Household Saving over the Life Cycle: International Evidence from Micro Data *

Oleksandr Movshuk
Department of Economics, University of Toyama,
3190 Gofuku, Toyama, 930-8555, Japan
E-mail: movshuk@eco.u-toyama.ac.jp

May 21, 2009

Abstract

In this paper I estimate age-saving profiles from micro data in six countries (Italy, Japan, Taiwan, Thailand, the UK, and the US) to verify whether households are saving as postulated by the life- cycle theory. The level of household savings depends on age, cohort and year effects, and the perfect collinearity among these effects is broken by applying a nonparametric regression model. In this model, the cohort effect is assumed to be an arbitrary smooth function, and the model is estimated by the generalized additive model with a penalized smoothing spline approach. Estimated saving-age profiles showed declining savings in the old age for the majority of examined countries. An interesting feature for Asian households was a double hump in savings, with a temporal dip for households for the age bracket at around mid-40s.

JEL classification: D91; E21
Keywords: Life-cycle hypothesis, household savings.

*This work was supported by a Grant-in-Aid for Scientific Research (No.19530180) from the Japan Society for the Promotion of Science. I am grateful to the Research Centre for Information and Statistics of Social Science, Institute of Economic Research, Hitotsubashi University for arranging the use of the National Survey of Family Income and Expenditure of Japan.

1 Introduction

This paper examines the saving behavior of households in six countries, and presents new evidence that age-saving profiles generally agree with predictions of the life-cycle hypothesis of Modigliani and Brumberg (1954). The life cycle hypothesis predicts that age-saving profiles have hump shape, with individuals saving most from the middle age to their retirement, and dissaving in young and old ages. Though I did not observe this hump-saving in each of six examined countries, the evidence of life-cycle savings agreed with theoretical predictions much more compared with previous studies that estimated age-saving profiles from micro data (Poterba, 1994; Borsch-Supan, 2003; Deaton and Paxson, 1994; Paxson, 1996). In particular, while the theory predicts negative savings among the elderly, most previous studies concluded that savings remained either flat, or even increasing, for aged households.

Most recent studies of household saving behavior follow Deaton and Paxson (1994) and study savings as a combination of age, cohort, and year effects. This decomposition produces estimates of age-saving profiles of households from estimates of age effect, but it also creates an identification problem due to the perfect collinearity among age, cohort, and year effects, since for every birth cohort, its year of birth is exactly the current year less its current age. The identification problem can be solved by imposing some restrictions on the data. A particularly popular approach follows Deaton and Paxson (1994), who suggested to impose orthogonality restrictions on year effects. The solution was used in many micro studies of household savings (Borsch-Supan, 2003; Paxson, 1996). In this paper, I apply an alternative solution that restricts cohort effect to be a smooth function that has no sudden jumps, and whose shape can be estimated by a nonparametric regression model. In contrast to the Deaton-Paxson approach, the smoothing cohort model leaves year effect unrestricted.

Applying the smoothing cohort model, I found that age-saving profiles showed a dip of around 20 percentage points among aged households in the United and Japan. The dip in the old age was also evident among household in Italy and Taiwan, but was less significant, at around 10 percentage points. On the other hand, age profiles of savings in the United Kingdom and Thailand turned out much more volatile, with much less clear evidence of dissaving in the old age. Another noteworthy finding was declining savings not only in the old age, but also in the middle age, especially among Asian households. This finding may indicate that changes in household behavior reflect not just the retirement motive (as postulated by the stripped-down model of household savings), but also motives that require substantial dissaving in the middle age, such as housing purchases, and support for children's education.

2 Model

I begin with the conventional model of Deaton and Paxson (1994), in which household savings depend on age, cohort and year effects. Consider a household that is observed in year t, with the head of household aged a and born in year b. Birth cohorts are defined by the year of birth of household head. The model explains the saving rate y, which equals the difference between disposable income and consumption, normalized by disposable income.

The shape of age, cohort and year effects on savings is not specified, and estimated by three sets of dummy variables for age, birth cohort, and current year. For example, age effects for ages between 25 and 70 are estimated with $70-25+1=46$ dummy variables for each age in this age span. Let this matrix of age dummies be D_a. Cohort and year effects are similarly defined by matrixes of dummy variables D_c and D_t.

The Deaton-Paxson model combines these three effects on savings, and yields

$$y=\beta_0+\sum_{a=1}^{A}\beta_a D_a+\sum_{c=1}^{C}\beta_c D_c+\sum_{t=1}^{T}\beta_t D_t+\varepsilon \tag{1}$$

For each dummy matrix D_a, D_c, and D_t, the sum across rows is always one, which results in the perfect collinearity between D_a, D_c, and D_t and the intercept term β_0. Typically, the problem is solved by dropping a single dummy variable from each of D_a, D_c, and D_t (such as the first age effect in D_a, and similarly for D_c, and D_t). The dropping of first terms in D_a, D_c, and D_t is another way to restrict parameters for the dropped terms to zero. This makes the dropped terms a benchmark to interpret estimates of β_a, β_c, and β_t. Suppose that age effect is estimated for age span 25 to 70, and the omitted dummy variable is for age 25. Then a positive estimate for dummy variable for age 26 means that compared with the benchmark age 25, savings are increased at age 26.

Using a particular age, cohort or year as a benchmark is not helpful in interpreting estimated parameters β_a, β_c, and β_t. More informative benchmark is produced by an alternative restriction that the sum of estimated coefficients for each of three effects is zero (Suits, 1984):

$$\sum_{a=1}^{A}\beta_a=\sum_{c=1}^{C}\beta_c=\sum_{t=1}^{T}\beta_t=0 \tag{2}$$

This approach keeps the full set of dummy variables for age, cohort and year effects, but restricts their sum to zero. The zero benchmark level is associated with the average effect across the full span of dummy variables for age, cohort or period effects. Then positive estimates of, say, age effect show positive deviations from the average saving rate across the estimated life cycle.

3 Identification problem and its solutions

Identification problem occurs in model (1) even with restriction (2), due to the exact linear relation between the current year t, age a and year of birth b (namely, $t=a+b$). Because of this perfect collinearity, it is impossible to find a unique explanation for examined data. Suppose, for example, that saving rate is increasing by 3 percent a year. This trend can be explained by year effects in savings that increase by 3 percent per year, with no changes in age and cohort effects. Another possible interpretation is by a combination of increasing age and cohort effects, with 3 percent growth per year of age, and the same 3 percent increase in each younger cohort, and no contribution from year effects. Deaton and Paxson (1994) and Paxson (1996) provide similar examples how the identification problem leads to alternative interpretations of observed trends in data.

The identification problem can be avoided by imposing restrictions on estimated regression coefficients in (1). The most common solution in studies of household savings follows Deaton and Paxson (1994), who suggested restrictions on year effects. Namely, Deaton and Paxson proposed the following two restrictions: (1) year effects are orthogonal to a linear time trend, and (2) the sum of year effects is zero. The first restriction is crucial, while the second restriction is not (in fact, it is identical to restriction on year effects in (2)). Due to the orthogonality restriction, any linear trends in data are removed from year effects, and attributed to both age and cohort effects. For example, in the previous example of the 3 percent growth in saving rate, the Deaton-Paxson approach will choose the second interpretation, with 3 percent increase in both age and cohort effects, and no growth in year effect. Essentially, the Deaton-Paxson approach postulates that time effects contain only cyclical variation. If any trends appear in data, they are forced to appear in age and cohort effects, since only these two effects are unrestricted. In consequence, the Deaton-Paxson approach may result in spurious trends in age and cohort effect, masking their original patterns.

In this paper I will use an alternative solution to the identification problem. The solution restricts the pattern of cohort effect, and leaves age and year effects unrestricted, so in contrast to the Deaton-Paxson approach, estimates of year effect may contain any kind of trend. The cohort effect is restricted to an arbitrary smooth function, with no sudden jumps, which can be estimated by a nonparametric regression. The solution is called the smoothing cohort model, and was suggested by Fu (2008). Essentially, the smoothing cohort model replaces the matrix of cohort dummies D_c in (1) with a single variable c for birth cohorts, and its effect is allowed to be nonlinear. The introduction of a smooth nonlinear function $f(c)$ in (1) produces the following smoothing cohort model of saving rate y:

$$y = \beta_0 + \sum_{a=1}^{A} \beta_a D_a + f(c) + \sum_{p=1}^{P} \beta_p D_p + \varepsilon \tag{3}$$

3.1 Estimation

The smoothing cohort model (3) is essentially a semiparametric regression model that consists of two parts: a nonparametric component $f(c)$ and a parametric component that combines two sets of dummy variables D_a and D_p. Originally, Fu (2008) suggested to estimate the smoothing cohort model as a generalized additive model (GAM). The GAM can fit regression model (3) by the backfitting algorithm of Hastie and Tibshirani (1990). However, in recent years the stability of the backfitting algorithm was questioned, particularly in datasets with high collinearity among explanatory variables (Schimek, forthcoming). Another limitation of the GAM estimator is the need to select a smoothing parameter (namely, the number of degrees of freedom λ). While Fu (2008) claimed that setting λ to 10 degrees of freedom 'yields good results' (p. 341), it is preferable if the degree of smoothing of $f(c)$ is determined endogenously, depending on analyzed data.

