Working Paper

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

Oleksandr Movshuk

Working Paper No. 258
December 28, 2010

FACULTY OF ECONOMICS
UNIVERSITY OF TOYAMA

3190 Gofuku, Toyama
930-8555 JAPAN

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

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 semiparametric 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 in 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, and to the International Centre for the Study of East Asian Development (ICSEAD) in Kitakyushu, Japan, for financial support for the Research Project on Comparative Analysis of Consumption and Saving Behavior in East Asian Countries.

1 Introduction

This paper examines the saving behavior of households in six countries, and presents favorable evidence about predictions of the life-cycle hypothesis (LCH) of Modigliani and Brumberg (1954). The LCH claims that age-saving profiles have a hump shape, with individuals saving between the middle age and retirement, and dissaving in young and old ages. Though this paper did not observe this hump-saving in every of six examined countries, the evidence of life-cycle savings agreed with the theoretical prediction of the LCH to a much greater extend, as compared with previous studies that examined age-saving profiles in household data (Poterba, 1994; Deaton and Paxson, 1994; Paxson, 1996; Borsch-Supan, 2003). In particular, while the LCH 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 followed Deaton and Paxson (1994) and studied 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 the decomposition 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 (Paxson, 1996; Borsch-Supan, 2003). In this paper, I apply an alternative solution that restricts cohort effect to be a smooth function, whose shape is estimated by a semiparametric 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 15 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. On the other hand, age profiles of savings in the United Kingdom and Thailand turned out more irregular, with 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 is explaining the saving rate y, which is 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 observed year. For example, age effects for ages between 25 and 70 are estimated with $70-25+1=46$ dummy variables for each age between 25 and 70. Let D_a be a matrix that combines these 46 age dummy variables. 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 in the following model:

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

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). By dropping first terms in D_a, D_c, and D_t, the corresponding parameter estimate is set to zero. This turns the dropped terms into a benchmark to interpret estimates of β_a, β_c, and β_t. However, this choice of a particular age (say, 25 years old) as a benchmark is not helpful in interpreting estimated parameters β_a, β_c, and β_t. A more informative benchmark is obtained 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 after imposing the restriction (2), because of the exact linear relation between observed year t, age a and year of birth b (namely, $t=a+b$). As a result of this perfect collinearity, it is not possible to get a unique explanation for examined data, especially if there are general trends in data. Suppose 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 that year effects are orthogonal to a linear time trend, and 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 precluded from appearing from year effects, and are attributed to a combination of 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. In consequence, the Deaton-Paxson approach postulates that time effects contain only cyclical variation. If any trends appear in data, they are forced to show up in age and cohort effects, since only these two effects are unrestricted. Thus, the Deaton-Paxson approach may result in spurious trends in age and cohort effect that may mask 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. In particular, estimates of year effect may contain any kind of trend.

In this solution, cohort effect is restricted to an arbitrary smooth function that is 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 the effect from c on the saving rate y is nonlinear. The introduction of a smooth nonlinear function $f(c)$ in (1) produces the following smoothing cohort model:

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

4 Estimation

4.1 Estimation of the basic model

The smoothing cohort model (3) is essentially a semiparametric regression model with a nonparametric term $f(c)$ and a parametric part that has two sets of dummy variables D_a and D_p. Originally, Fu (2008) suggested to fit the smoothing cohort model as a generalized additive model (GAM), using the backfitting algorithm of Hastie and Tibshirani (1990). However, the stability of the backfitting algorithm was questioned in recent years, particularly in datasets with high collinearity among explanatory variables (Schimek, forthcoming). Another limitation of the GAM estimator is requirement to select a smoothing parameter (namely, the number of degrees of freedom v). While Fu (2008) claimed that setting v to 10 degrees of freedom 'yields good results' (p. 341), there is no guarantee that the value of smoothing parameter will be an accurate in describing the actual shape of cohort effect. A more preferable approach is to determine the degree of smoothing of $f(c)$ in an endogenous way that depends on examined data.

The automatic selection of the smoothness criteria in the GAM model is possible with the Modified Generalized Cross Validation (*MGCV*) algorithm of Wood (2004). The MGCV has superior numerical stability compared to the backfitting algorithm, especially when explanatory variables are correlated (Schimek, forthcoming). In addition, the MGCV algorithm selects an appropriate degree of smoothness using a large variety of selection methods, including the generalized cross validation (GCV) criterion of Craven and Wahba (1979), or restricted maximum likelihood (REML) methods that represent the nonparametric part as random effects (Ruppert et al., 2003). In this section I discuss how the smoothing cohort model is estimated by the MGCV algorithm, with smoothness selected by minimizing the GCV criterion.

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

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

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. In the MGCV algorithm, the smooth function $f(x)$ is approximated by a sequence of cubic splines. In general, splines are piecewise polynomials that are joined at the 'knots'. Due to special restrictions, the cubic splines are continuous at the knots, and also have continuous first and second derivatives. Let K denote the number of knots. Then a cubic spline can be represented by truncated cubic basis functions:

$$f(x) = \beta_0 + \beta_1 x + \beta_2 x^2 + \beta_3 x^3 + \sum_{k=1}^{K} \beta_{k+3}(x - \kappa_k)_+^3 \tag{5}$$

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 of a *global* cubic polynomial $\beta_0 + \beta_1 x + \beta_2 x^2 + \beta_3 x^3$ and K *local* polynomial deviations $\sum_{k=1}^{K} (x - \kappa_k)_+^3$.

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. A common penalty is $\lambda \int [f''(x)]^2 dx$, which has a smoothing parameter λ and an integrated squared second derivative of $\hat{f}(x)$. This results in the penalized least-squares criterion $Q(f, \lambda) = \|y - X\beta\|^2 + \lambda \int [f''(x)]^2 dx$.

If $\hat{f}(x)$ is too rough, this will increase the penalty term $\int [f''(x)]^2 dx$. The smoothing parameter λ controls the trade-off between the model fit $\|y - X\beta\|$ and the roughness penalty $\int [f''(x)]^2 dx$. When $\lambda = 0$, the roughness penalty has no effect on the minimization criterion $Q(f, \lambda)$, producing unpenalized estimates $\hat{f}(x)$ that just interpolate data. In contrast, when $\lambda = +\infty$, this results in the perfectly smooth line, *i.e.*, in a linear regression line with a constant slope.

The minimization of the penalized criterion $Q(f, \lambda)$ is simplified by noting that derivatives and integrals of $f(x)$ are linear transformations of parameters $\beta_k(x)$ in the cubic spline basis, with $f''(x) = \sum_{k=1}^{K} \beta_k b_k''(x_i)$ and $\int f(x) = \sum_{k=1}^{K} \beta_k \int b_k(x_i) dx$, where $b_k(x_i)$ denotes a particular form of basis function (such as the truncated cubic basis function in (5)). Thus, $f''(x) = \sum_{k=1}^{K} \beta_k b_k''(x_i) = b''(x)^T \beta$, from which it follows that $[f''(x)]^2 = \beta^T b''(x)^T b''(z)\ \beta = \beta^T F(x)$ β. Finally, $J = \int f''(x) = \beta^T \int F(x) dx\ \beta = \beta^T S \beta$. Thus, the roughness penalty J can be represented as a quadratic form in the parameter vector β and matrix S of known coefficients that are derived from the basis function $b_k(x)$.

