Federal Reserve Bank of Minneapolis
Research Department

The Age-Time-Cohort Problem and the Identification of Structural Parameters in Life-Cycle Models*

Sam Schulhofer-Wohl

Working Paper 707

June 2013

ABSTRACT

The standard approach to estimating structural parameters in life-cycle models imposes sufficient assumptions on the data to identify the "age profile" of outcomes, then chooses model parameters so that the model's age profile matches this empirical age profile. I show that the standard approach is both incorrect and unnecessary: incorrect, because it generally produces inconsistent estimators of the structural parameters, and unnecessary, because consistent estimators can be obtained under weaker assumptions. I derive an estimation method that avoids the problems of the standard approach. I illustrate the method's benefits analytically in a simple model of consumption inequality and numerically by reestimating the classic life-cycle consumption model of Gourinchas and Parker (2002).

Keywords: Age-time-cohort identification problem; Life-cycle models
JEL classification: C23, D91, J1

---

*Schulhofer-Wohl: Federal Reserve Bank of Minneapolis. I thank Greg Kaplan, Elena Pastorino, Yang Yang, Motohiro Yogo, and participants at the Society for Economic Dynamics annual meeting for helpful comments, Tayyar Buyukbasaran for excellent research assistance, and Joan Gieseke for editorial assistance. The views expressed herein are those of the author and not necessarily those of the Federal Reserve Bank of Minneapolis or the Federal Reserve System.

1. Introduction

A large literature investigates how economic choices and characteristics change over the life cycle. A well-known difficulty in such research is that it is impossible to separately identify the effects of age, time, and birth cohort on the outcome of interest. In this paper, I show that the literature's standard solution to this age-time-cohort identification problem will, in general, cause researchers to make incorrect inferences about the structural parameters of their economic models. I provide a simple alternative that allows accurate identification of the structural parameters, even though age, time, and cohort effects remain unidentified.

Consider an economic model that describes how age affects some outcome of interest, all else equal. Canonical examples include models that describe how age affects the share of a portfolio allocated to stocks (Ameriks and Zeldes, 2004), how inequality among a fixed group of people changes as they age (Deaton and Paxson, 1994a), or how a household optimally arranges consumption over the course of its life (Gourinchas and Parker, 2002). Suppose that, according to the model, an outcome y depends on age a according to

$$y(a) = \xi_0 + q(a; \boldsymbol{\theta}^*), \tag{1}$$

where ξ_0 is an intercept, q is a known function, and $\boldsymbol{\theta}^*$ is a vector of structural parameters of the model, such as parameters of a utility function or of the stochastic process for income. A researcher who has data on the age profile $y(a)$ might seek to estimate $\boldsymbol{\theta}^*$ by choosing $\boldsymbol{\theta}^*$ so that the model's predicted age profile, as given by the right-hand side of (1), comes as close as possible to the observed age profile.

Example. Gourinchas and Parker (2002) build a model in which a household's consumption at age a depends on its rate of time preference and coefficient of relative risk aversion, as well as other parameters. They estimate the model by finding the rate of time preference, risk aversion coefficient, and other parameters that make the model's predicted age profile of consumption come as close as possible to the observed empirical age profile of consumption.[1]

[1]Other papers that estimate or calibrate structural parameters by minimizing the distance between a model's age profile and an empirical age profile include the study of health expenses and saving among the elderly by De Nardi, French, and Jones (2010); the study of household investments by Wachter and Yogo

Equation (1) is oversimplified: In the real world, outcomes y depend not only on age but also on a host of other variables. In particular, outcomes may depend on time and birth cohort. For example, in a study of portfolio choice, an investor's allocation to stocks may depend not only on her age but also on expected returns this year (time) and on whether she is averse to stocks because she grew up during the Great Depression (birth cohort). Thus, observed outcomes $y_{a,t}$ at age a and time t may not satisfy the theoretical relationship (1) but instead may satisfy

$$y_{a,t} = \xi_0 + q(a; \boldsymbol{\theta}^*) + \beta_t + \gamma_c + u_{a,t}, \tag{2}$$

where $c = t - a$ is birth cohort and where β_t and γ_c are time effects and cohort effects, respectively.