In this paper the smoothing cohort model (3) is estimated by the Modified Generalized Cross Validation (MGCV) algorithm of Wood (2004, 2006). The MGCV has superior numerical stability compared to the backfitting algorithm (*ibid.*). In addition, the algorithm selects smoothing parameters endogenously, by minimizing the prediction error criteria (such as the generalized cross validation or Akaike Information Criteria). In this section I discuss the estimation of the smoothing cohort model by the MGCV.

Consider a reduced specification of (3), with only nonparametric term $f(x_i)$. Once this basic case is introduced, its extension to the full semiparametric model is trivial. In the reduced specification, the dependent variable y depends on a single explanatory variable x:

$$y = f(x) + \varepsilon_i \tag{4}$$

where $f(\cdot)$ is an arbitrary smooth function and ε_i is the error term with zero mean and variance σ^2.

Let $\kappa_1 < ... < \kappa_K$ be a sequence of breakpoints ('knots') that are distinct numbers that span the range of x. The smooth function $f(x)$ is estimated by cubic splines, which are cubic piecewise polynomials that are joined at the 'knots'. Due to special restrictions, these polynomials join smoothly at the knots, and also have continuous first and second derivatives. Let K denote the number of knots. The a cubic spline can be represented by truncated cubic basis functions:

$$s(x) = \beta_0 + \beta_1 x + \beta_2 x^2 + \beta_3 x^3 + \sum_{k=1}^{K} \beta_{k+3}(x - \kappa_k)_+^3$$

where

$$(x - \kappa_k)_+ = \begin{cases} 0, & x \le k \\ (x - \kappa_k), & x > k \end{cases}$$

In this representation, the cubic spline has a simple interpretation, as a global cubic polynomial $\beta_0 + \beta_1 x + \beta_2 x^2 + \beta_3 x^3$ and a set of K local polynomial deviations $\sum\limits_{k=1}^{K} (x - \kappa_k)_+^3$. In matrix form, the truncated cubic basis becomes $y = X\beta + \varepsilon$, where X is design matrix with ith row $x_i = \begin{bmatrix} 1 & x_i & x_i^2 & x_i^3 & (x_i - \kappa_1)_+^3 & \dots & (x_i - \kappa_K)_+^3 \end{bmatrix}$, β is the corresponding vector of regression parameters, and ε is the error term. The smooth function $f(x,\beta)$ is linear in $K+4$ regression parameters, and can be fitted by minimizing the sum of squared residuals $(y - X\beta)^T (y - X\beta) = \|y - X\beta\|^2$.

By increasing the number of knots K, the model becomes more flexible in approximating y. But if the number of knots is too large, the estimates $\hat{f}(x)$ may follow y too closely. In the limit, when $K = n$, the cubic spline simply interpolates y. To prevent too much wiggliness in the estimated curve, a special term that penalizes rapid changes in $\hat{f}(x)$ is added to the fitting criteria. The most common penalty is by $\lambda \int [f''(x)]^2 dx$, resulting in the penalized least-squares criterion

$$Q(f,\lambda) = \|y - X\beta\|^2 + \lambda \int \left[f''(x)\right]^2 dx$$

The penalty contains integrated squared second derivative of $\hat{f}(x)$. If the regression fit produces estimates $\hat{f}(x)$ that are too rough, this will increase $\int [f''(x)]^2 dx$. The parameter λ in the penalty term controls the trade-off between the model fit $\|y - X\beta\|$ and the wiggliness penalty $\int [f''(x)]^2 dx$. When $\lambda = 0$, the wiggliness penalty has no effect on the minimization criterion

$Q(f,\lambda)$, resulting in unpenalized estimates of $f(x)$ that just interpolate data. In contrast, $\lambda = +\infty$ produces the perfectly smooth line, *i.e.*, a linear regression line with constant slope.

The calculation of the wiggliness penalty is simplified by noting that derivatives and integrals of $f(x)$ are linear transformations of estimated parameters $\beta_k^*(x)$, with $f''(x) = \sum_{k=1}^K \beta_k^* b_k''(x_i)$ and $\int f(x) = \sum_{k=1}^K \beta_k^* \int b_k(x_i)dx$. Thus, $f''(x) = \sum_{k=1}^K \beta_k^* b_k''(x_i) = b''(x)^T \beta$, from which follows that $[f''(x)]^2 = \beta^T b''(x)^T b''(z)\ \beta = \beta^T F(x)\ \beta$. Finally, $J = \int f''(x) = \beta^T \int F(x)dx\ \beta = \beta^T S\beta$. Thus, the wiggliness penalty J is a quadratic form in the parameter vector β and matrix S of known coefficients, derived from the basis function $b_k(x)$. The objective function becomes $\|y - X\beta\|^2 + \lambda\beta^T S\beta$. Differentiating the objective function with respect to β and setting the derivative to zero produces the estimate of regression parameter β :

$$\hat{\beta} = \left(X^T X + \lambda S\right)^{-1} X^T y \tag{5}$$

The estimate depends on the unknown smoothing parameter λ. Let H be a hat matrix that projects the vector y into the vector of predicted values $\hat{y} = X\hat{\beta}$, where estimate of $\hat{\beta}$ is given by (5). Then the hat matrix H is $H = X\left(X^T X + \lambda S\right)^{-1} X^T$. In the MGCV algorithm, the optimal value of λ is found by minimizing the generalized cross validation (GCV) criteria $V_g(\lambda)$:

$$V_g(\lambda) = \frac{n\left\|y - X\hat{\beta}\right\|^2}{[n - tr(H)]^2} \tag{6}$$

where n is the number of observations, and $tr(H)$ is the trace of hat matrix H. Approximately, $tr(H)$ equals to the number of degrees of freedom λ, used in approximating the smooth function $f(x)$. Note that while Fu (2008) suggested to set the smoothing parameter λ to 10, the MGCV algorithm selects λ that minimizes the GCV criterion $V_g(\lambda)$.

In practice, the use of the GCV criteria results in undersmoothing (Kim and Gu, 2004), but the drawback can be fixed by multiplying $tr(H)$ in (6) by a parameter γ that increases the cost per trace of H:

$$V_g'(\lambda) = \frac{n\left\|y - X\hat{\beta}\right\|^2}{[n - \gamma tr(H)]^2} \tag{7}$$

Following Kim and Gu (2004) and Wood (2006), I set the parameter γ to 1.4, but in practice the modification has little effect on estimated saving-age profiles.

The use of spline basis functions to estimate the smooth function $f(x)$ in the basic nonparametric model (4) can be easily extended to a semiparametric model that, apart from the non-parametric part $f(x)$, also has a parametric part Z (for instance, in the smoothing cohort model, matrix Z includes the combination of matrixes $[D_a, D_p]$). In the semiparametric case, the truncated cubic basis still has the form $y = \tilde{X}\tilde{\beta} + \varepsilon$, where $\tilde{X}$ is an expanded design matrix $\tilde{X} = [X, Z]$. The estimate of $\tilde{\beta}$ is obtained from (5), where the smoothing parameter λ is found by minimizing either $V_g(\lambda)$ or $V_g'(\lambda)$.

I applied the MGCV algorithm by using R software (R Development Core Team, 2009) with *mgcv* library (Wood, 2009). The MGCV algorithm allows various additions to the basic model (3).

In this paper, I report results for the basic model (I will refer to it as Model 1), and the following three extensions to it.

It is known that changes in demographic structure of households may have large impact on estimated age-saving profiles (Paxson, 1996). Therefore, I extended the Model 1 with a demographic variable q that is the number of children per household. This extension yields Model 2 as follows:

$$y = \beta_0 + \sum_{a=1}^{A} \beta_a D_a + f(c) + \sum_{p=1}^{P} \beta_p D_p + \beta_q q + \varepsilon \tag{8}$$

The impact on number of children on the household saving rate is likely to be negative. It is possible that a kind of 'economy of scale' exists for the increased number of children q, with the scale of the negative impact getting progressively smaller. To account for the possible nonlinearity, I modified Model 2 by introducing nonparametric term $f(q)$, similarly to the smooth cohort effect $f(c)$, which produced Model 3 as follows:

$$y = \beta_0 + \sum_{a=1}^{A} \beta_a D_a + f(c) + \sum_{p=1}^{P} \beta_p D_p + f(q) + \varepsilon \tag{9}$$

In Model 3, the two smooth nonparametric terms $f(c)$ and $f(q)$ have additive affect on the saving rate y. The final model introduces the joint effect between these two nonparametric terms, since it is possible that the effect from the number of children q may be conditional on the birth cohort c. For example, household cohorts that were born more recently may have a larger negative effect from q on the saving rate. To account for this joint effect, I estimated Model 4 with a joint term $f(c,q)$:

$$y = \beta_0 + \sum_{a=1}^{A} \beta_a D_a + f(c) + \sum_{p=1}^{P} \beta_p D_p + f(q) + f(c,q) + \varepsilon \tag{10}$$

4 Data

4.1 Construction of pseudo-panel dataset

To study the saving behavior of households, I use time series of cross-sectional household surveys in six countries: the United States, the United Kingdom, Italy, Japan, Taiwan, and Thailand. In every country, the composition of households changes between successive surveys, making it impossible to trace individual households over time. Instead of individual households, I analyzed the saving behavior of household groups (or 'cohorts') that were born in the same year. The idea to construct 'pseudo-panels' of different birth cohorts goes back to Deaton (1985), and has become a standard approach in estimating life-cycle models of savings. While panel data trace the same individual (or household) over time, the pseudo-panel approach traces groups of individuals who share a common trait (such as the same year of birth). These cohorts are analyzed as they are aging over time. Using this approach, cohort cells can be calculated by averaging data across households for specific age a and time t. Alternatively, cohort cells can be obtained from medians of households for specific age a and time t.

