches. An F statistic in excess of the critical value implies that we reject the model being considered under the null against the alternative hybrid model. The only model not rejected by any hybrid model is the quadratic in age. On the other hand, all the other representations are rejected against some hybrid alternative. In particular, they are all rejected against a hybrid that includes the quadratic in age. This is despite the fact that the quadratic in age is one of the two most parsimonious models, using only two summary variables, while the principal components model uses four and the general to specific approach uses six.

Our selection criteria therefore selects the quadratic in age. It cannot be rejected against unrestricted estimation of the age structure effects, or against a cubic in age. Neither can it be rejected against an encompassing model that includes one of the alternative approaches.

Both the quadratic in age and two age group representations do reasonably well according to our criteria giving reasonable parsimonious representations with just two summary variables. The relative failure of the other approaches is somewhat surprising. In particular, using principal components captures most of the variation in the age shares in a small number of variables and should, a priori, provide a very parsimonious representation.

An intuitive explanation of what is happening is that the underlying age effects are in fact continuous in age. Certainly, average savings rate of age groups in individual and household studies vary quite continuously. This suggests that a polynomial in age exploits this continuity of the effect to give a very parsimonious representation. Age groupings also work well because they also impose some continuity of effect within groups. Principal components treats each age share symmetrically, not taking account of their natural ordering and does not impose any structure across age share effects. The general to specific approach imposes zero coefficients on many age shares and, as the spikes in figure 5 show, implies that effects are very discontinuous with age. Imposing such discontinuities does not fit the data very well.

Exploiting the continuity of the effects with age seems to be a good strategy in this example; of course, in general, the effects may actually not be continuous and a different approach to parsimonious representation may be better. The quadratic selected, together with its associated confidence interval is drawn in figure 6. The estimated confidence interval is clearly much smaller than in figure 1. This is the efficiency gain from using a parsimonious specification. It is important to realize, however, the estimates and confidence interval we obtain when we estimate a parsimonious model are *conditional* on the model being correctly specified. Our selection criteria ensures we have no strong evidence that our selected model is false. However, if we were simply to use a different representation, without testing, we are likely to make incorrect inferences.

5. Conclusion

In a linear model a parsimonious representation of age structure effects is equivalent to a set of restrictions on the coefficients on age each share. We can test the validity of a particular proposed parsimonious representation by testing these restrictions. We can test between different representations by forming an encompassing model and test each against this hybrid. This ensures that the final formulation we use is compatible with the data.

At present, different papers impose different parsimonious specifications, usually either age groupings or a quadratic or cubic polynomial. The methodology presented here gives a straightforward mechanism for deciding which specification is most appropriate in each setting. It is worth reiterating that the appropriate parsimonious representation is entirely dependent on the relationship being estimated. We have a methodology for selecting an appropriate representation rather than a prescription for which representation to use.

References

Behrman J.R. , Duryea S. and Szekely M. (1999), “Aging and Economic Opportunities; Major World Regions Around the Turn of the Century,” Inter-American Development Bank, Working Paper #405.

Bloom, D., and Sachs, J. (1998): “Geography, Demography, and Economic Growth in Africa,” *Brookings Papers on Economic Activity,* 2: 207–273.

Deaton A. and Paxson C. (1997): “The Effects of Economic and Population Growth on National Savings and Inequality,” *Demography,* 34: 97–114.

Doshi K. (1994), “Determinants of the Saving Rate – An International Comparison,” *Contemporary Economic Policy*,” Vol. 12, pp 37-45.

Fair R.C. and Dominguez K.M. (1991), “Effects of the Changing U.S. Age Distribution on Macroeconomic Equations,” *American Economic Review*, Vol. 81, pp 1276-94.

Fry, M., and Mason, A. (1982): “The Variable Rate of Growth Effect in the Life-cycle Model,” *Economic Inquiry* 20: 426-42.

Hendry, D.F. (1993), *Econometrics: Alchemy or Science?* Oxford, Blackwells.