Substituting the roughness penalty J with $\beta^T S \beta$, the penalized least-squares criterion $Q(f, \lambda)$ becomes $\|y - X\beta\|^2 + \lambda \beta^T S \beta$. Differentiating $Q(f, \lambda)$ with respect to β and setting the derivative to zero produces an estimate of β :

$$\hat{\beta} = \left(X^T X + \lambda S\right)^{-1} X^T y \tag{6}$$

The estimate of β depends on the value of unknown smoothing parameter λ. The MGCV algorithm selects an appropriate value of λ by using the concept of hat matrix from the ordinary least-squares model. In the model, the hat matrix H projects the vector of dependent variable y into the vector of predicted values $\hat{y} = Hy$, with $H = X\left(X^T X\right)^{-1} X^T$. Using the estimate of $\hat{\beta}$ from (6), the hat matrix of the penalized spline model can be similarly defined as $H_S =$

$X\left(X^TX+\lambda S\right)^{-1}X^T$. Since the matrix H_S transforms the vector of y into the vector of its smoothed values, the matrix H_S is often called a smoother matrix. In the MGCV algorithm, the optimal value of λ is found by minimizing the GCV criteria $V_g(\lambda)$ that depends on the sum of squared residuals $\left\|y-X\hat{\beta}\right\|^2$ and the trace of smoother matrix H_S:

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

where n is the number of observations, and $tr(H_S)$ is the trace of H_S.

Though the MGCV algorithm selects an appropriate degree of smoothness with respect to parameter λ, this parameter is not informative in evaluating the estimated degree of smoothness. It is much easier to interpret the trace of the smoother matrix $tr(H_S)$, since it is equal to the number of degrees of freedom, needed to approximate the smoothed function $f(x)$ (Ruppert et al., 2003). Let $v=tr(H_S)$. Since the smoothing parameter λ is a part of H_S, λ and v are correlated. In particular, a small degree of smoothing is indicated by $\lambda\to 0$ and $v\to\infty$. Conversely, a high degree of smoothing corresponds to $\lambda\to\infty$ and $v\to 0$. An important special case is when $v\leq 1$. This range of v indicates a parametric effect, when a single variable is sufficient to approximate the smoothed function $f(x)$.

In summary, while Fu (2008) suggested to estimate the smoothing cohort model with v fixed at 10, the MGCV algorithm searches for an optimal values of smoothing parameter λ, which in practice can produce any degree of smoothing.

The GCV criterion $V_g(\lambda)$ has one problem in selecting an optimal smoothness. Monte Carlo studies by Kim and Gu (2004) and Bacchini et al. (2007) demonstrated that $V_g(\lambda)$ may choose too small values of λ, which results in undersmoothing. The problem can be solved by multiplying $tr(H_S)$ in (7) by a parameter γ that increases the cost per trace of H_S:

$$V_g'(\lambda)=\frac{n\left\|y-X\hat{\beta}\right\|^2}{[n-\gamma tr(H_S)]^2} \tag{8}$$

In estimating the smoothing cohort model, I followed the suggestion of Wood (2006) that a good value of γ is 1.4. But in practice, the modification had little effect on estimated saving-age profiles.

Once the use of spline basis functions in estimating the smooth function $f(x)$ is introduced, the basic model (4) can be easily extended to the full semiparametric model with a parametric part. In the case of smoothing cohort model, the parametric part Z includes the combination of matrixes with dummy variables $[D_a,D_p]$. Then the truncated cubic basis (5) still has the form $y=\tilde{X}\tilde{\beta}+\varepsilon$, but the basis $\tilde{X}$ includes an expanded design matrix $\tilde{X}=[X,Z]$. The estimate of $\tilde{\beta}$ is obtained from (6), 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 with *MGCV* library, version 1.3 (Wood, 2006). The MGCV algorithm allows various additions to the basic model (3). In this paper, I report results for the original specification (3). It will be called Model 1. In addition, I will consider three modifications to the basic model.

4.2 Extensions to the basic empirical model

In addition to three major effects on saving rate in Model 1, changes in household's demographic structure also may have large effect on estimated age-saving profiles (Paxson, 1996). To account for the demographic change, Model 1 was extended with a demographic variable q, which is the number of children per household. This extension produced Model 2 as follows:

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

Larger number of children will increase household consumption. Other things being equal, the increased consumption will depress the saving rate, so the impact of q on the saving rate is likely to be negative. But it is also possible that an 'economy of scale' exists for the increased number of children, with the scale of the negative impact on the saving rate getting progressively smaller for each additional child. To account for this possible nonlinearity, Model 2 was modified with a nonparametric term $f(q)$, similarly to the smooth cohort effect $f(c)$. This produced Model 3 as follows:

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

In Model 3, the two smooth nonparametric terms $f(c)$ and $f(q)$ have additive affect on the saving rate y, and there is no interaction between these two nonparametric terms. However, the effect from the number of children q may be conditional on the birth cohort c. For example, household cohorts that were born in more recent years may need to spend more on educating their children. This will result in a higher household expenditures among these birth cohorts, and consequently, a larger negative effect from q on the saving rate. To account for this kind of joint effect, Model 4 was augmented with a joint term $f(c,q)$ for cohort effect and demographics:

$$y = \beta_0 + \sum_{a=1}^{A} \beta_a D_a + f(c) + \sum_{p=1}^{P} \beta_p D_p + f(q) + f(c,q) + \varepsilon \tag{11}$$

5 Data

5.1 Construction of pseudo-panel dataset

To study the saving behavior of households, I used 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 age over time. In this approach, cohort cells can be calculated by averaging data across households for specific age a and time t. Alternatively, cohort cells can be taken from medians of households for specific age a and time t.

5.2 Definitions of common variables

For all countries, the saving rate was defined as saving 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, since household surveys rarely report such data. 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 typically included housing, vehicles, furniture, and household equipment, but in some countries the information for some of these durable categories was not available.

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. 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. Below I discuss details of constructing pseudo-panels for specific countries.

5.3 United States

Household data for the US households were taken from the Consumer Expenditure Survey (CEX) from 1984 to 2003. The survey is a rotating panel that collects data over 5 quarters.

Each quarter, 20 percent of households are replaced. The first interview collects only basic household characteristics, while income and consumption data are collected during the following four interviews. One potential problem of the CEX data is that data are not complete for many households. Two patterns of missing data are common. First, some households fail to report the 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 survey data contain adjusted weights that take into account the attrition problem.

The CEX data was downloaded from the homepage of the National Bureau of Economic Analysis (`http://www.nber.org/data/ces_cbo.html`). The full dataset contains CEX data from 1980 to 2003. 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 (Attanasio and Paiella, 2001).

Income and consumption was calculated according to 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 whose disposable income or nondurable consumption were negative.

5.4 United Kingdom

Data for the United Kingdom were obtained from the Family Expenditure Survey (FES) from 1975 to 2003. The data were obtained from the homepage of Central Statistical Office at the UK 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 minus 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.

5.5 Italy

Household data for Italy was taken from the 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 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.

5.6 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 in the Institute of Economic Research of 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 the 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, as the ratio of income and consumption categories in NSFIE to same categories from another household survey in Japan, the Family Income and Expenditure Survey (FIES). The FIES 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 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 report this expenditure category.

Income was disposable income, defined as 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 was 80% of the tax rate of worker households.

Total consumption expenditure was defined 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 by

$$\overline{YD^{nw}} = [1 - 0.8\tau^w] Y/12 - \alpha_{TR} TR^h + IR^h$$

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 $SR_w^h = \left(\overline{YD_w^h} - \overline{CN^h}\right) / \overline{YD_w^h}$ and $SR_{nw}^h = \left(\overline{YD_{nw}^h} - \overline{CN^h}\right) / \overline{YD_{nw}^h}$.

5.7 Taiwan

Household data for Taiwan were taken from the annual 'Survey of Personal Income Distribution' from 1978 to 2004. The survey included data for 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.

5.8 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 returned to the two-year interval between 2002-2004. Earlier surveys included around 12,000 households, and their number increased substantially in recent years, reaching more than 35,000 households in 2004. Income was calculated as gross income less taxes and social security contributions. Because a large number of Thai households live in rural areas, where many households grow their own food, the definition of consumption was 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 originally used an unusual definition of household head. While household surveys in other countries use economic definitions of household head (as the primary earner, or as the person who rents or 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 are the elderly. Fortunately, the survey data reports income data for each household member. With this information, it was possible to reclassify the heads of households as household members with largest income among household members.