4.2 Definitions of common variables

For all countries, saving rate was defined as savings divided by non-durable consumption. Saving was measured by the residual method, as disposable income less nondurable consumption. The measure includes only discretionary savings, and omits mandatory savings to various pension plans. Disposable income was current income less direct taxes and social security contributions. Nondurable consumption was the total consumption expenditures on goods and services less expenditures on durables. Durable consumption included housing, vehicles, furniture, and household equipment, but in some countries the information for some of these categories was missing.

In general, pseudo-panel datasets were constructed as follows. Let A and T be the number of ages and cross-sectional surveys, respectively. First, saving rates for individual households were calculated. Second, these individual saving rates were used to create $A \times T$ cohort cells. Though one can use means to calculate cohort cells, I opted to use medians, because they have high robustness to outlying observations (and household data usually contain plenty of outliers). So in practice each cohort cell contained the median saving rate for specific age and year. The medians of demographic variable q was similarly calculated for different cohort cells.

Details of constructing pseudo-panels for specific countries are discussed below.

4.3 United States

Household data for the US households were taken from the Consumer Expenditure Survey (CEX) from 1984 to 2003 that is collected by the Bureau of Labor Statistics (BLS). The survey is a rotating panel that collects data during 5 quarters. Each quarter, 20 percent of households are replaced by new households. The first interview collects only basic household characteristics, while income and consumption data are collected during the following four interviews. The survey data may become incomplete in two respects. First, some households do not report complete income information about income sources (in fact, many of them report no information about their income). Second, many households do not participate in all interviews. These two groups of households represent around half of all households, and this creates a serious attrition problem. However, the BLS provides adjusted weights that take into account the attrition problem.

The CEX data was downloaded from the National Bureau of Economic Analysis (NBER) homepage (`http://www.nber.org/data/ces_cbo.html`). The full dataset contains CEX data from 1980 to 2003, but I did not use cross-sections for 1980-1983, because of low data quality in 1980-1981 surveys, and the omission of non-urban households in 1982-1983 surveys (Attansio and Paiella, 2001).

Income and consumption was calculated by following the documentation of the CEX dataset. Total income included cash income, net cash transfers and other money received. Disposable income was total income minus personal taxes and social insurance contributions. Consumption included all expenditures on goods and services, less the following durables: rent, furniture, household equipment, and personal transportation equipment.

Typically, CEX surveys around 5000 households. I dropped households that did participate in all interviews and who did not provide complete information about income sources. These selection criteria reduced the sample size by around half. In addition, I omitted student households, and households who reported negative disposable income or nondurable consumption.

4.4 United Kingdom

Data for the United Kingdom were obtained from the Family Expenditure Survey (FES) from 1975 to 2003. The data were downloaded from the homepage of Central Statistical Office at the UKDA data archive (`http://www.data-archive.ac.uk/findingData/fesTitles.asp`). The FES collects income and consumption for around 7000 households. Disposable income was measured as 'normal gross income, excluding tax and national insurance contributions, but including income in-kind'. Consumption was defined as all expenditures on goods and services, less durables. In practice, the durable consumption in the U.K. included only housing expenditures. Similarly to the US data, I omitted households who reported negative disposable income or nondurable consumption.

4.5 Italy

Household data for Italy was taken from various round of Survey of Household Income and Wealth (SHIW). The SHIW data was downloaded from the homepage of the National Bank of Italy (`http://www.bancaditalia.it/statistiche/indcamp/bilfait`). The survey collects data for various social and demographic characteristics of around 8000 households, including their consumption, income, and wealth. I used 10 SHIW cross-sections for 1987, 1989, 1991, 1993, 1995, 1998, 2000, 2002, 2004, and 2006. The definition of income included wages, property income, net transfers, and fringe benefits. Consumption was measured by total expenditures on goods and services, less durable consumption. Durables included housing, personal transport equipment, and furniture.

4.6 Japan

Data for Japanese households were taken from the National Survey of Family Income and Expenditure (NSFIE) for 1989, 1994, 1999, and 2004. The access to the micro data was arranged by the Research Centre for Information and Statistics of Social Science, Institute of Economic Research, Hitotsubashi University.

The survey collects data from more than 50,000 households, and includes information on various household characteristics, such as income, consumption, financial assets and liabilities. One limitation of the survey is that it collects household data only for a three-month period, from September to November. To convert the NSFIE data to the full year period, I followed Kitamura et al. (2003), and calculated seasonal adjustment coefficients by comparing income and consumption categories in NSFIE to the same categories from another household survey in Japan, the Family Income and Expenditure Survey (FIES). This survey collects data for the whole year, but covers only worker households, while the NSFIE includes also non-worker households. In applying the adjustment coefficients, I assumed that they are the same for worker and non-worker households.

In practice, the seasonal adjustment proceeded as follows. For consumption expenditures, I calculated adjustment coefficients for major 10 consumption categories, and then summed them up to obtain the seasonally-adjusted total consumption. Non-durable consumption was calculated as the total consumption less consumption of durables. Durable consumption included expenditures on housing (including imputed rent from owner-occupied housing), furniture, and personal transportation equipment. Categories of durable consumption were seasonally adjusted by comparing them with the same expenditure category in the FIES. The seasonal adjustment was not possible for imputed rent from owner-occupied housing, since the FIES does not estimate this expenditure category.

Income was disposable income, equal to the difference between gross income and non-living expenditures (essentially, taxes and social security contributions). Gross income included wages and salaries, income from assets (such as dividend income, and the rent from owner-occupied housings), social security benefits, and private donations. Transfer expenditures were deducted from the total income. Whenever possible, I applied seasonal adjustment to income categories by comparing them to the same income categories in the FIES. The adjustment was not possible for non-living expenditures of non-worker households. Similarly to Mason et al. (2004), I assumed that the tax rate of non-worker households is 80% of the tax rate of worker households.

Total consumption expenditure was calculated as $\overline{C^h} = \sum_{i=1}^{10} \alpha_{C,i} C_i^h + IR^h$, where $\overline{C^h}$ is total, seasonally-adjusted consumption expenditures of household h, C_i^h is unadjusted household expenditure in the NSFIE on a major consumption category, $\alpha_{C,i}$ is the adjustment coefficient for the consumption category, defined as the ratio of expenditures on the ith category in the FIES and NSFIE, and IR^h is imputed rent of household h.

Nondurable consumption was calculated as $\overline{CN^h} = \overline{C^h} - \left(\sum_{i=1}^{3} \alpha_{CD,i} CD_i^h + IR^h\right)$, where $\overline{CN^h}$ is the total nondurable consumption of household h, CD_i^h are three categories of durable consumption (namely, housing, furniture, and personal transportation equipment), $\alpha_{CD,i}$ is corresponding seasonal adjustment factor, derived as the ratio of average expenditures on CDi in the FIES and the NSFIE.

Disposable income for worker household was calculated as $\overline{YD^w} = Y/12 - \left(\alpha_{NL} Y_{NL} + \alpha_{TR} TR^h\right) + IR^h$, where $\overline{YD^w}$ is seasonally-adjusted disposable income of worker household h, Y^h is annual gross income, Y_{NL} is non-living expenditures, while α_{NL} is seasonal coefficient for Y_{NL}, and TR^h is transfer expenditures.

Disposable income for non-worker household was calculated as $\overline{YD^{nw}} = [1 - 0.8\tau^w] Y/12 - \alpha_{TR} TR^h + IR^h$, where YD^{nw} is seasonally-adjusted disposable income of non-worker household h, and τ^w is the average tax rate for worker households.

Saving rates SR_w^h and SR_{nw}^h of worker and nonworker households were defined as $SR_w^h = \left(\overline{YD_w^h} - \overline{CN^h}\right) \Big/ \overline{YD_w^h}$ and $SR_{nw}^h = \left(\overline{YD_{nw}^h} - \overline{CN^h}\right) \Big/ \overline{YD_{nw}^h}$.