To connect the observed data to the model — which says nothing about time effects or cohort effects — the researcher must first empirically estimate "the effect of age on y, holding time and cohort constant," and then compare the model's predictions with these estimates. Unfortunately, it is not possible to identify "the effect of age, holding time and cohort constant," even with a controlled experiment: "What is the effect of age on y, holding time and cohort constant?" is a "fundamentally unidentified question" in the sense of Angrist and Pischke (2009, p. 5). To find the effect of age on y, all else equal, a researcher must collect data at the same instant on two people who were born simultaneously but are now different ages. But this is impossible: If the people are different ages, either they were born at different times or the researcher collected the data at different times. The researcher cannot vary age without varying time or birth cohort.

Recall, though, that estimating the effect of age, holding time and cohort constant, was meant to be only an intermediate step toward estimating the structural parameters $\boldsymbol{\theta}^*$. The parameters may be identified even if the age profile is not. This paper analyzes methods for identifying $\boldsymbol{\theta}^*$ despite the impossibility of identifying the age profile.

(2010); and the studies of life-cycle consumption and inequality by Huggett, Ventura, and Yaron (2011), Kaplan (2012), and Aguiar and Hurst (2013). Deaton and Paxson (1994a,b), Ameriks and Zeldes (2004), and Heathcote, Storesletten, and Violante (2005) follow a similar but more qualitative procedure by comparing various models' broad predictions to the observed relationship between y and a.

A. The standard solution

Suppose that $y_{a,t}$, an outcome of interest for people who are age a in year t, depends on their age a, on the year t, and on their birth year or cohort c in a linear, additively separable manner:

$$y_{a,t} = \xi_0 + \alpha_a + \beta_t + \gamma_c, \tag{3}$$

where ξ_0 is an intercept and α_a, β_t, and γ_c are the coefficients on dummy variables for age, period, and cohort, respectively. The age coefficients α_a represent the age profile of y after controlling for period and cohort effects, and given estimates of the α_a's, one could estimate the structural parameters $\boldsymbol{\theta}^*$ by choosing them to minimize the difference between $q(a;\boldsymbol{\theta}^*)$ and α_a.

However, the age effects α_a in (3) are not identified. One problem is that a complete set of dummy variables would be collinear with the intercept; to avoid this issue, I impose throughout the innocuous normalization that each set of effects is normalized to sum to zero: $\sum_a \alpha_a = \sum_t \beta_t = \sum_c \gamma_c = 0$, and assume that $q(a;\boldsymbol{\theta}^*)$ is normalized similarly.[2] Even with this normalization, however, equation (3) is not identified because, if (3) holds, then for any real number k, the following equation also holds:

$$y_{a,t} = \xi_0 + (\alpha_a + ka - k\bar{a}) + (\beta_t - kt + k\bar{t}) + (\gamma_c + kc - k\bar{c}), \tag{3$'$}$$

where $\bar{a}$, $\bar{t}$, and $\bar{c}$ are the arithmetic means of the possible values of a, t and c.[3] Thus, if $(\xi_0, \alpha_a, \beta_t, \gamma_c)$ are coefficients that fit the data on $y_{a,t}$, then for any real number k, $(\xi_0, \alpha_a + ka - k\bar{a}, \beta_t - kt + k\bar{t}, \gamma_c + kc - k\bar{c})$ are an alternative set of coefficients that fit the data equally well.

The standard method for solving this identification problem is to impose a normalization on either the period or the cohort effects to pin down the unknown real number k, so that the age effects α_a can be identified and the structural parameters $\boldsymbol{\theta}^*$ can be estimated. Two commonly used normalizations are the following:

[2] Equivalently, one could omit one dummy variable from each category, but this would complicate the notation.