6 Results

Figure 1 compares estimates of age-saving profiles in Models 1 and 2, as specified by equations (3) and (9). These models differ only in the addition of the number of children in Model 2, with a linear effect from the demographic variable.

To interpret the estimated saving profiles, note that the sum of estimated effects is restricted to zero by (2). For age effects, this implies that an estimate for a specific age shows a deviation from the average level of savings over the whole age span (specifically, between ages 25 and 70). The deviation is measured in percentage points.

Saving-age profiles in Figure 1 did not show a uniform pattern, but some countries demonstrate a pronounced drop in the saving rate among aged households. In particular, households in the United States and Japan reduced their saving rate in old age by about 15 percentage points, while for Italian and Taiwanese households, the drop was around 10 percentage points. The relatively sharp decline among the aged in Japan agrees with micro evidence from Horioka (2009) that Japanese households in recent years had large dissavings in old age, mainly due reduced public pensions, and increased taxes and social security payments, as well as higher consumption expenditures. As for the similarly sharp decline in the saving rate among the aged U.S. households, it turned out deeper compared with age-saving profiles in previous studies (Attanasio, 1998; Attanasio and Paiella, 2001).

Figure 1 also shows that in most countries, the addition of demographic variable in Model 2 had little effect on saving-age profiles. However, there was a shift in the saving profile of U.S. households that moved the profile upward for young households. The effect was also evident for the U.K. households, with a similar upward shift for young households. For subsequent ages, the U.K. saving profile was unusually odd. The saving rate started to decline for age in the late 30s, and then jumped conspicuously for the old age. The unusual pattern for the eldest U.K. households has been previously reported by Paxson (1996, Figure 14), and may indicate a particularly severe selection bias among the oldest households, or some other unusual features of the U.K. data. Another curious finding in Figure 1 is that the saving rate in some countries was decreasing not only in old age, but also in middle age, with the middle-age decline in savings particularly evident among Asian households. Possible causes of the double hump in savings will be further discussed in subsection 6.4.

Figure 2 reports estimates of cohort effects from Model 2. Similarly to estimates in Figure 1, the sum of estimated cohort effects is constrained to zero. Therefore, an estimate for a specific birth cohort shows a deviation from the average level of cohort effect for all birth cohorts.

Compared with estimates of age effects in Figure 1, estimates of cohort effects turned out much smaller. Typically, the spread between the largest and smallest effects was less than 10 percentage points, with the largest variation among households in Japan and the United States. Among Japanese households, the cohort effect was the smallest for households born in the 1930s, with a sharp subsequent increase up to the youngest cohorts that were born in the early 1980s. A similar pattern was evident among the American households, with the relatively low saving rate for households born between 1920s and 1940s. It is noteworthy that Attanasio (1998) also found a similar depressed level of U.S. household savings among the interwar cohorts.

6.1 Comparison with estimates based on the Deaton-Paxson approach

Figure 3 reports estimates of year effects on saving rates, and compares them with estimates from the Deaton-Paxson solution to the identification problem. In the Deaton-Paxson approach, year effects are restricted to be orthogonal to a linear trend, and this removes linear time trends from estimated year effects. Figure 3 reports year effects that are based on the smoothing cohort model and the Deaton-Paxson approach. The contrast between these two approaches is most evident for Thai households. The smoothing cohort model does not restrict the year effect, so its estimates of year effect have a significant upward trend. In contrast, the orthogonality restriction in the Deaton-Paxson approach precludes this upward trend from appearing, so estimated year effects are tilted clockwise, and become essentially flat. Similar clockwise rotations in estimates of year effects are evident in the United States, Japan, and Taiwan, while estimates for Italy produced a counter-clockwise rotation. The between the two approaches was small only for the U.K. households.

As discussed in section 3, the Deaton-Paxson solution to the identification problem not only prevents linear trends from appearing in year effect, but also may superimpose them over unrestricted estimates of age and cohort effects. For example, any positive trends in data will produce flat year effects, and rotate counter-clockwise the original estimates of age and cohort effects.

These superimposed time trends in age effects are illustrated in Figure 4, which report age-saving profiles from Model 2 with the Deaton-Paxson approach. Comparable estimates from the smoothing cohort model are shown in Figure 1 (presented with thick lines).

In several countries, the rotations in age effects were so large that they greatly distorted the original saving-age profiles. In particular, saving-age profiles of old households were markedly shifted upward in several countries, making it difficult to discern the original decline in savings among aged households. For instance, the smoothing cohort model found a significant upward trends in year effects for Thailand, and, to a lesser degree, in the United States, Japan, and Taiwan (as shown in Figure 3). In each of these countries, the age profile of savings with the Deaton-Paxson approach rotated counter-clockwise.

In sum, while the smoothing cohort model identified decreasing savings in the old age in these countries, the pattern became much less evident with the Deaton-Paxson approach. The most striking contrast between two approaches was in Thailand and Taiwan, and to less degree – in Japan and the United States.

6.2 Consequences of introducing nonlinear demographic effects

Figure 5 returns to estimates from the smoothing cohort model, and reports age profiles from Model 3 that allows a nonlinear impact from the number of children. Estimates from Model 2, which models the demographic effect in a linear way, are also shown for comparison. In most countries, there was little difference between age-saving profiles from these two models. In fact, it is difficult to tell apart estimates for the United States, the United Kingdom, Italy, and Thailand. This happened because when the MGCV algorithm searched for an appropriate degree of smoothing in Model 3, it eventually selected the smoothing parameter λ with linear effects from the number of children, thus going back to Model 2. These outcomes are illustrated in Figure 6, which reports the estimated shape of demographic effect in Model 3, including the number of degrees of freedom v for the demographic variable 'children'.

As discussed in section 4, the MGCV algorithm identifies a linear effect from a variable if the effective degrees of freedom v is less or equal to one. The effective degrees of freedom v for the demographic variable are shown in vertical axis of Figure 6. For example, for Japan

$v = 8.75$, indicating a highly nonlinear pattern that required more than 8 degrees of freedom. On the other hand, it turned out that $v = 1$ in the U.S., U.K., and Italy, implying a linear impact from the number of children.

Overall, the demographic effect on the saving rate turned out clearly nonlinear only in Japan and Taiwan, which explains slightly different age-saving profiles for Models 2 and 3 for these two countries in Figure 5. Therefore, while the increase in number of children had negative effect on the saving rate, the effect for most countries could be represented by a simple linear function.

6.3 The joint nonlinear impact of demographic and cohort effects

Figure 7 reports whether age-saving profiles changed after including in Model 4 an additional term $f(c,q)$ that accounts for the joint impact of demographic and cohort effects. Using age-saving profiles from Model 3 for comparison, it is evident that the joint term had little effect for households in the United States, Japan, in Thailand, implying that the impact of demographic and cohort effects can be modeled in an additive way in these countries. On the other hand, the difference between Models 3 and 4 was particularly large for Taiwanese households. Their age-saving profile still preserved a double-hump profile, but its trough shifted to older households.

The extend of interaction between birth cohorts and the number of children is illustrated in Figure 8. The simplest interaction is evident in the United States and Japan, where the pattern of decreasing saving rates with more children (as shown by arrow next to the axis with title 'children') remained basically the same for different birth cohorts (as can be seen by moving along the axis with title 'cohort').

Conversely, the interaction pattern between c and q turned out more complex in the U.K., Italy, and Taiwan. A particularly noteworthy result was that in younger cohorts, the drop in saving rates with more children was relatively larger, as compared with older cohorts. In other words, the negative effect of larger families on savings became more distinct among recently-born households in these countries.