4.7 Taiwan

Household data for Taiwan were taken from the 'Survey of Personal Income Distribution'. The survey is conducted annually, and I analyzed household surveys from 1978 to 2004. The survey typically covered around 9000 households. Disposable income was calculated as gross income minus personal taxes and social security contributions. Consumption was total consumption expenditures less three categories of durables: housing, furniture, and personal transportation equipment.

4.8 Thailand

Household data for Thailand were taken from Socio-Economic Survey of Thailand between 1986 and 2004. The survey was conducted in irregular intervals, every two years between 1986 and 1998, then annually between 1999 and 2002, and then resumed in the two-year interval starting from 2002. Earlier surveys included around 12,000 households, but their number increased substantially in recent years, and reached more than 35,000 households in 2004. Income was calculated as gross income less taxes and social security contributions. Because a large number of households were from rural areas, where many of them were growing their own food, the definition of consumption was relatively wider than in other countries. Specifically, consumption included not only purchased items, but also items produced at home. Consumption excluded the following categories of durable goods: housing, household equipment, vehicles, and recreation equipment.

The Thai data applies an unusual definition of household head. While household surveys in other contrives define the household head as either the primary earner, or the person who rents ow owns the housing, the Thai survey uses non-economic definition, as 'the person recognized as such by other members, whether he or she was responsible for financial support or welfare of the household members or not' (National Statistical Office of Thailand, 2003, p. A2). Since a large number of Thai households consist of three-generation households, a disproportionately large number of household heads include the elderly. However, the survey data reports income data for each household member, and I used this information to apply the economic definition of household head, using the age of household members with largest income.

5 Results

Figure 1 compares estimated age-saving profiles by Model 1 and 2 (specified by equations (3) and (8)). The models differ only in the addition of number of children per household in Model 2. The sum of age effects are constrained by restriction (2) to zero, implying that an estimate for a specific age shows deviation from the average level of savings over the estimated age span (namely, between ages 25 and 70). The deviation is measured in percentage points.

As shown in Figure 1, the addition of average number of children did not change substantially saving-age profiles in Italy, Japan, Taiwan, and Thailand. However, the addition of demographic variable shifted saving profile for U.S. households upward up to the age of 50, while for the U.K. households, the saving profile in Model 2 became downward-sloping.

Overall, saving-age profiles in Figure 1 did not show a close resemblance to the humped saving profile, but a few notable patterns are noteworthy. Households in the United States and Japan reduced their saving rate in old age by about 20 percentage points, while for Italian and Taiwanese households, the drop in saving rate was around 10 percentage points. Saving profile in the U.K. differed from the life-cycle theory the most, with decline in saving rate beginning much too early, and with a conspicuous jump in saving rate for the eldest households. The odd finding for the U.K. households has been previously reported by Paxson (1996, Figure 14), and may indicate a particularly severe selection bias among the eldest households in the U.K.

A particularly curious finding in Figure 1 is that in some countries the saving rate decreased not only in the old age, but also in the middle age, when the head of household was in mid-40s. The decline of savings in the mid-age is especially evident in Asian countries.

Figure 2 reports estimates of cohort effects, derived from Model 2. Overall, changes in saving rate due to different birth cohorts is typically less than 10 percentage points. The variation is much smaller compared with changes due to age effects in Figure 1. The largest change occurred among households in the United States, with relatively low savings for households that were born between 1920s and 1940s (a similar pattern was also reported by Attanasio (1998)).

Figure 3 shows estimates of year effects from Model 2. The figure also provides estimates of year effects that were obtained by using the Deaton-Paxson solution to the identification problem. Recall that Deaton and Paxson suggested to make year effects orthogonal to a linear trend, essentially removing time trends from estimated year effects. The impact of this restriction is most evident in estimated year effect in Thailand. With the smoothing cohort model, the year effect contains a significant upward trend (shown by thin line). In contrast, the Deaton-Paxson approach removes this upward trend, by tilting down estimates of year effect (shown by thick line). Similar clockwise rotations in estimates of year effects are evident for the United States, Japan, and Taiwan, while estimates for Italy showed the counter-clockwise rotation. The estimates of year effects were similar only in the United Kingdom.

Consequences of removing time trend from year effects in the Deaton-Paxson approach are illustrated in Figure 4. As previously discussed, the Deaton-Paxson solution to the identification problem prevents linear trends from appearing in year effect, but if any trend is present in the data, it appears in unrestricted estimates of age and cohort effects. For example, any positive trends in the data will rotate counter-clockwise both age and cohort effects. Figure 4 demonstrates that such rotations in age effect can be quite large, with substantial distortions in original pattern of saving-age profiles (particularly for aged households). For instance, the smoothing cohort model produced significant upward trends in year effects for Thailand, and, to a lesser degree, in the United States, Japan, and Taiwan (Figure 3). In each of these countries, the age profile of savings with the Deaton-Paxson approach rotated counter-clockwise. While the smoothing cohort model found decreasing savings in the old age in these countries (shown in Figure 1), the pattern became much less evident with the Deaton-Paxson approach, with the most striking contrast between two approaches for the United States, Japan, and Thailand.

Results of applying the smoothing cohort model to Model 3 are reported in Figure 5. Model 3 relaxes the linearity assumption for the impact from the number of children. However, age profiles of savings turned out identical for Models 2 and 3 in the United States, the United Kingdom, Italy, and Japan. Figure 6 provides the background for this result. In the MGCV algorithm, the degree of smoothness is determined by the number of degrees of freedom λ for which the modified GCV criterion $V_g'(\lambda)$ is minimized. In countries that produced identical saving-age profiles in Model 2 and 3, $V_g'(\lambda)$ was the smallest when the number of degrees of freedom was one, which corresponds to the linear effect from the number of children. In other words, as a result of searching for the smallest GCV statistics, Model 3 with nonlinear effects from children was reduced to Model 2, where the effect was linear. Nonlinear effect from demographics was evident only in Taiwan and Thailand. In sum, while the increase in number of children had negative effect on the saving rate, for most countries the effect can be represented by a linear function.

Figure 7 demonstrates consequences of allowing the joint impact of nonlinear demographic and cohort effects on saving rates, represented by term $f(c,q)$ in equation (10). For comparison, the figure also contains saving-age estimates from Model 3, where demographics and birth cohorts have only additive effects on saving rates. Estimates for the United States show the most notable change. Age effects for young households became flat in Model 4, and then showed a larger drop for elder households. In Taiwan, age-saving profile preserved a double-hump profile, but its trough in the 40s shifted by around 6 years to older households. On the other hand, age profiles from Models 3 and 4 were almost identical in Thailand, indicating that the impact of demographic and cohort effects can be modeled in an additive way for this country.

The interaction between the number of children and birth cohorts in Model 4 is illustrated in Figure 8. The simplest interaction is shown in Thailand, where the pattern of decreasing saving rates with more children remained similar for different birth cohorts. For other countries, the interaction pattern was more complicated. Taiwan and the United States demonstrate the contrasting patterns. In Taiwan, more recently-born households had a significant drop in saving rates with increased number of children. For older cohorts, the drop is much smaller. On the other hand, older households in the United States had lower saving rates when the number of children was large. For more recently-born households, the decline in saving rates with large number of children turned out much smaller.

Out of four considered models of household savings, which one can be considered preferable? Table 1 compares the models by generalized cross-validation criterion $V_g'(\lambda)$, used in selecting the degree of smoothness. In four countries the criterion was smallest for Model 4, and in one countries the criterion was smallest for Model 2 (USA) and Model 3 (Thailand). Note in the United States, the criterion was the same for Models 2 and 3. This happened because the search of optimal degree of smoothness in Model 3 by the MGCV algorithm selected a model with linear impact from the number of children, which is exactly Model 2.

6 Conclusions

This paper reports two major findings. First, the use of the smoothing cohort model to solve the identification problem produced more favorable evidence for the life-cycle model compared with previous studies of household savings. In particular, the Deaton-Paxson approach to solve the identification problem resulted in spurious trends in age-saving profiles of countries with rapidly changing saving rates over time (such as Thailand). Second, the paper found that for some countries the saving rate was decreasing not only in the old age, bit also in the middle age. This 'double-hump' in age-saving profiles was particularly pronounced in Japan, Taiwan, Thailand, and, to a less degree, Italy. At the current stage, we can only speculate why life-cycle savings go through two stages. The first hump in the age-saving profiles may indicate savings to take care of the growing-up children, particularly, the need for parents to finance their children's education in countries where educational loans are difficult to obtain. Savings for retirement are postponed until children become grown-up, and leave their households. Only at this point the retirement savings become the major motive for savings, and their accumulation is reflected in the second hump in age-saving profiles.

These results have several implications for thinking about the life-cycle in savings. First, the focus chiefly on the retirement motive in savings appears to be too narrow, and may miss important factors of savings, especially for households who are bringing up their children (particularly households that have to shoulder costs of their children's education). Second, the 'M-shape' in saving profiles implies that households go through two stages of savings and dissavings, with a particularly heavy financial burden for households in their 40s. Finally, the shortfall of positive savings in the 40s implies much smaller impact of income growth on savings compared with the conventional 'stripped-down' theory of life-cycle savings. In particular, the double hump may weaken the impact of population growth on savings, and on balance may even decrease them due to the depressed savings of households who bring up their children.