[3] Notice that (3$'$) continues to satisfy the normalization that each set of effects sums to zero.

Cohort view: Secular trends appear only in cohort effects, so that the period effects either are orthogonal to a time trend ($\sum_t \beta_t(t-\bar{t}) = 0$), are all zero ($\beta_t = 0$ for all t), or can be replaced with observed variables such as the unemployment rate that measure cyclical economic variation.

Period view: Secular trends appear only in period effects, so that the cohort effects either are orthogonal to a time trend ($\sum_c \gamma_c(c-\bar{c}) = 0$) or are all zero ($\gamma_c = 0$ for all c).

Some authors maintain one of these normalizations throughout the analysis; for example, Deaton and Paxson (1994a,b), Gourinchas and Parker (2002), and De Nardi et al. (2010) adopt the cohort view. Others, such as Ameriks and Zeldes (2004), Heathcote et al. (2005), Wachter and Yogo (2010), Huggett et al. (2011), Kaplan (2012), and Aguiar and Hurst (2013), investigate how their results depend on the choice between the cohort view and the period view; if similar estimates of $\boldsymbol{\theta}^*$ are obtained using both normalizations, the researcher typically argues that the results are not sensitive to the choice of normalization.[4]

B. The flaw in the standard solution

The period view and the cohort view do not span the space of the possible restrictions that could be imposed to identify the age effects in (3). Therefore, even if the period view and the cohort view lead to similar estimates of the structural parameters $\boldsymbol{\theta}^*$, other restrictions on (3) might have led to entirely different estimates of $\boldsymbol{\theta}^*$.

To see this point more clearly, let α_a, β_t, and γ_c be the true age, period, and cohort effects. Equation (3′) shows that, for any real number k, the alternative parameters

$$\alpha_a(k) = \alpha_a + ka - k\bar{a}, \quad \beta_t(k) = \beta_t - kt + k\bar{t}, \quad \gamma_c(k) = \gamma_c + kc - k\bar{c}$$

[4] There is an important but sometimes overlooked difference between normalizing the period or cohort effects to be orthogonal to a trend and normalizing the period or cohort effects to be all zero. If the data cover T time periods and C cohorts, the all-zero normalization imposes T or C restrictions on the coefficients in (3), whereas the orthogonal-to-trend normalization imposes just one restriction. Only one restriction is needed to identify the coefficients, so the all-zero normalization involves overidentifying restrictions. Conclusions about the structural parameters $\boldsymbol{\theta}^*$ might depend on these overidentifying restrictions; thus, if one takes the period or cohort view, it would generally be better to use the orthogonal-to-trend normalization and avoid imposing additional, unnecessary restrictions. In the remainder of the paper, I abstract from the consequences of the overidentifying restrictions in the all-zeros normalizations and assume that researchers have imposed only one linear restriction on the coefficients in (3).

will fit the data on y just as well as α_a, β_t, and γ_c. Effects estimated under the cohort-view normalization correspond to a particular value of k, call it k^{cohort}, that solves $\sum_t \beta_t(k^{cohort})(t-\bar{t}) = 0$. Effects estimated under the period-view normalization correspond to a different value, $k = k^{period}$, that solves $\sum_c \gamma_c(k^{period})(c - \bar{c}) = 0$. The standard approach is to choose the structural parameters $\boldsymbol{\theta}^*$ so that the model's predicted age profile comes as close as possible to either $\alpha_a(0)$ or $\alpha_a(k^{period})$, and if $\alpha_a(0)$ and $\alpha_a(k^{period})$ have similar implications for $\boldsymbol{\theta}^*$, to conclude that the results are not sensitive to the normalization. This reasoning is flawed because there is a continuum of possible normalizations indexed by k,[5] with a corresponding continuum of estimated age profiles $\alpha_a(k)$, and we can conclude that the results are not sensitive to the normalization only if we obtain similar values of $\boldsymbol{\theta}^*$ for all values of k.