Higgins M. and Williamson J.G. (1997), "Age Structure Dynamics in Asia and Dependence on Foreign Capital," *Population and Development Review*, Vol. 23, pp 261- 293.

Higgins M. (1998), "Demography, National Savings, and International Capital Flows," *International Economic Review*, Vol. 39, pp 343-369.

Kelly A.C. and Schmidt R.M. (1996), "Saving, Dependency and Development," *Journal of Population Economics*, Vol. 9, pp 365-386.

Leff, Nathaniel H. (1969), "Dependency Rates and Savings rates," *American Economic Review*, Vol. 59, pp 886 – 96.

Lindh Thomas and Malmberg Bo, (1999), "Age Structure Effects and Growth in the OECD, 1950-1990, *Journal of Population Economics*, Vol .12, pp 431-449.

Maddala G.S. (2001), *Introduction to Econometrics*, third edittition, Chichester, John Wiley and Sons.

Mason A. (1988), "Saving, Economic-Growth, and Demographic Change," *Population and Development Review*, Vol. 14, pp 113-144.

Masson P.R., Bayoumi T. and Samiei H. (1998), "International Evidence on the Determinants of Private Saving," *World Bank Economic Review*, Vol. 12, pp 483-501.

Paxson C.H. (1996), "Saving and Growth: Evidence for Micro Data," *European Economic Review*, Vol. 40, pp 255-288.

United Nations, (1998): *Demographic Indicators 1950-2050,*. New York, United Nations.

World Bank, (1999), *World Development Indicators 1999*, Washington D.C., World Bank.

Table 1
Corrolations in Age Shares

Ages	0-4	5-9	10-14	15-19	20-24	25-29	30-34	35-39	40-44
0-4	1.000								
5-9	0.923	1.000							
10-14	0.749	0.875	1.000						
15-19	0.515	0.598	0.788	1.000					
20-24	0.158	0.119	0.234	0.549	1.000				
25-29	-0.212	-0.293	-0.355	-0.130	0.517	1.000			
30-34	-0.467	-0.489	-0.569	-0.531	0.005	0.655	1.000		
35-39	-0.664	-0.677	-0.652	-0.640	-0.277	0.301	0.713	1.000	
40-44	-0.783	-0.802	-0.744	-0.618	-0.362	0.073	0.443	0.770	1.000
45-49	-0.814	-0.849	-0.794	-0.651	-0.330	0.016	0.292	0.588	0.829
50-54	-0.818	-0.846	-0.811	-0.665	-0.380	0.035	0.264	0.487	0.715
55-59	-0.839	-0.842	-0.793	-0.673	-0.430	-0.034	0.287	0.496	0.681
60-64	-0.850	-0.851	-0.770	-0.654	-0.432	-0.069	0.245	0.523	0.689
65-69	-0.853	-0.851	-0.774	-0.625	-0.409	-0.064	0.215	0.484	0.706
70-74	-0.843	-0.848	-0.778	-0.626	-0.380	-0.041	0.216	0.452	0.666
75-80	-0.830	-0.832	-0.767	-0.625	-0.374	-0.015	0.236	0.449	0.633
80+	-0.818	-0.803	-0.734	-0.595	-0.344	0.009	0.277	0.470	0.619

Ages	45-49	50-54	55-59	60-64	65-69	70-74	75-80	80+
45-49	1.000							
50-54	0.868	1.000						
55-59	0.812	0.907	1.000					
60-64	0.787	0.868	0.932	1.000				
65-69	0.791	0.841	0.899	0.941	1.000			
70-74	0.798	0.840	0.885	0.909	0.947	1.000		
75-80	0.752	0.827	0.860	0.891	0.919	0.951	1.000	
80+	0.701	0.746	0.801	0.842	0.882	0.910	0.944	1.000