References

Attanasio, O. P., 1998. Cohort analysis of saving behavior by US households. Journal of Human Resources 33 (3), 575–609.

Attansio, O. P., Paiella, M., 2001. Household savings in the USA. Research-in-Economics 55 (1), 109–132.

Borsch-Supan, A., 2003. Life-cycle savings and public policy : a cross-national study of six countries. Academic Press, Amsterdam, Boston.

Deaton, A. S., 1985. Panel data from time-series of cross-sections. Journal of Econometrics 30, 109–126.

Deaton, A. S., Paxson, C., 1994. Saving, growth, and aging in Taiwan. In: Wise, D. (Ed.), Studies in the Economics of Aging. University of Chicago Press, Chicago, pp. 331–357.

Fu, W. J., 2008. A smoothing cohort model in age–period–cohort analysis with applications to homicide arrest rates and lung cancer mortality rates. Sociological Methods and Research 36 (3), 327–361.

Hastie, T. J., Tibshirani, R. J., 1990. Generalized Additive Models. Chapman and Hall–CRC, London.

Kim, Y.-J., Gu, C., 2004. Smoothing spline gaussian regression: more scalable computation via efficient approximation. Journal of Royal Statistical Society (Series B) 66, 337–356.

Kitamura, Y., Takayama, N., Arita, F., 2003. Household savings and wealth distribution in japan. In: Borsch-Supan, A. (Ed.), Life-cycle savings and public policy : a cross-national study of six countries. Academic Press, Amsterdam, Boston, pp. 148–203.

Mason, A., Ogawa, N., Fukui, T., 2004. Aging, Family Support Systems, Savings and Wealth: Is Decline on the Horizon for Japan? Working paper, University of Hawaii, `http://www2.hawaii.edu/~amason/Research/Mason.Ogawa.Fukui.pdf`.

Modigliani, F., Brumberg, R. H., 1954. Utility analysis and the consumption function: an interpretation of cross-section data. In: Kurihara, K. K. (Ed.), Post-Keynesian Economics. Rutgers University Press, New Brunswick, pp. 388–436.

National Statistical Office of Thailand, 2003. Report of the 2002 Household Socio-Economic Survey. National Statistical Office, Bangkok.

Paxson, C., 1996. Saving and growth: Evidence from micro data. European Economic Review 40, 255–288.

Poterba, J. (Ed.), 1994. International Comparison of Household Savings. Chicago University Press, Chigago.

R Development Core Team, 2009. R 2.8.1: a Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria, `http://www.R-project.org`.

Schimek, M. G., forthcoming. Semiparametric penalized generalized additive models for envoronmental research and epidemiology. Envinmetrics, accepted for publication.

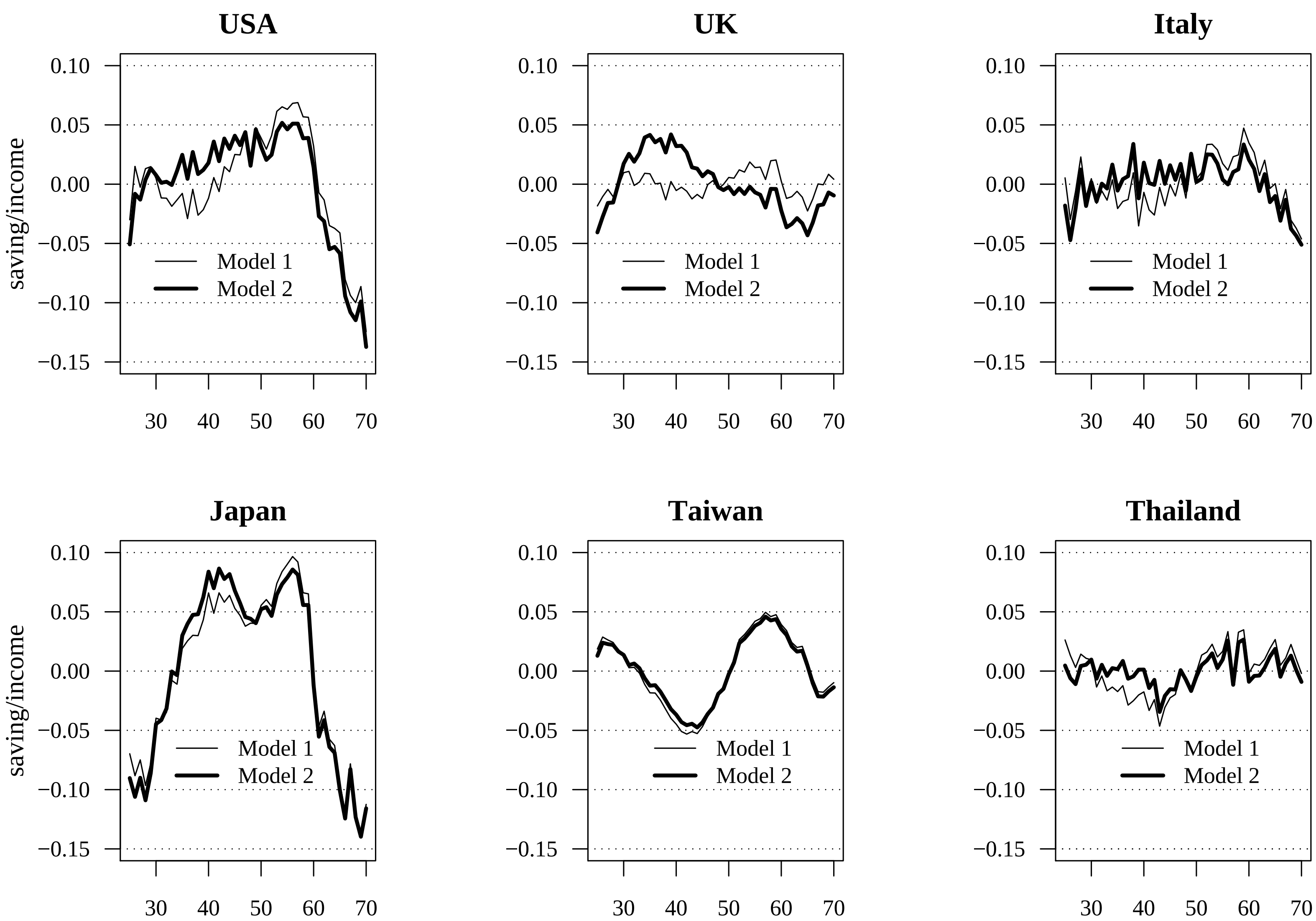
Suits, D. B., 1984. Dummy variables: mechanics versus interpretation. Review of Economics and Statistics 66, 177–180.

Wood, S., 2004. Stable and efficient multiple smoothing parameter estimation for generalized additive models. Journal of the American Statistical Association 99, 673–686.

Wood, S., 2006. Generalized Additive Models. An Introduction with R. Chapman and Hall–CRC, Boca Raton, Florida.