6.4 What explains the double hump in savings of Asian households?

As shown in Figure 7, the double hump in savings continued to emerge for Asian households in Model 4. This unusual pattern can be attributed to the common practice in many Asian countries that households provide the major financial support for children's education. The increased financial burden is especially heavy for university education, when parents are in their 40s-early 50s. Since educational expenditures are usually classified as consumption expenditures, the increased support of children's education raises the overall consumption expenditures of household, and depresses its saving rate.

To verify whether increased educational expenditures can account for the double hump among Asian households, I used an alternative definition of the saving rate. In this definition, educational expenditures are no longer classified as consumption, but as a part of savings[1].

Among countries with a substantial double hump in savings, it was possible to calculate the alternative definition of saving rate for Japan and Taiwan. Figure 9 reports age-saving profiles from Model 4, with the original and modified definition of saving rate (denoted as 'Model 4' and 'Model 4 (educ)', respectively).

Figure 9 demonstrates that once educational expenditures are counted as savings, the double hump disappeared in Taiwan, and was reduced by about one-third in Japan. Evidently, the temporal burden of financing children's education can account for all of the double hump in Taiwan, while in Japan some other motives apart from education continued to depress savings of middle-age households.

6.5 Robustness to an alternative specification of the spline basis.

I conclude with a small robustness study that examined whether an alternative specification of spline basis could result in different estimates of age-saving profiles. Up to now, the spline function had a cubic spline basis (5), with penalty specified by the integrated square second derivative $\int [f''(x)]^2 dx$. As an alternative basis function, I considered P-splines that were proposed by Eilers and Marx (1996). Compared with cubic splines, P-splines are relatively simple to setup, and have a particularly simple penalty term. The term does not depend on a particular functional form, and is simply a difference penalty on changes in regression parameters β_i for adjacent knots. Large changes between parameters for adjacent knots are used as a measure of function wiggliness. In this robustness study, I used the default specification of P-splines in *mgcv* package, which penalizes second-order differences in regression coefficients.

Overall, estimated saving profiles with P-splines turned out very similar to previously reported estimates with cubic splines[2]. The only noticeable difference was for age-saving profiles in Japan and Taiwan, when expenditures on education were classified as savings. As shown in 9, the alternative definition of savings no longer had the double hump in Taiwan, but the effect was much smaller for Japan. In Figure 10 makes the same comparison of age-saving profiles, but with P-splines in the spline basis. With the alternative specification for spline basis, the double hump disappeared in Japan as well, once educational expenditures are counted as savings.

[1]I am grateful to Charles Yuji Horioka for suggesting this test.

[2]Results with P-splines specification of the basis function are available upon request.

7 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 paper showed that the solution of the identification problem by the Deaton-Paxson approach is not always good, since it can superimpose spurious trends on age-saving profiles in countries with rapidly changing saving rates over time (such as Thailand).

Second, the paper found that the saving rate was decreasing not only in the old age, but 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. The paper provides a tentative evidence for this interpretation, but a more extensive study has to be done.

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.

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

Bacchini, M., Biggeri, A., Lagazio, C., Lertxundi, A., Saez, M., 2007. Parametric and semiparametric approaches in the analysis of short-term effects of air pollution on health. Computational Statistics and Data Analysis 51, 4324–4336.

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

Craven, P., Wahba, G., 1979. Smoothing noisy data with spline functions: estimating the correct degree of smoothing by the method of generalized cross validation. Numerische Mathematik 31, 377–403.

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.

Eilers, P., Marx, B., 1996. Flexible smoothing using b-splines and penalized likelihood (with comments and rejoinder). Statistical Science 11, 89–121.

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.

Horioka, C. Y., 2009. The (Dis)Saving Behavior of the Aged in Japan. Working paper, Institute of Social and Economic Research, Osaka University.

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.

Ruppert, D., Wand, M. P., Carroll, R. J., 2003. Semiparametric Regression. Cambridge University Press, Cambridge.

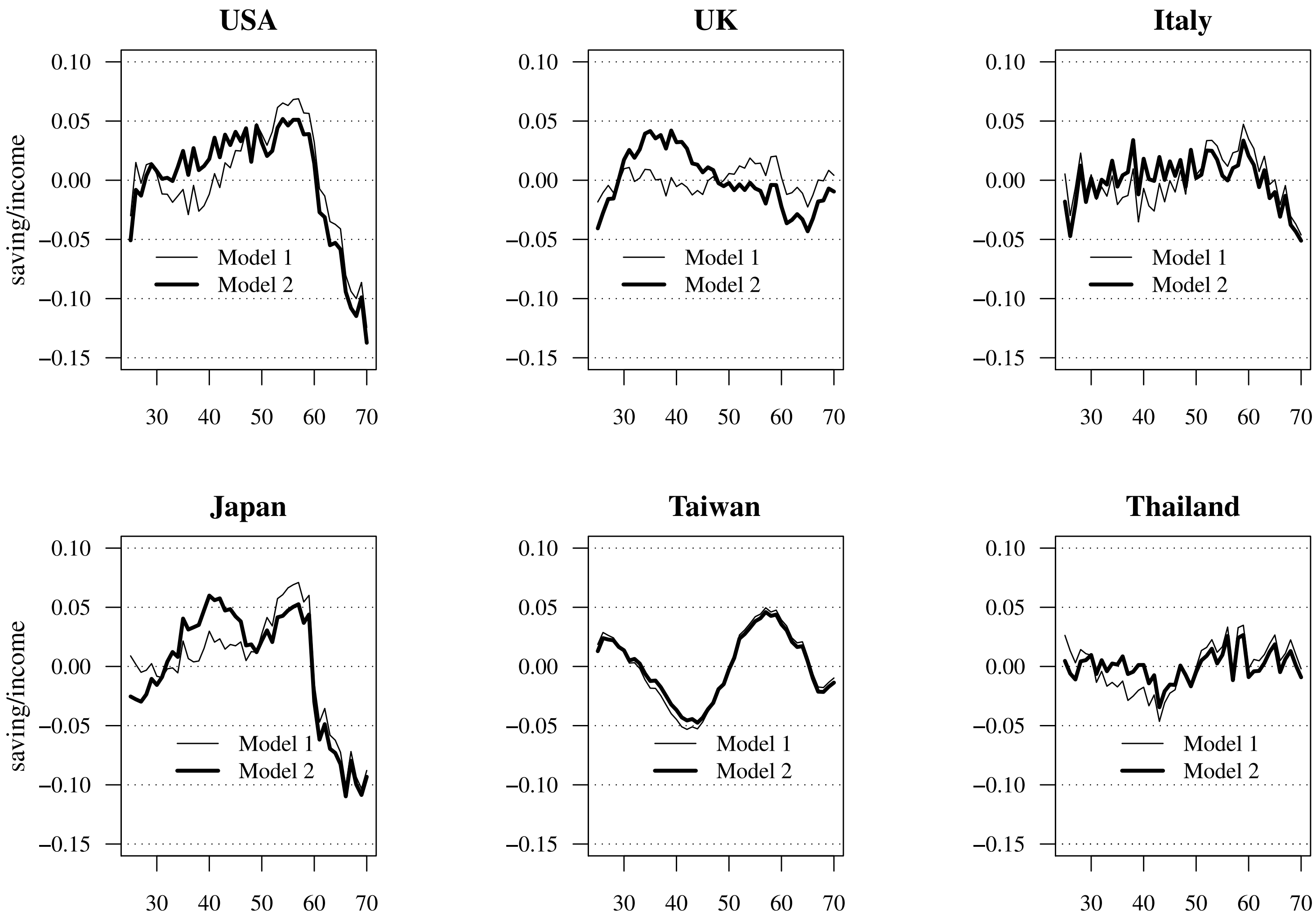
Schimek, M. G., forthcoming. Semiparametric penalized generalized additive models for environmental research and epidemiology. Envinmetrics, accepted for publication.