As a trivial example, suppose that the theoretical model depends on a scalar parameter θ^*, and suppose the model predicts that the outcome y increases with age if and only if $\theta^* > 0$. If the age effects as estimated under both the period view and the cohort view increase with age, it would be tempting to conclude that $\theta^* > 0$. But this conclusion would be incorrect. For k sufficiently negative, $\alpha_a(k) = \alpha_a + ka - k\bar{a}$ decreases with age, and if a restriction were chosen that corresponded to such a negative value of k, one would obtain age effect estimates that imply $\theta^* \leq 0$. Thus, in this example, the conclusion about the structural parameter θ^* is not truly robust to changes in the assumptions used to identify the age effects, even though the period view and the cohort view give similar results.

C. An alternative solution

The method proposed in this paper exploits the fact that, as (3′) shows, the age effects in (3) are identified up to a single constant k. My method treats this constant as a nuisance parameter to be estimated. In other words, to estimate the structural parameters $\boldsymbol{\theta}^*$, my method estimates the age effects α_a using any one normalization on (3), then chooses k and $\boldsymbol{\theta}^*$ such that $\alpha_a + ka - k\bar{a}$ is as close as possible to $q(a, \boldsymbol{\theta}^*)$.

There may or may not be a unique pair $k, \boldsymbol{\theta}^*$ that minimizes the distance between $\alpha_a + ka - k\bar{a}$ and $q(a, \boldsymbol{\theta}^*)$. If the solution is not unique, then $\boldsymbol{\theta}^*$ is not identified. My

[5] Given any real number k, if we impose the normalization $\sum_t \beta_t(t - \bar{t}) = -k \sum_t (t - \bar{t})^2$, the estimated age, period, and cohort effects will be $\alpha_a(k)$, $\beta_t(k)$, and $\gamma_c(k)$.

method therefore does not guarantee identification of the structural parameters. However, my method makes clear whether identification of the structural parameters relies on the choice of a normalization for the age, time, and cohort effects: If the solution is not unique, then it is impossible to identify the structural parameters without an arbitrary normalization, whereas if the solution is unique, the structural parameters are identified even though the age effects themselves are not identified.

My method amounts to identifying $\boldsymbol{\theta}^*$ from second and higher derivatives of the age profile. McKenzie (2006) shows that the second derivative of the age profile is identified even though the first derivative is not and uses the second derivative to characterize the reduced-form relationship between a and y. The innovation here is that I show how to use the second and all higher derivatives to identify structural parameters of economic models.

The paper proceeds as follows. Section 2 formally describes my proposed method for estimating $\boldsymbol{\theta}^*$ and states conditions under which the structural parameters are identified. The section also shows that, in general, the standard method does not identify the structural parameters. Section 3 illustrates the benefits of my method relative to the standard method by analytically solving a simple life-cycle model of consumption inequality. Section 4 shows quantitatively that my method produces substantially different results when estimating the life-cycle consumption model of Gourinchas and Parker (2002). Section 5 concludes.

2. The method

I assume the researcher has data on a variable $y_{a,t}$ for various ages $a = 1, \ldots, A$ in various time periods t. For example, $y_{a,t}$ could be the cross-sectional variance of log consumption among individuals who are age a in year t. The researcher defines cohorts by $c = t - a$. The researcher also has a theoretical model that says that, in the absence of time effects, cohort effects, and measurement error, y is related to age a and a parameter vector $\boldsymbol{\theta}^*$ according to

$$y = \xi_0 + q(a; \boldsymbol{\theta}^*), \tag{1}$$

where the functional form of q is known a priori and where we normalize $\sum_a q(a; \boldsymbol{\theta}^*) = 0$.

My method for estimating $\boldsymbol{\theta}^*$ is as follows.