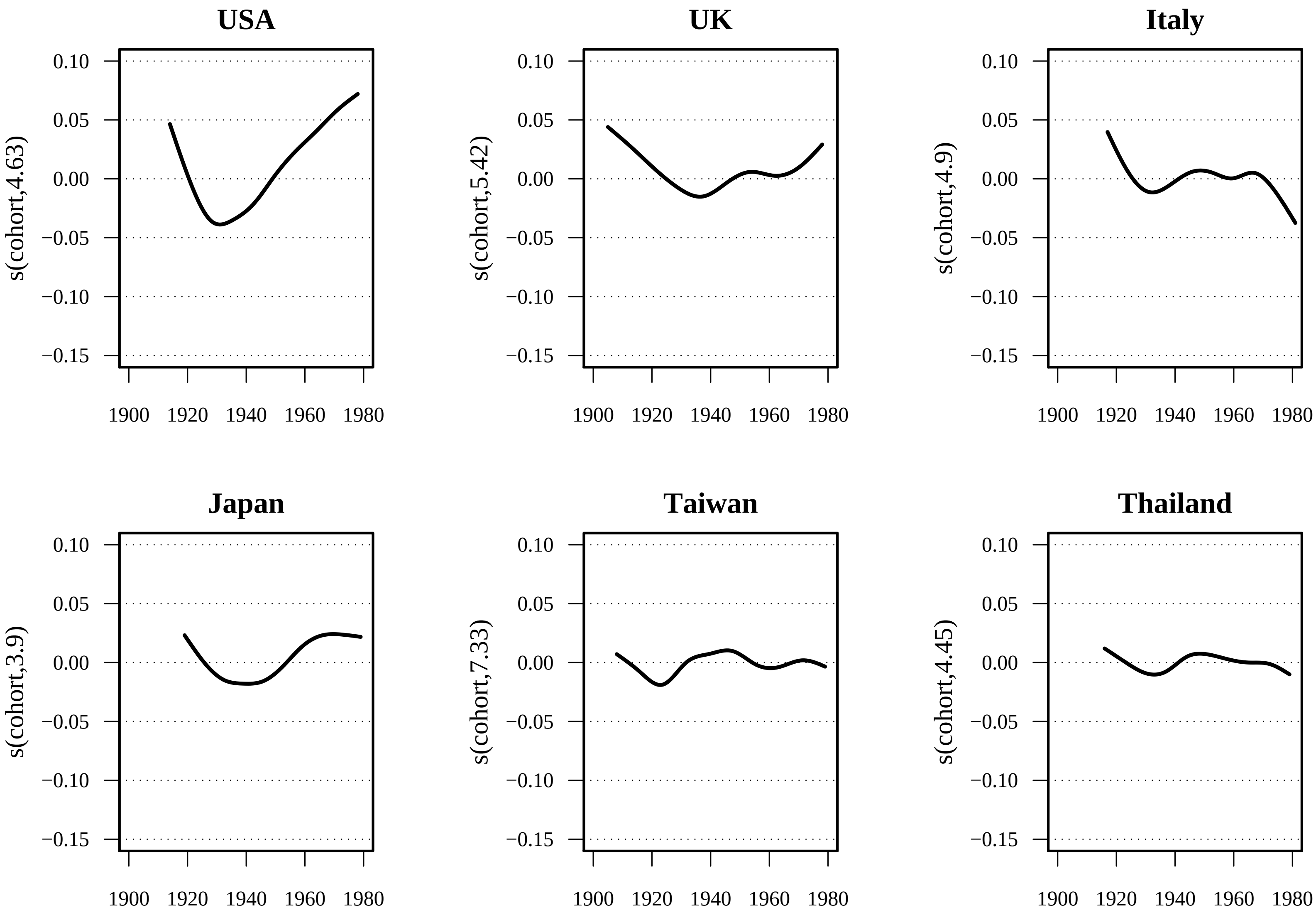
Wood, S., 2009. The mgcv Package. `http://cran.r-project.org/web/packages/mgcv/mgcv.pdf`.

Figure 1. *Age effects in saving rate (Models 1 and 2).*



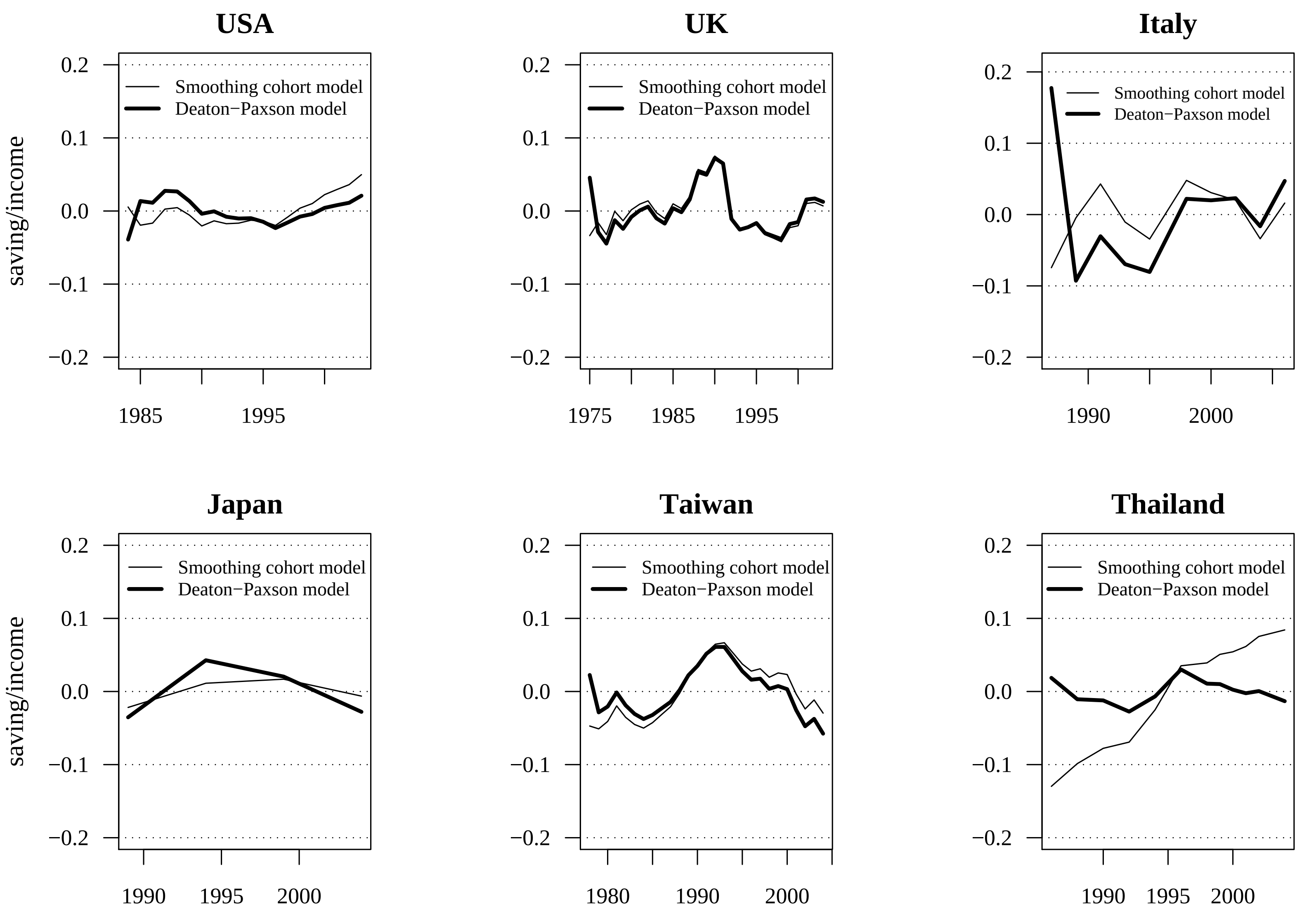
Note: Age effects in models 1 and 2 are estimated by (3) and (8).

Figure 2. *Cohort effects in saving rate (Model 2).*



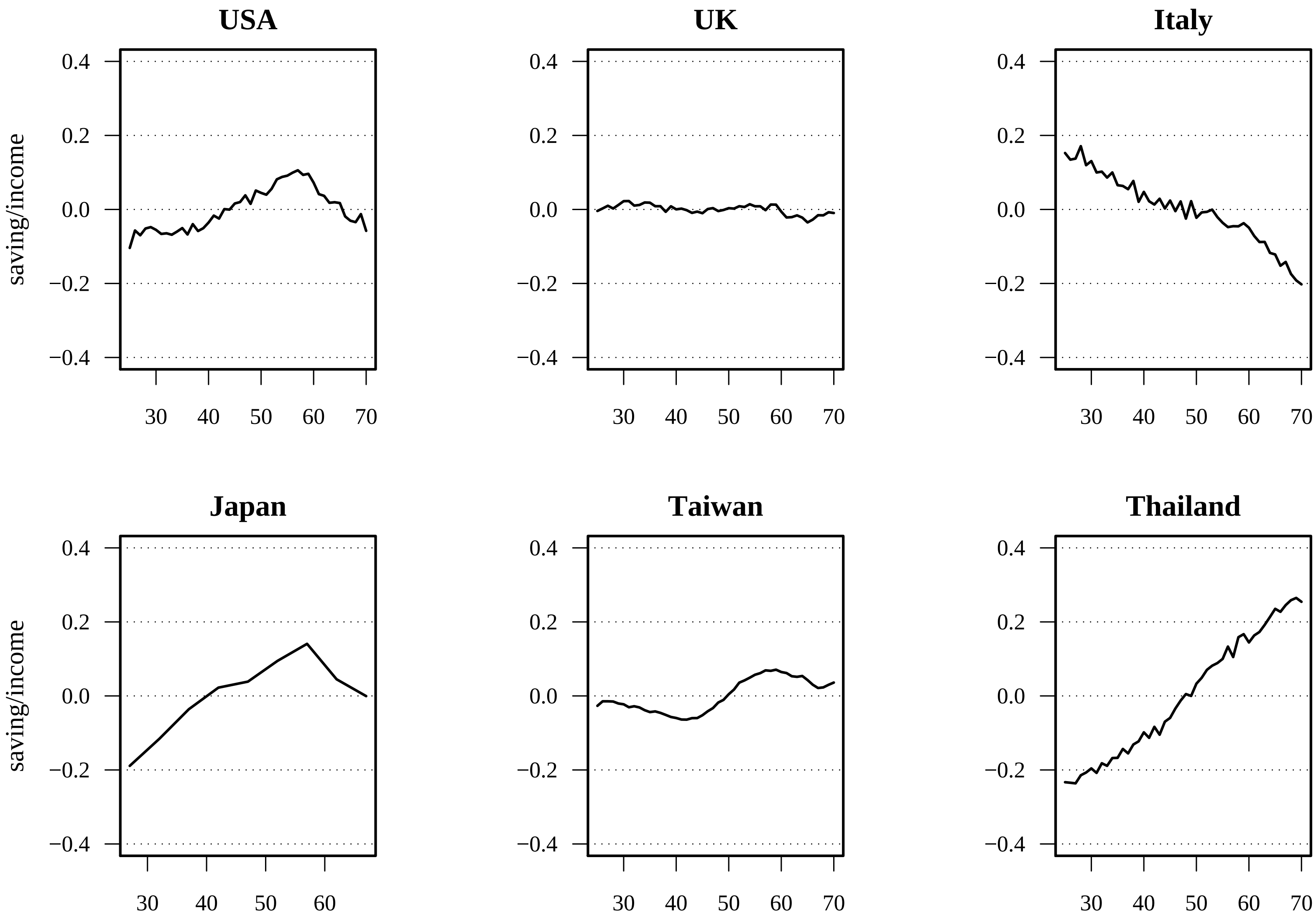
Note: Cohort effects are estimated with model 2, given by (8).