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.

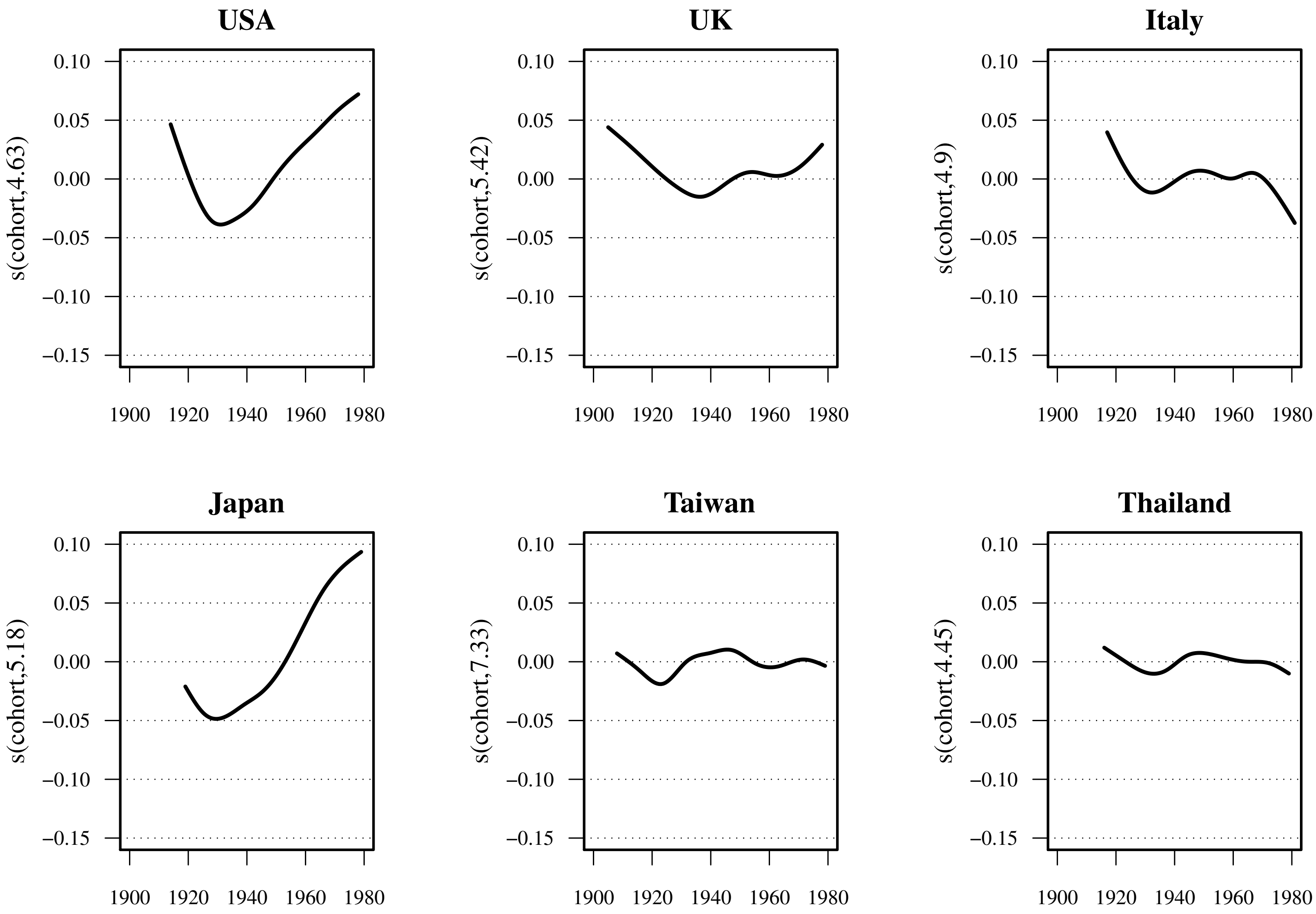
Wood, S., 2006. Generalized Additive Models. An Introduction with R. Chapman and Hall–CRC, Boca Raton, Florida.

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



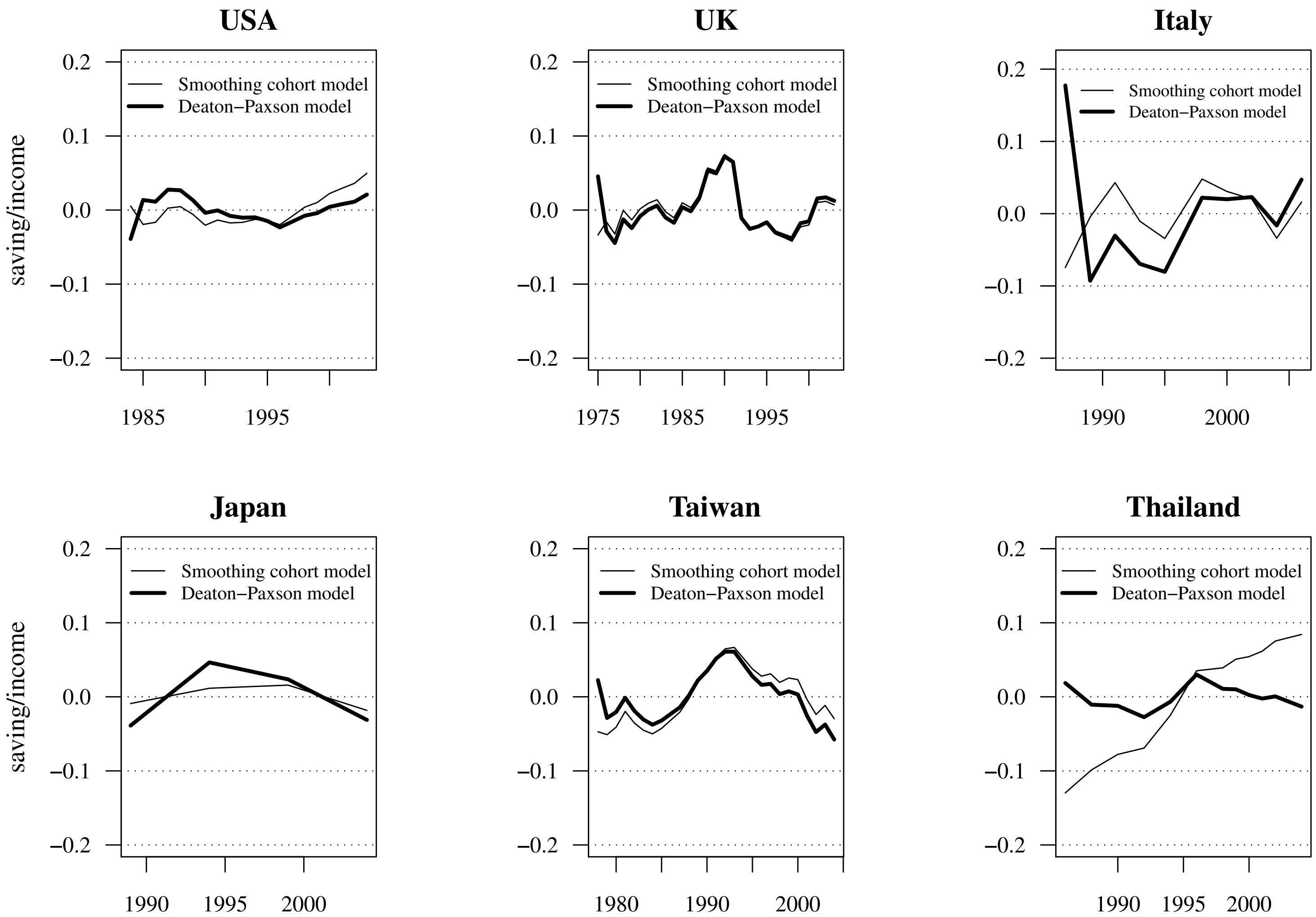
Note: Age effects in models 1 and 2 are estimated by (3) and (9).