Table 2
Age Structure Variances Explained by a Single Summary Variable

Age Group	Average Age	Dependency Ratio	1 Principal Component
0-4	0.885	0.831	0.836
5-9	0.896	0.906	0.878
10-14	0.754	0.817	0.804
15-19	0.479	0.488	0.578
20-24	0.119	0.019	0.161
25-29	0.011	0.163	0.011
30-34	0.163	0.432	0.197
35-39	0.409	0.619	0.465
40-44	0.646	0.698	0.688
45-49	0.771	0.713	0.791
50-54	0.825	0.683	0.827
55-59	0.867	0.655	0.859
60-64	0.886	0.627	0.868
65-69	0.895	0.587	0.867
70-74	0.888	0.570	0.858
75-80	0.864	0.547	0.834
80+	0.806	0.511	0.775
Total	0.657	0.580	0.664

Table 3
Age Structure Variances Explained by Two Summary Variables

Age Group	Quadratic in Age	Young and Old	2 Principal Components
0-4	0.936	0.895	0.852
5-9	0.942	0.947	0.916
10-14	0.769	0.838	0.868
15-19	0.485	0.513	0.588
20-24	0.347	0.218	0.439
25-29	0.454	0.451	0.822
30-34	0.533	0.604	0.793
35-39	0.639	0.657	0.661
40-44	0.726	0.702	0.703
45-49	0.782	0.760	0.793
50-54	0.826	0.787	0.842
55-59	0.880	0.832	0.889
60-64	0.914	0.875	0.912
65-69	0.940	0.945	0.917
70-74	0.949	0.964	0.905
75-80	0.938	0.955	0.874
80+	0.875	0.911	0.797
Total	0.761	0.756	0.798

Table 4
Age Structure Variances Explained by Three Summary Variables

Age Group	Cubic in Age	Young, Old and Middle Aged	3 Principal Components
0-4	0.969	0.896	0.897
5-9	0.943	0.947	0.958
10-14	0.807	0.839	0.868
15-19	0.679	0.527	0.804
20-24	0.594	0.454	0.915
25-29	0.512	0.806	0.874
30-34	0.533	0.757	0.872
35-39	0.665	0.661	0.820
40-44	0.779	0.769	0.746
45-49	0.842	0.879	0.793
50-54	0.875	0.892	0.847
55-59	0.905	0.898	0.892
60-64	0.921	0.911	0.918
65-69	0.940	0.952	0.936
70-74	0.955	0.965	0.937
75-80	0.961	0.956	0.907
80+	0.945	0.928	0.827
Total	0.813	0.826	0.871