1. Estimate the linear model

$$y_{a,t} = \xi_0 + \alpha_a + \beta_t + \gamma_c + u_{a,t}, \tag{4}$$

   where $\xi_0, \alpha_a, \beta_t, \gamma_c$ are parameters and $u_{a,t}$ is an unobservable measurement error, by ordinary least squares subject to the normalization $\sum_a \alpha_a = \sum_t \beta_t = \sum_c \gamma_c = 0$ and to any one additional linear restriction that identifies the parameters. For example, one could assume that the time effects are orthogonal to a trend or that the cohort effects for two adjacent cohorts are equal. The choice of restrictions does not matter so long as there is exactly one, the minimum number required for the matrix of regressors in (4) to be nonsingular given the normalization $\sum_a \alpha_a = \sum_t \beta_t = \sum_c \gamma_c = 0$.
2. Let $\hat{\boldsymbol{\alpha}}$ be the vector of estimated age effects from step 1. Also define the column vectors $\mathbf{a} = [1, \ldots, A]'$ and, for any $\boldsymbol{\theta}$, $\mathbf{q}(\theta) = [q(1, \boldsymbol{\theta}), \ldots, q(A, \boldsymbol{\theta})]'$. Choose $\hat{\boldsymbol{\theta}}$ and $\hat{k}$ to solve

$$(\hat{\boldsymbol{\theta}}, \hat{k}) \in \arg\min_{\boldsymbol{\theta},k} [\mathbf{q}(\boldsymbol{\theta}) - \hat{\boldsymbol{\alpha}} - k\mathbf{a} + k\bar{a}]' \mathbf{W} [\mathbf{q}(\boldsymbol{\theta}) - \hat{\boldsymbol{\alpha}} - k\mathbf{a} + k\bar{a}], \tag{5}$$

   where $\mathbf{W}$ is an $A \times A$ symmetric, positive definite weighting matrix. (For example, $\mathbf{W}$ could be the identity matrix or could be an efficient weighting matrix based on the variance-covariance matrix of $\hat{\boldsymbol{\alpha}}$.) If this problem has a unique solution $\hat{\boldsymbol{\theta}}$, then that solution is my estimator of $\boldsymbol{\theta}^*$. If the solution for $\hat{\boldsymbol{\theta}}$ is not unique, then I conclude that $\boldsymbol{\theta}^*$ is not identified.

The estimated structural parameters will not depend on the normalization used to estimate (4) in step 1. Changing the normalization merely adds a linear trend to the estimated age effects. This trend can be removed by changing the choice of k in problem (5). Thus, the normalization affects the estimator of the nuisance parameter $\hat{k}$ but not the estimator of the parameters of interest $\hat{\boldsymbol{\theta}}$.

Notice, also, that the standard method is identical to my method but imposes $k = 0$ in (5). Thus my method relaxes the assumptions of the standard method.

A. Identification

I have claimed that, if time and cohort effects are additive, my method correctly identifies $\boldsymbol{\theta}^*$ or correctly reports that the parameters are not identified, whereas the standard method may not do so. I now formalize this claim.

Time and cohort effects may enter the data in many ways. I assume that they are additive, so that the linear model (4) is appropriate. Specifically:

ASSUMPTION 1. The observed data satisfy

$$y_{a,t} = \xi_0 + q(a;\boldsymbol{\theta}^*) + \beta_t + \gamma_c + u_{a,t} \tag{6}$$

for some intercept ξ_0, time effects β_t, cohort effects γ_c, and measurement errors $u_{a,t}$ satisfying

$$\mathrm{E}[u_{a,t}|a,t] = 0 \tag{7}$$

and the normalizations

$$\forall\boldsymbol{\theta}\ \sum_a q(a;\boldsymbol{\theta}) = 0, \quad \sum_t \beta_t = \sum_c \gamma_c = 0. \tag{8}$$

Measurement errors also may arise in many ways. I assume that the observed data $y_{a,t}$ are means or other moments calculated from a random sample of $N_{a,t}$ individuals who are age a in time period t and that $u_{a,t}$ arises from sampling error, so that in the limit as the sample size goes to infinity, $u_{a,t}$ is asymptotically normal:

ASSUMPTION 2. For all a and t, $\sqrt{N_{a,t}}u_{a,t} \xrightarrow{d} N(0, s_u^2)$ in the limit as $N_{a,t} \to \infty$.