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



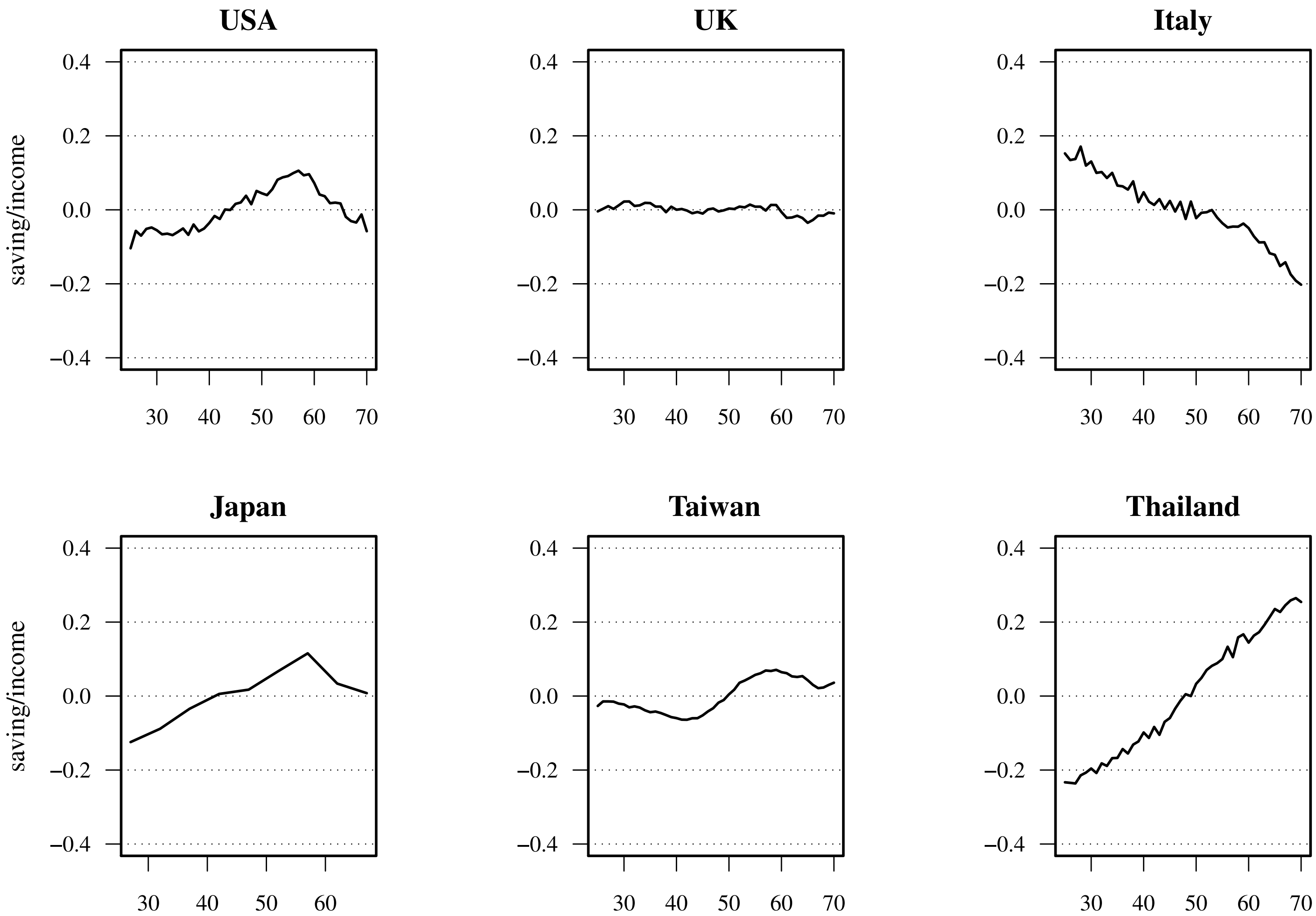
Note: Cohort effects are estimated with Model 2, given by (9).

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



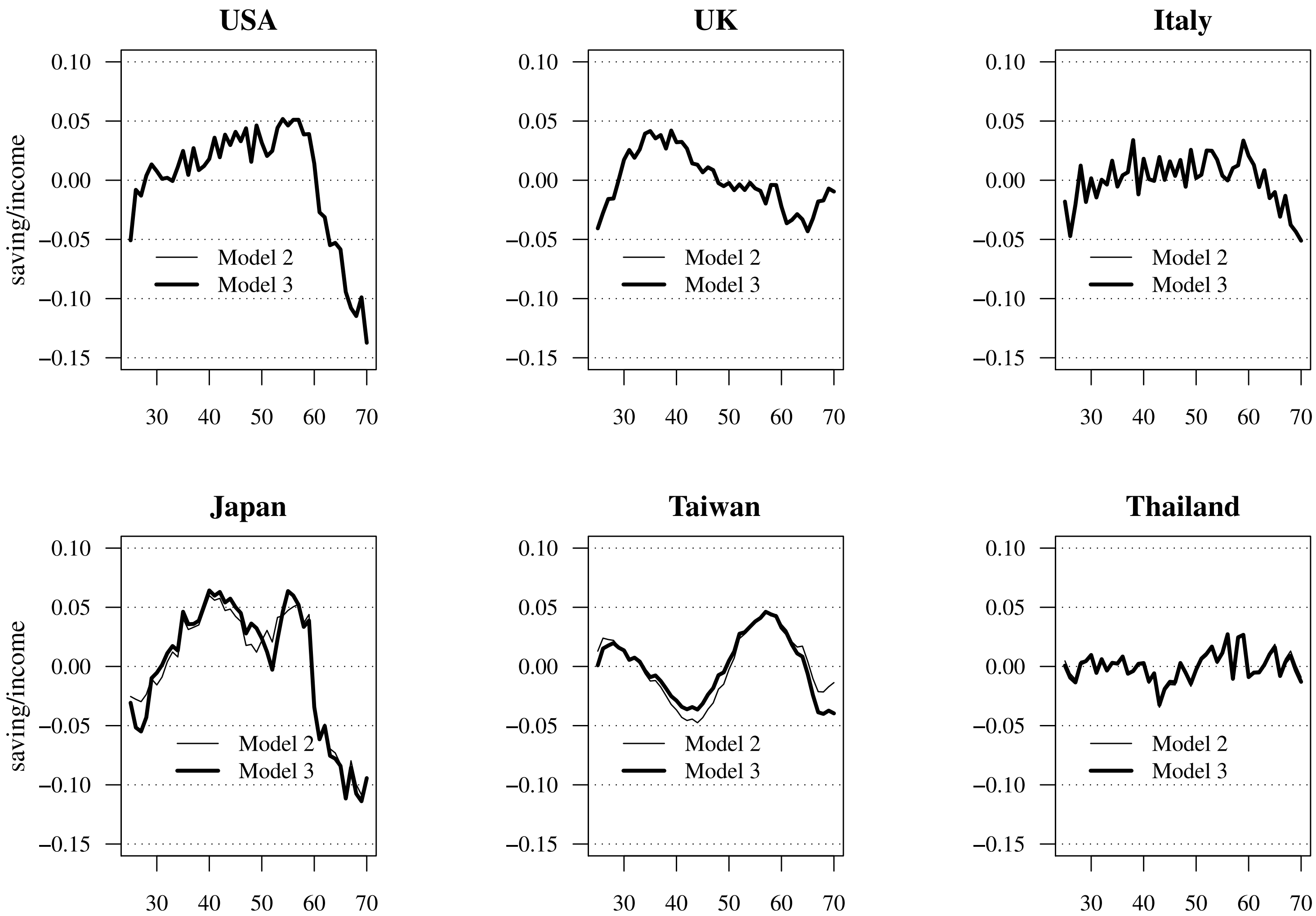
Note: Year effects are estimated with Model 2, given by (9).