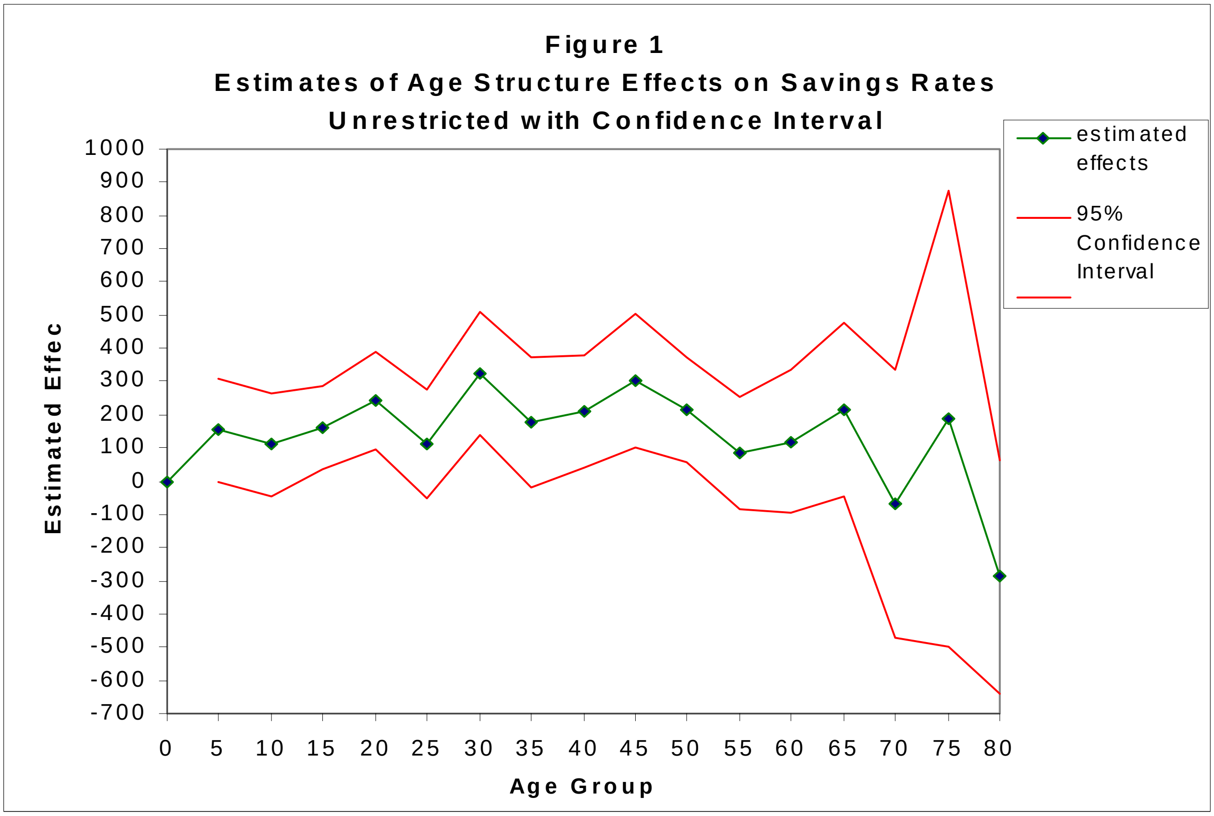
Figure 1
Estimates of Age Structure Effects on Savings Rates
Unrestricted with Confidence Interval
estimated effects
95% Confidence Interval
Estimated Effec
1000
900
800
700
600
500
400
300
200
100
0
-100
-200
-300
-400
-500
-600
-700
0 5 10 15 20 25 30 35 40 45 50 55 60 65 70 75 80
Age Group

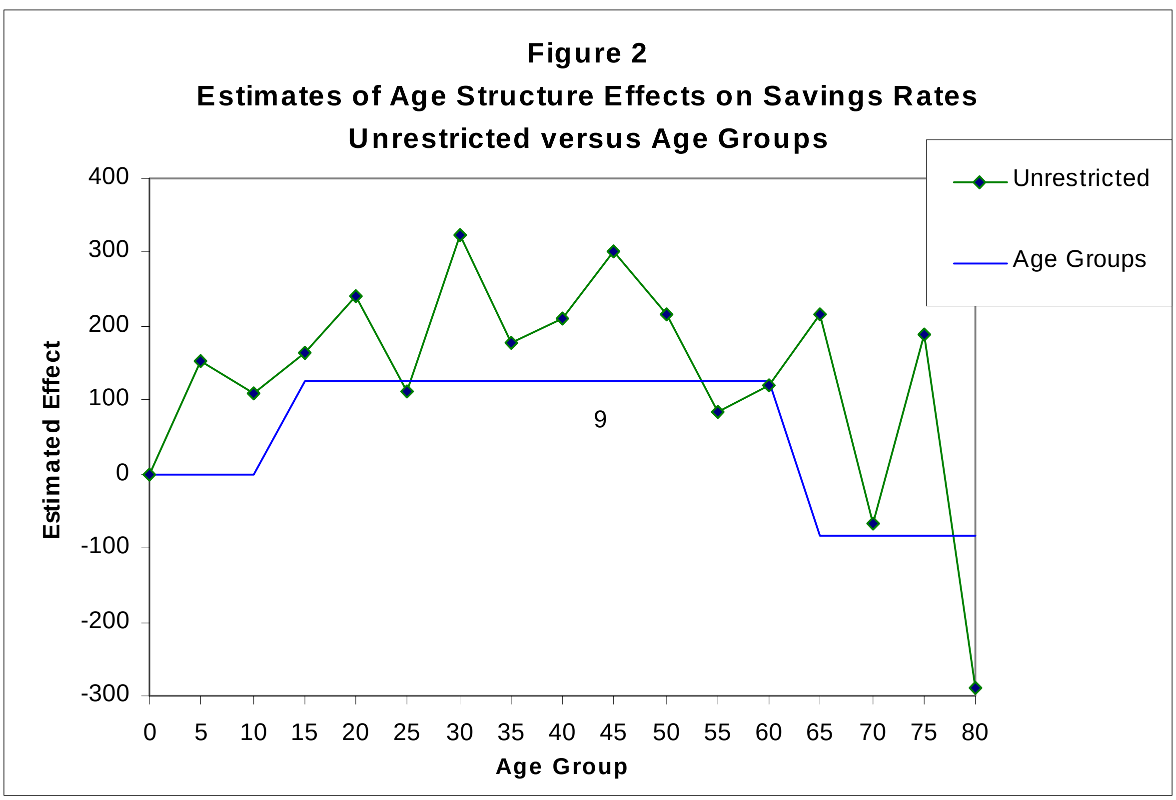
Figure 2
Estimates of Age Structure Effects on Savings Rates
Unrestricted versus Age Groups
Unrestricted
Age Groups
Estimated Effect
400
300
200
100
0
-100
-200
-300
9
0 5 10 15 20 25 30 35 40 45 50 55 60 65 70 75 80
Age Group