We then have the following result:

PROPOSITION 1. *Under assumptions 1 and 2 and standard regularity assumptions, in the limit as $N_{a,t}$ goes to infinity for all a and t, either the solution $\hat{\boldsymbol{\theta}}$ to problem* (5) *converges in probability to a unique vector, which is $\boldsymbol{\theta}^*$, or the solution to problem* (5) *does not converge to a unique vector and $\boldsymbol{\theta}^*$ is not identified.*

Proof. Under assumptions 1 and 2 and standard regularity assumptions, when the researcher follows step 1 of my method, his estimates satisfy

$$\hat{\boldsymbol{\alpha}} = \mathbf{q}(\boldsymbol{\theta}^*) - k^*\mathbf{a} + k^*\bar{a} + \boldsymbol{\epsilon} \tag{9}$$

for some constant k^* and some asymptotically normal random vector $\boldsymbol{\epsilon}$ that is orthogonal to $\mathbf{a}$ in finite samples and that converges in probability to zero as $N_{a,t}$ goes to infinity. Problem (5) in step 2 then becomes

$$(\hat{\boldsymbol{\theta}}, \hat{k}) \in \arg\min_{\boldsymbol{\theta},k} [\mathbf{q}(\boldsymbol{\theta}) - \mathbf{q}(\boldsymbol{\theta}^*) + (k^* - k)(\mathbf{a} - \bar{a}) - \boldsymbol{\epsilon}]'\mathbf{W}[\mathbf{q}(\boldsymbol{\theta}) - \mathbf{q}(\boldsymbol{\theta}^*) + (k^* - k)(\mathbf{a} - \bar{a}) - \boldsymbol{\epsilon}]. \tag{10}$$

Under standard regularity conditions, the solution to (10) converges in probability to the solution to

$$(\tilde{\boldsymbol{\theta}}, \tilde{k}) \in \arg\min_{\boldsymbol{\theta},k} [\mathbf{q}(\boldsymbol{\theta}) - \mathbf{q}(\boldsymbol{\theta}^*) + (k^* - k)(\mathbf{a} - \bar{a})]'\mathbf{W}[\mathbf{q}(\boldsymbol{\theta}) - \mathbf{q}(\boldsymbol{\theta}^*) + (k^* - k)(\mathbf{a} - \bar{a})]. \tag{11}$$

One solution to (11) is $\tilde{k} = k^*$ and $\tilde{\boldsymbol{\theta}} = \boldsymbol{\theta}^*$. If this is the unique solution, then we have shown that $\hat{\boldsymbol{\theta}}$ converges in probability to $\boldsymbol{\theta}^*$. If (11) has multiple solutions, there exist $(\tilde{k}, \tilde{\boldsymbol{\theta}}) \neq (k^*, \boldsymbol{\theta}^*)$ such that

$$\mathbf{q}(\tilde{\boldsymbol{\theta}}) - \mathbf{q}(\boldsymbol{\theta}^*) = (\tilde{k} - k^*)(\mathbf{a} - \bar{a}). \tag{12}$$

If (12) holds, then either there are two parameter vectors that generate the same age profile (so $\mathbf{q}(\tilde{\boldsymbol{\theta}}) - \mathbf{q}(\boldsymbol{\theta}^*) = \mathbf{0}$) or the difference between the age profiles generated by the two parameter vectors is linear in age. In the former case, the age profile is clearly not sufficient to identify the parameters. In the latter case, the fact that the age profile is identified only up to an unknown linear trend is an insurmountable obstacle to identifying the parameters; only by imposing an untestable, possibly incorrect normalization would we be able to identify the parameters. Hence, if (11) has multiple solutions, $\boldsymbol{\theta}^*$ is not identified.