Figure 4. Age effects with Deaton-Paxson approach



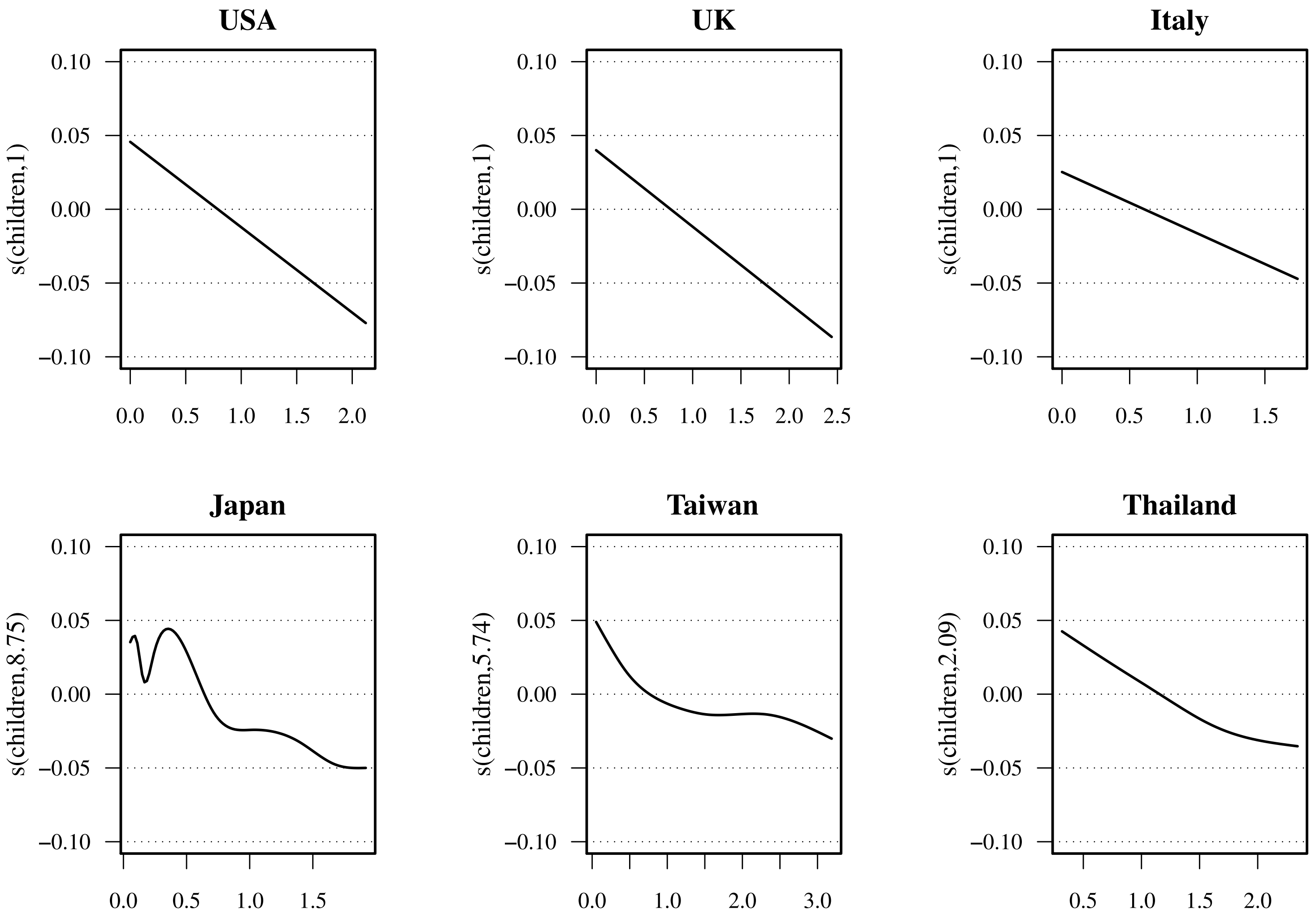
Note: Age effects were estimated with the Deaton-Paxson approach, assuming that year effect are orthogonal to a time trend.

Figure 5. Age effects with Models 2 and 3.



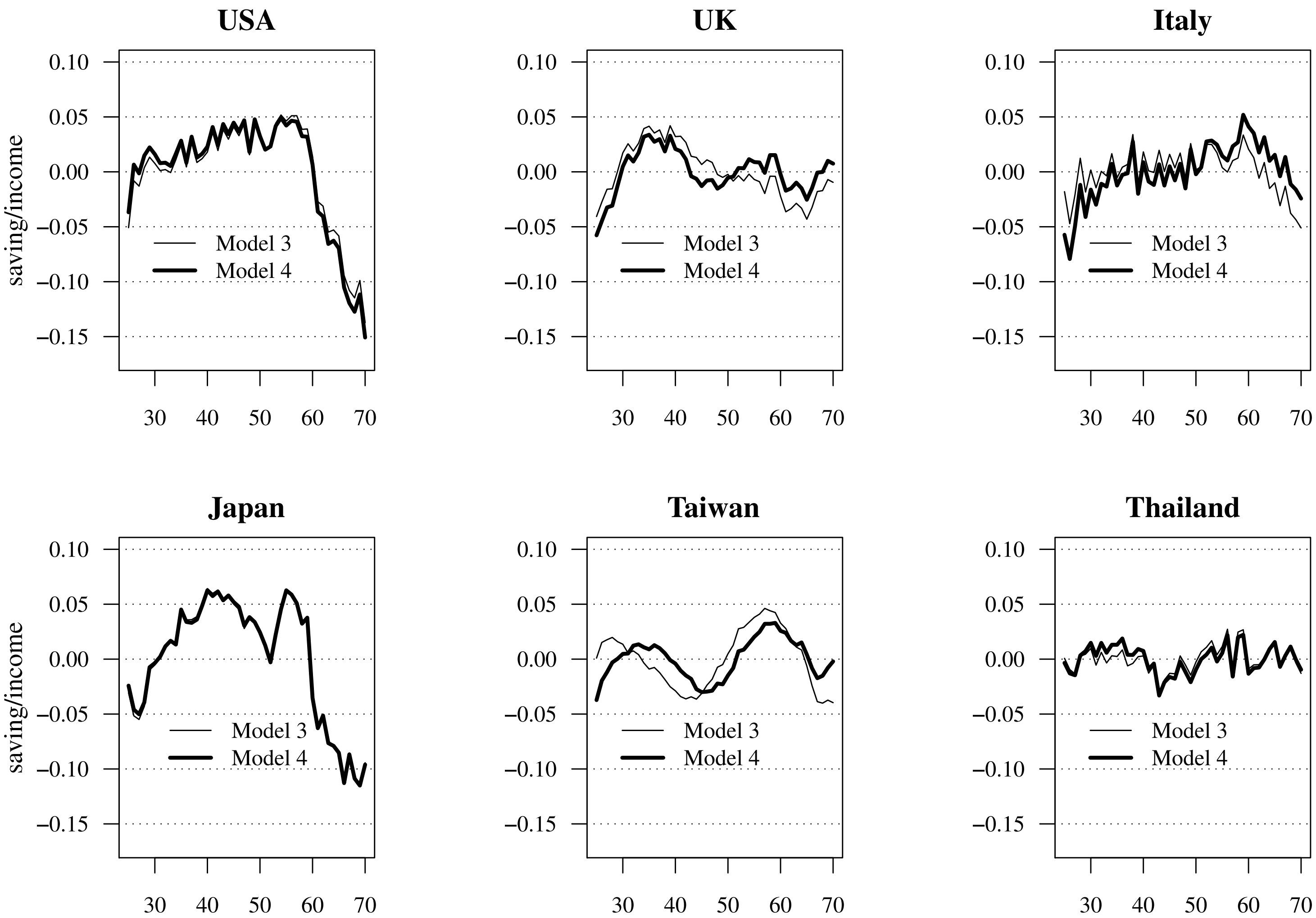
Note: Age effects in models 2 and 3 are estimated by (9) and (10).