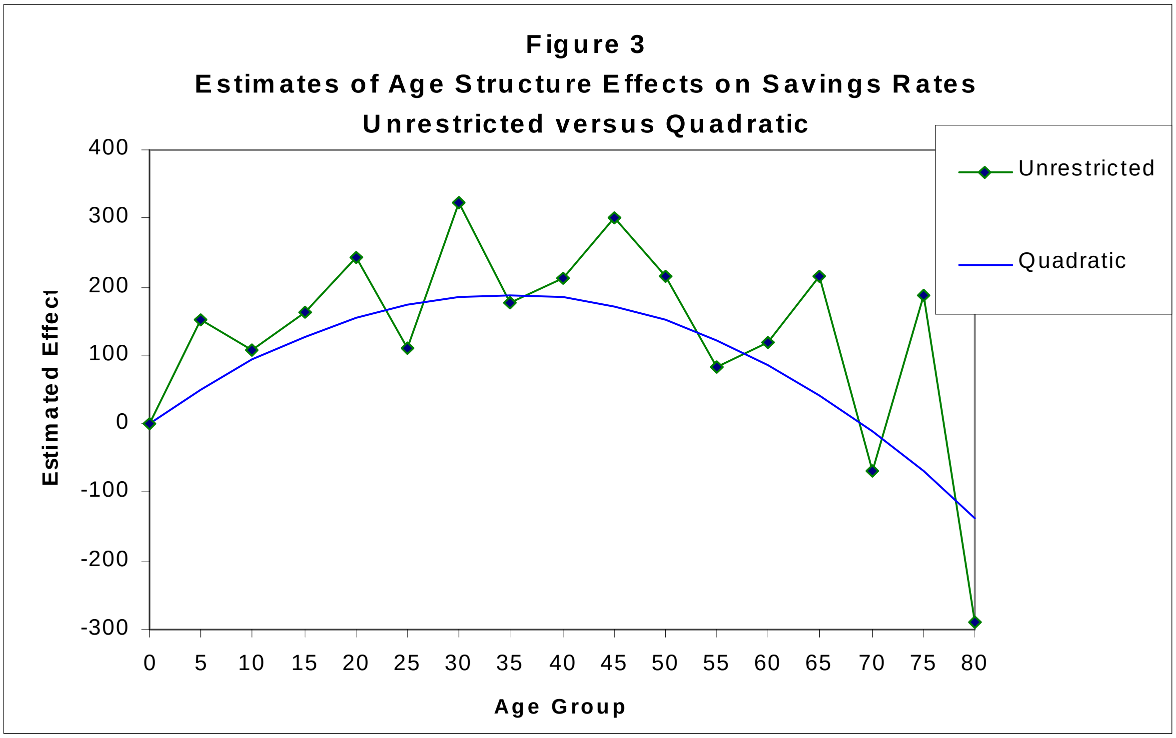
Figure 3
Estimates of Age Structure Effects on Savings Rates
Unrestricted versus Quadratic
Unrestricted
Quadratic
Estimated Effect
400
300
200
100
0
-100
-200
-300
0 5 10 15 20 25 30 35 40 45 50 55 60 65 70 75 80
Age Group