Figure 3. *Year effects in saving rate (Model 2).*



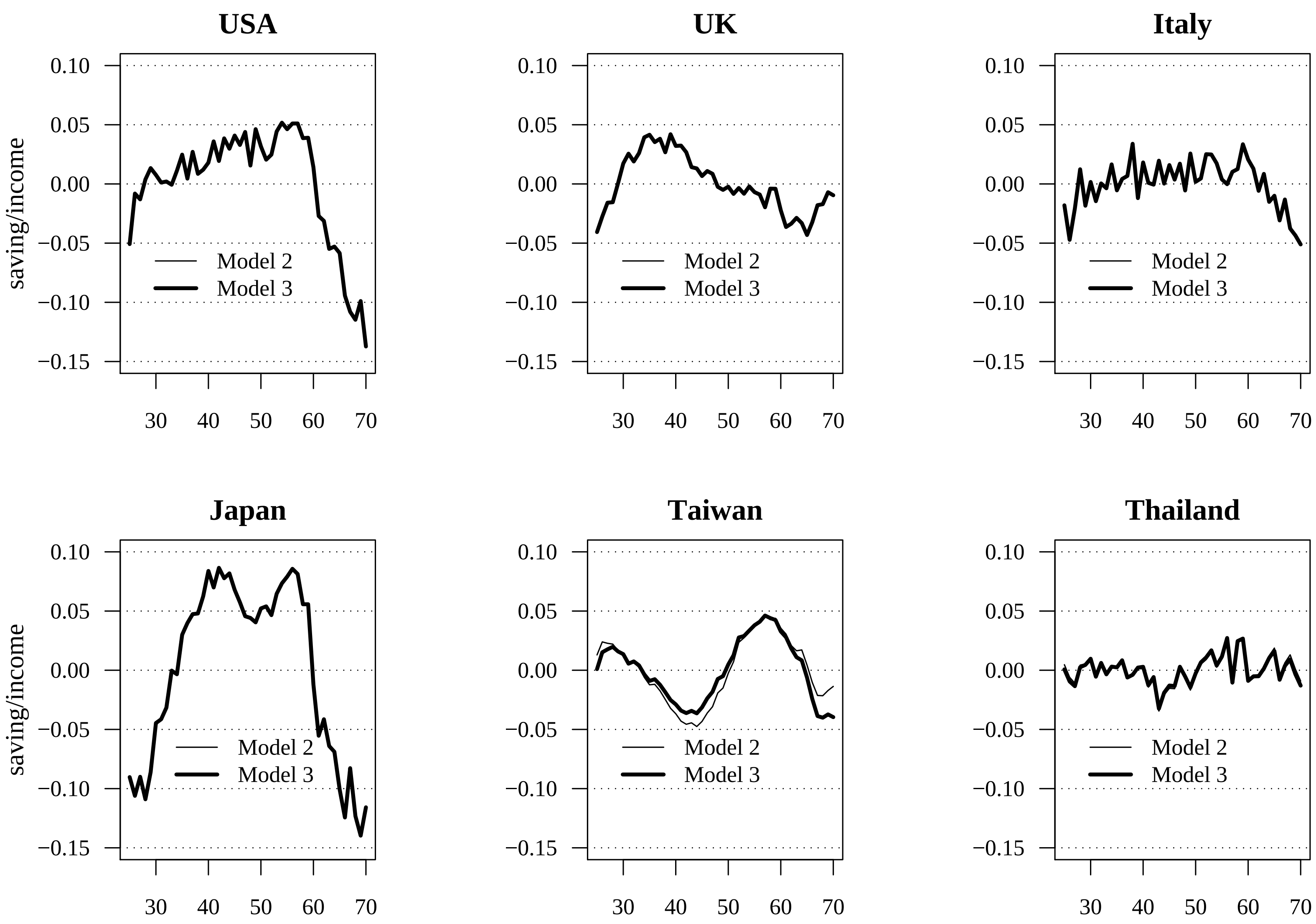
Note: Year effects are estimated with model 2, given by (8).

Figure 4. *Age effects with Deaton-Paxson approach*



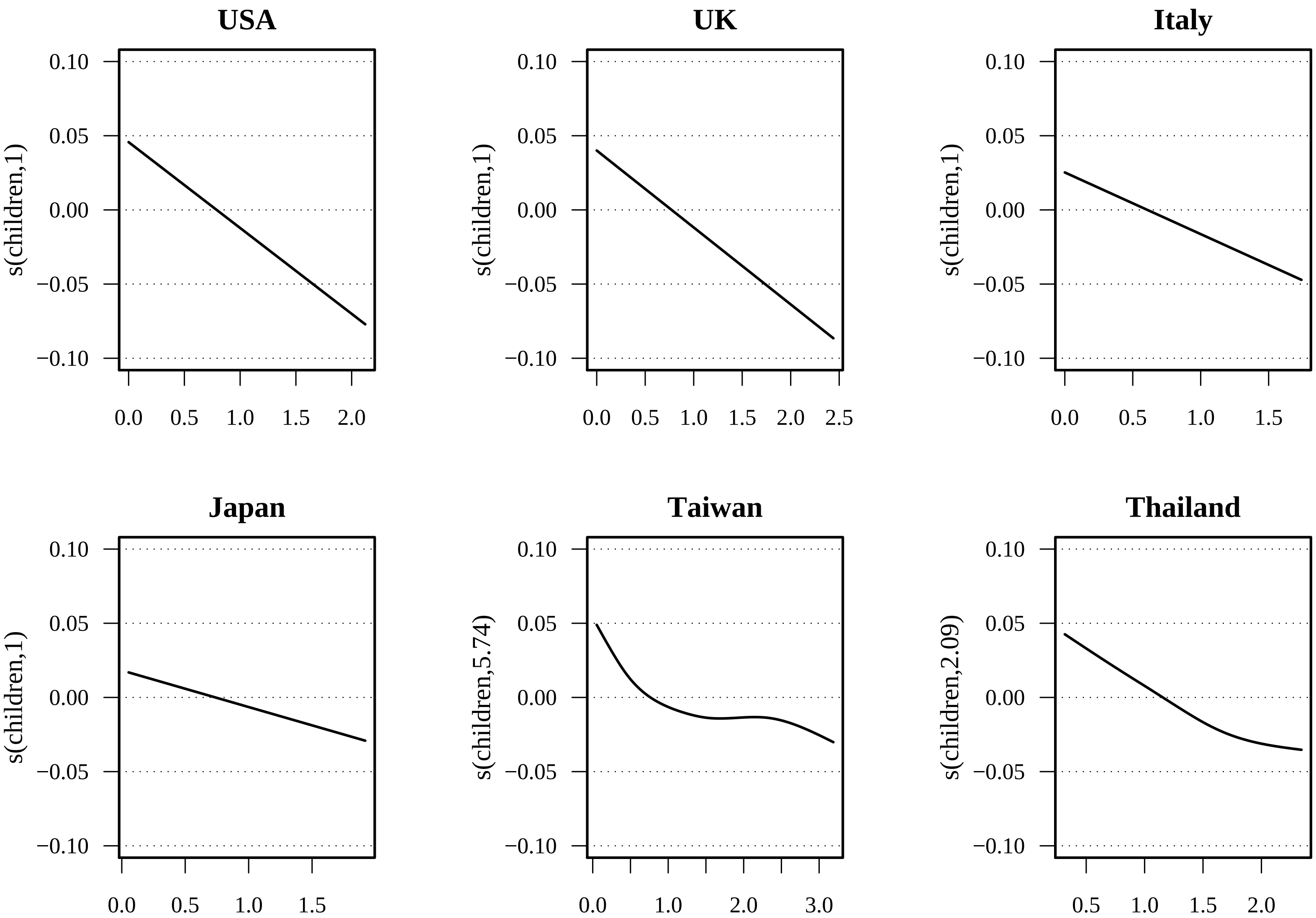
Note: Age effects were estimated with the Deaton-Paxson approach, assuming that year effect are orthogonal to a time trend.