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



Note: Demographic effects are from Model 3, estimated by (10).

Figure 7. Age effects in Models 3 and 4.



Note: Age effects in models 3 and 4 are estimated by (10) and (11).

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

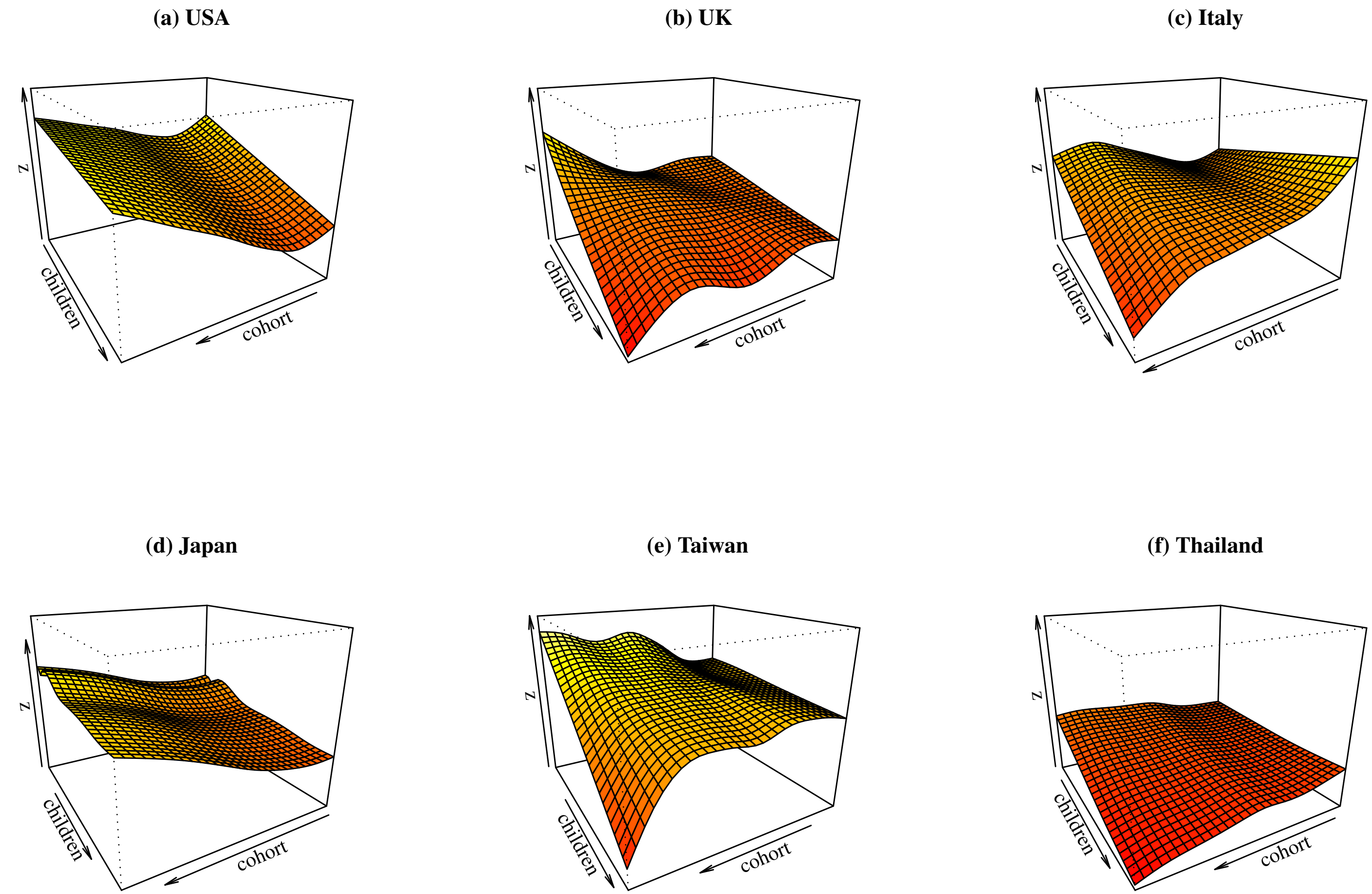
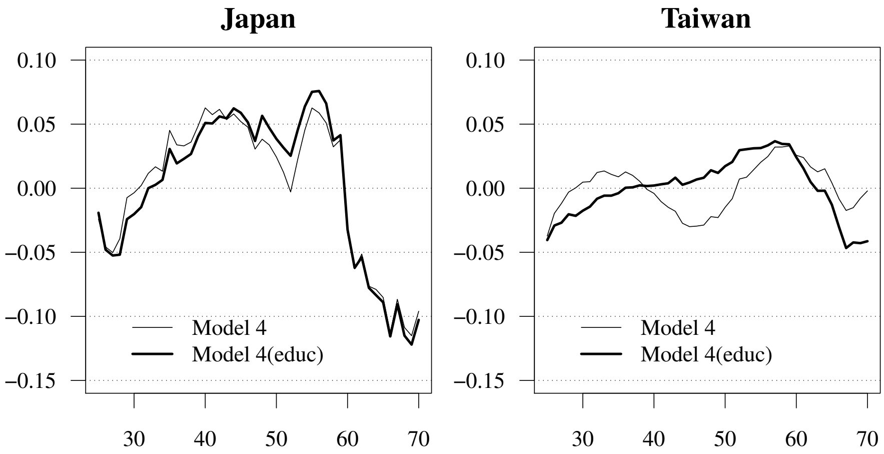
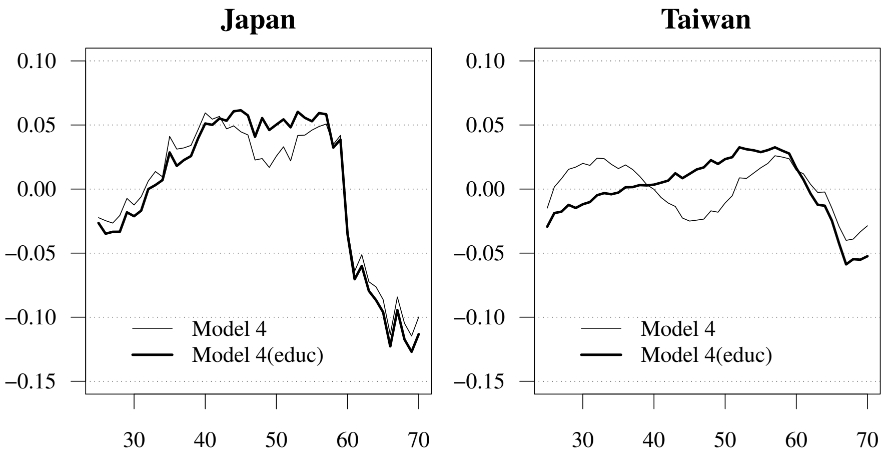


Figure 9. Age effect on savings with educational expenditures included in savings.



Note: The figure compares age-saving profiles, when educational expenditures are counted as a part of non-durable consumption (denoted 'Model 4'), or classified as a part of savings (denoted as 'Model 4 (educ)').

Figure 10. Age effect with educational expenditures classified as savings, and an alternative definition of the spline basis function.



Note: Similarly to figure 9, this figure compares age-saving profiles with different classifications of educational expenditures, with the only difference that P-splines were used in the spline basis function. In contrast, cubic spline basis was used to get age-saving profiles in figure 9.