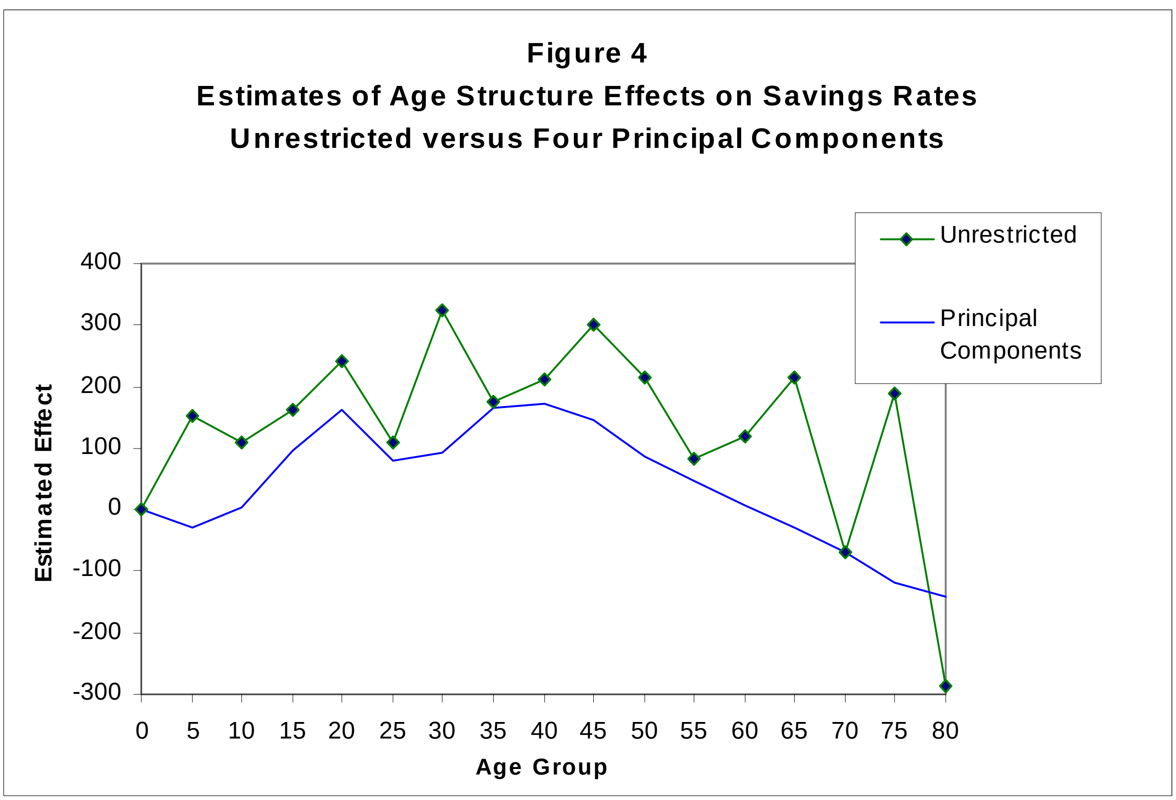
Figure 4
Estimates of Age Structure Effects on Savings Rates
Unrestricted versus Four Principal Components
Unrestricted
Principal Components
Estimated Effect
400
300
200
100
0
-100
-200
-300
0 5 10 15 20 25 30 35 40 45 50 55 60 65 70 75 80
Age Group