Figure 5. *Age effects with Models 2 and 3.*



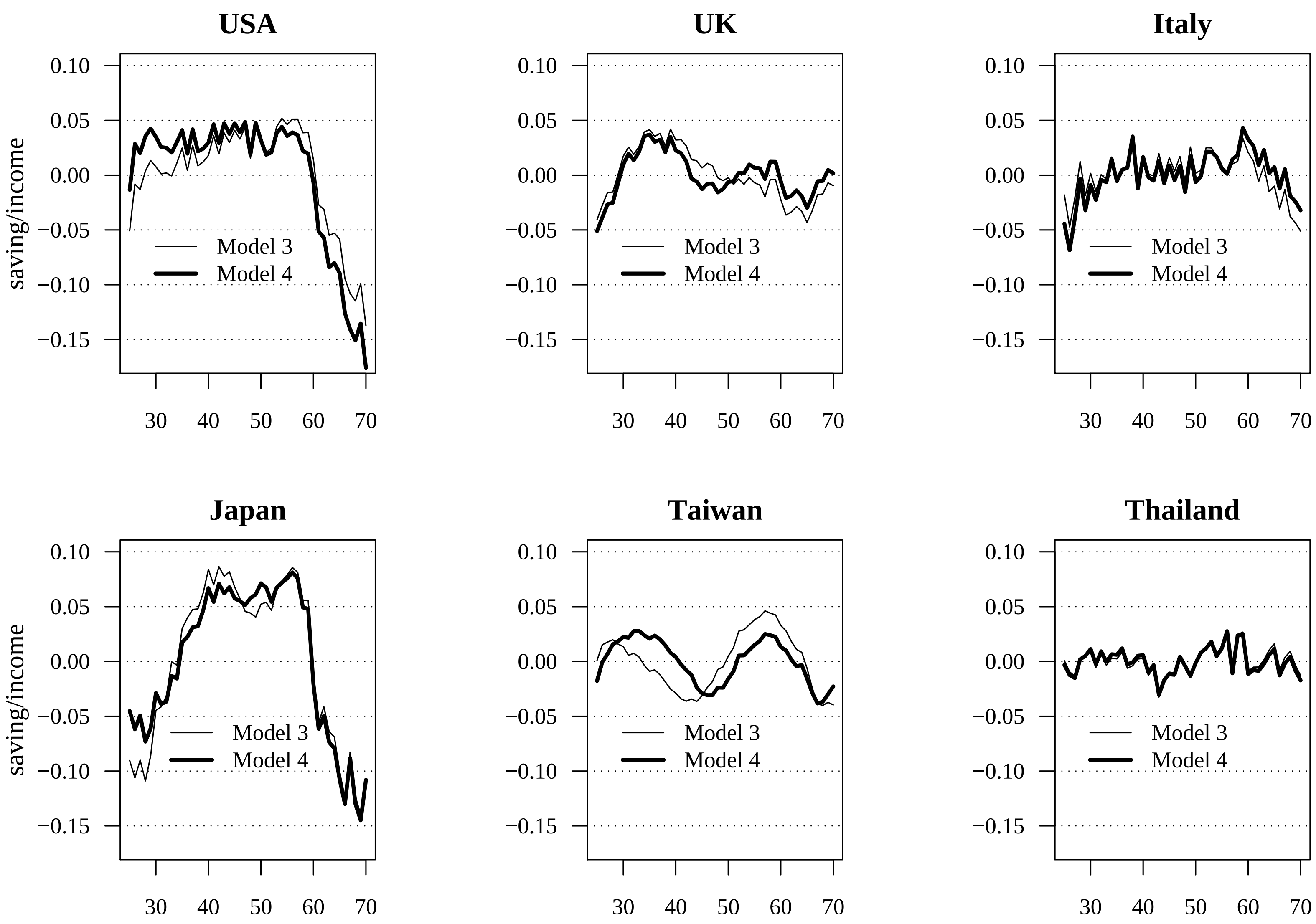
Note: Age effects in models 2 and 3 are estimated by (8) and (9).

Figure 6. *The impact of the number of children on saving rates (Model 3).*



Note: Demographic effects are from model 3, estimated by (9).

Figure 7. *Age effects in Models 3 and 4.*



Note: Age effects in models 3 and 4 are estimated by (9) and (10).

Figure 8. *Joint effect on saving rate from the number of children and birth cohorts.*

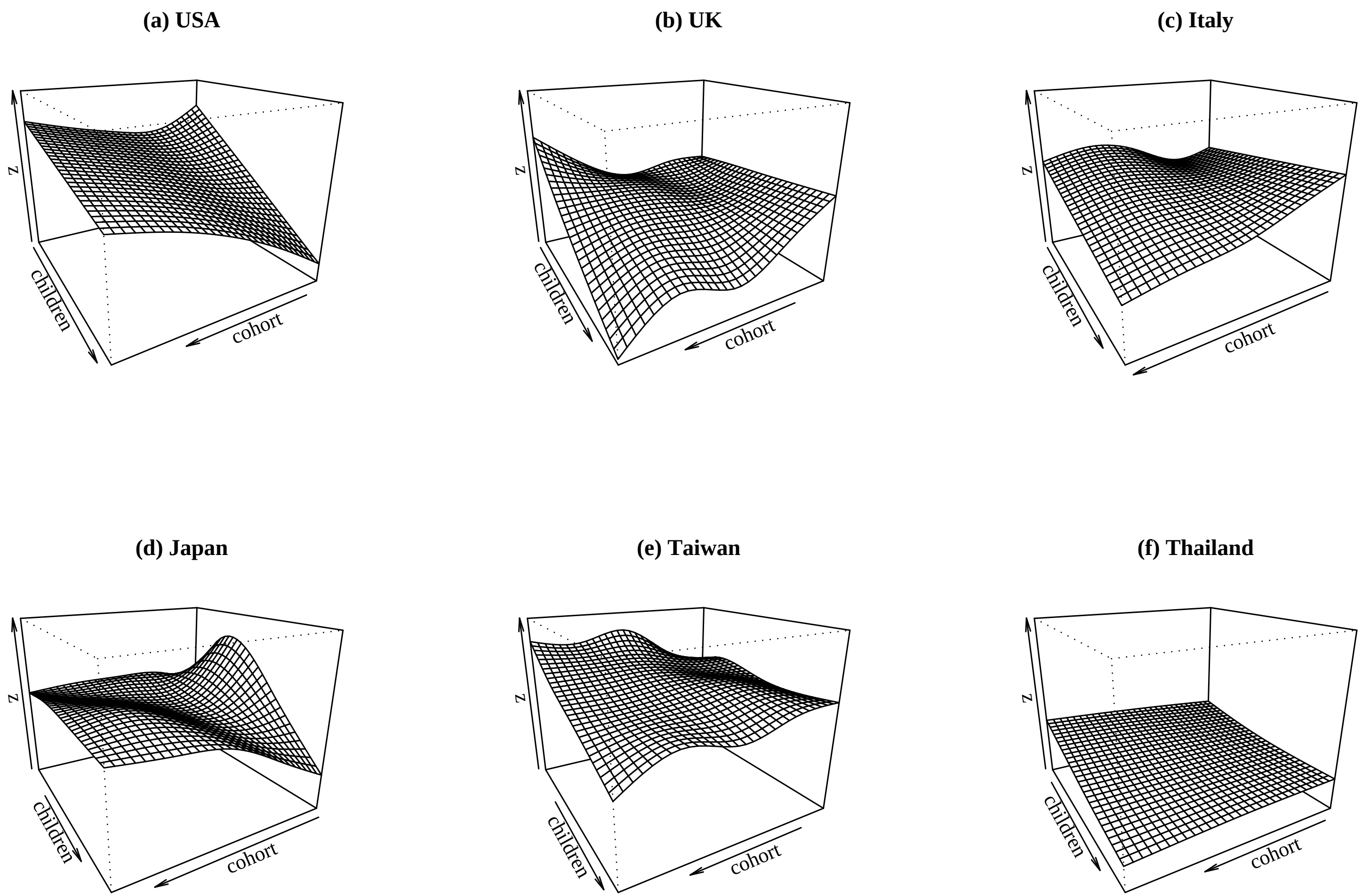


Table 1. *Comparison of modified GCV criterion $V'_g(\lambda)$ for alternative models of household saving rate*

	USA	UK	Italy	Japan	Taiwan	Thailand
Model 1	0.22448	0.17389	0.34082	0.79041	0.08879	0.46967
Model 2	**0.22025**	0.17103	0.34045	0.80104	0.08889	0.46589
Model 3	**0.22025**	0.17103	0.34045	0.80104	0.08393	**0.46457**
Model 4	0.22087	**0.17062**	**0.33790**	**0.71061**	**0.08076**	0.46769

Note: modified GCV criterion $V'_g(\lambda)$ was calculated by (7). Smallest values are shown in bold font.