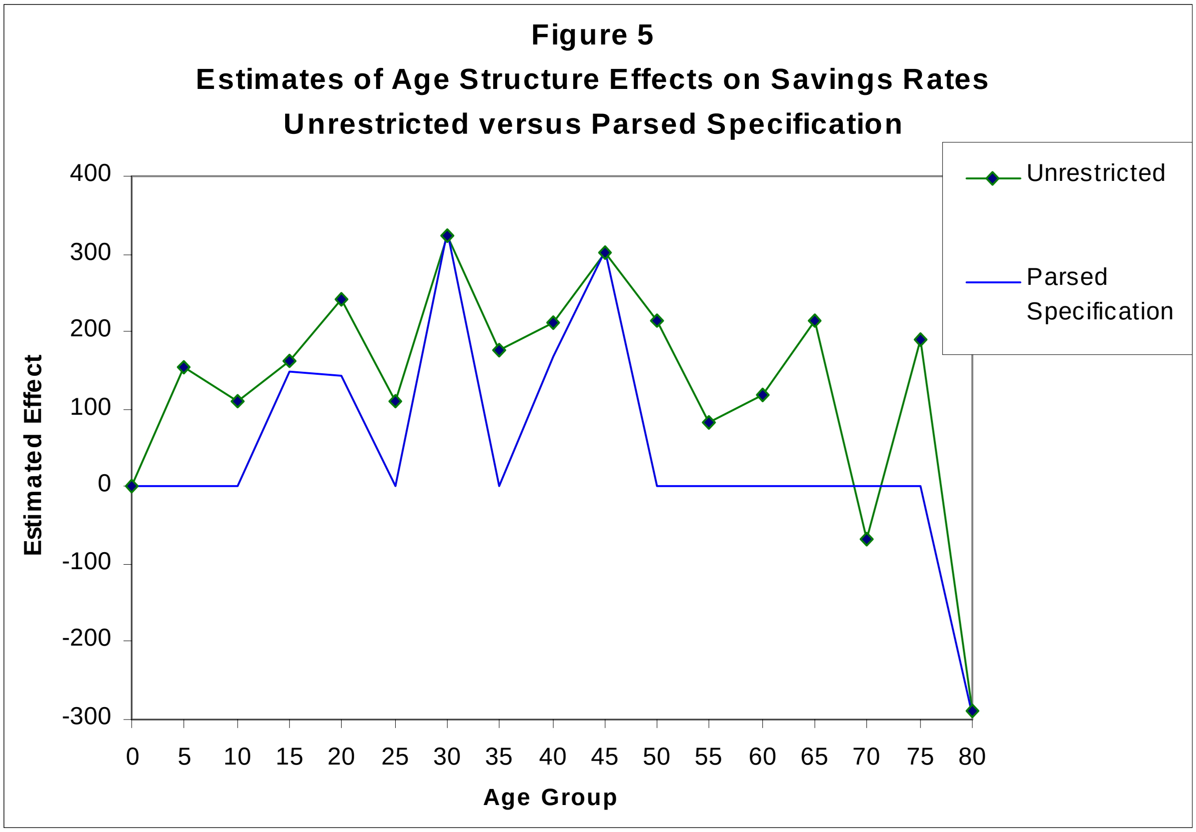
Figure 5
Estimates of Age Structure Effects on Savings Rates
Unrestricted versus Parsed Specification
Unrestricted
Parsed Specification
Estimated Effect
400
300
200
100
0
-100
-200
-300
0
5
10
15
20
25
30
35
40
45
50
55
60
65
70
75
80
Age Group

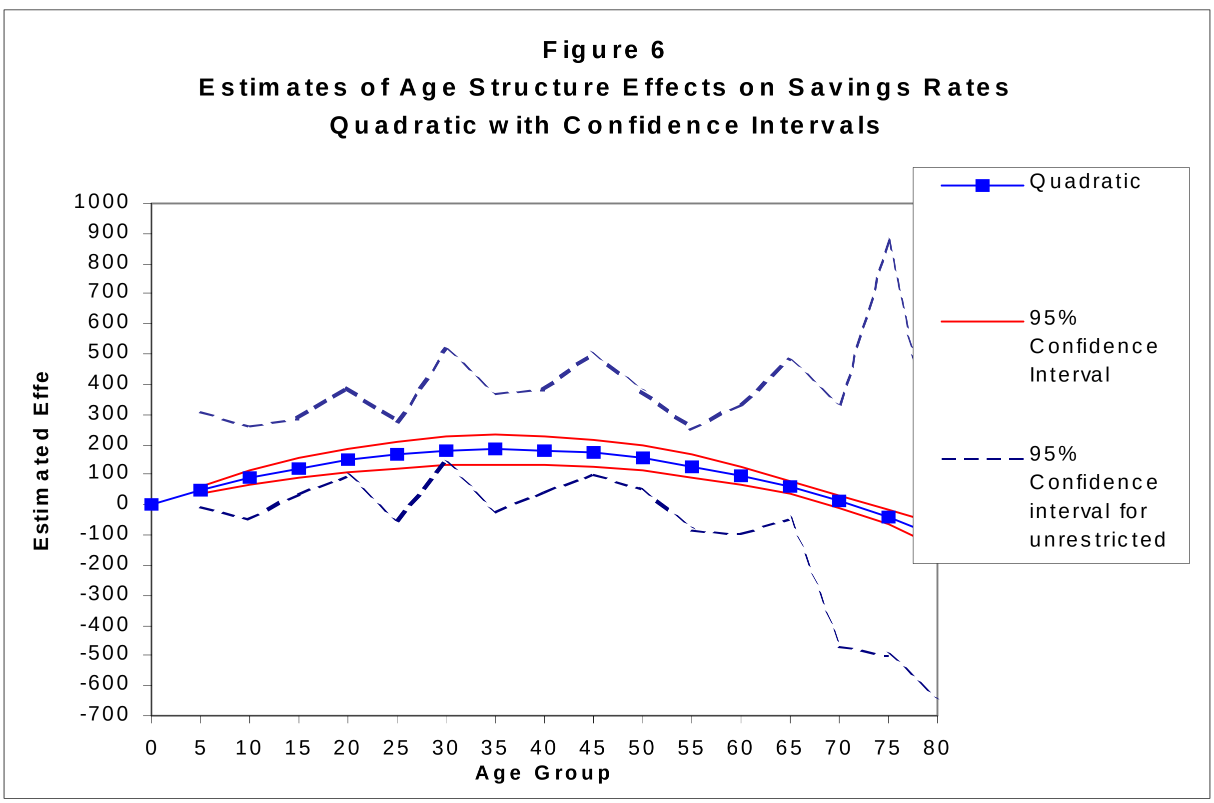
Figure 6
Estimates of Age Structure Effects on Savings Rates
Quadratic with Confidence Intervals
Quadratic
95% Confidence Interval
95% Confidence interval for unrestricted
Estimated Effe
1000
900
800
700
600
500
400
300
200
100
0
-100
-200
-300
-400
-500
-600
-700
0
5
10
15
20
25
30
35
40
45
50
55
60
65
70
75
80
Age Group

[1] Bloom, Canning and Graham (2000) find in a more complex model of saving that age groupings appear to outperform the polynomial specification.

[2] Actually, $2.5+5m_1$ is the mean age but this in co-linear with $m_{1.}$

[3] In fact we follow the usual custom and construct the principal components from the normalized share variables where normalization implies they are linearly transformed to have mean zero and variance one.