B. Remarks

Interpretation in terms of detrended age profiles. One way to interpret my method is that it chooses the structural parameters $\boldsymbol{\theta}^*$ so that the *detrended* age profile from the model matches, as closely as possible, the *detrended* age profile in the data. Specifically, we can decompose the model's age profile as

$$q(a;\boldsymbol{\theta}^*) = c_{model}(a-\bar{a}) + \check{q}(a;\boldsymbol{\theta}^*), \tag{13}$$

where c_{model} is the slope in a linear regression of $q(a;\boldsymbol{\theta}^*)$ on a and, therefore, by construction, $\check{q}(a;\boldsymbol{\theta}^*)$ is orthogonal to a linear trend in a. Because $\check{q}(a;\boldsymbol{\theta}^*)$ is orthogonal to a, we can describe $\check{q}(a;\boldsymbol{\theta}^*)$ as the detrended age profile from the model. Similarly, if $\hat{\boldsymbol{\alpha}}$ is the vector of age effects estimated in the data under any just-identified normalization on (4), then we can write

$$\hat{\alpha}_a = c_{normalization}(a-\bar{a}) + \check{\alpha}_a, \tag{14}$$

where $c_{normalization}$ is the slope in a linear regression of $\hat{\alpha}_a$ on a and, by construction, $\check{\alpha}_a$ is orthogonal to a. The minimization problem (5) in step 2 of my method can now be rewritten as

$$\begin{aligned}(\hat{\boldsymbol{\theta}},\hat{k}) \in \arg\min_{\boldsymbol{\theta},k} [\check{\mathbf{q}}(\boldsymbol{\theta}) + c_{model}(\mathbf{a}-\bar{a}) - \check{\boldsymbol{\alpha}} - c_{normalization}(\mathbf{a}-\bar{a}) - k\mathbf{a} + k\bar{a}]' \\ \mathbf{W}[\check{\mathbf{q}}(\boldsymbol{\theta}) + c_{model}(\mathbf{a}-\bar{a}) - \check{\boldsymbol{\alpha}} - c_{normalization}(\mathbf{a}-\bar{a}) - k\mathbf{a} + k\bar{a}],\end{aligned} \tag{15}$$

and because $\check{\alpha}_a$ and $\check{q}$ are both orthogonal to a, the solution to (15) is

$$k = c_{model} - c_{normalization}, \tag{16a}$$

$$\hat{\boldsymbol{\theta}} \in \arg\min_{\boldsymbol{\theta}} [\check{\mathbf{q}}(\boldsymbol{\theta}) - \check{\boldsymbol{\alpha}}]'\mathbf{W}[\check{\mathbf{q}}(\boldsymbol{\theta}) - \check{\boldsymbol{\alpha}}]. \tag{16b}$$

Equation (16b) shows that my method detrends the age profiles from both the data and the model, then chooses the structural parameters to make the two detrended age profiles as close as possible. This procedure is invariant to the normalization that is used to estimate the age profile in the data, because equation ($3'$) demonstrates that changing the normalization will change $c_{normalization}$ but not the detrended age profile $\breve{\alpha}_a$.

Incorrect results from the standard method. The standard method will generally produce incorrect results even when my method produces correct results. The standard method proceeds as follows: Let $\tilde{\boldsymbol{\alpha}}$ be the vector of estimated age effects obtained by imposing the period-view normalization. (Analogous results apply if one uses the cohort view.) The standard method estimates the structural parameters by

$$\tilde{\boldsymbol{\theta}} = \arg\min_{\boldsymbol{\theta}} [\mathbf{q}(\boldsymbol{\theta}) - \tilde{\boldsymbol{\alpha}}]' \mathbf{W} [\mathbf{q}(\boldsymbol{\theta}) - \tilde{\boldsymbol{\alpha}}]. \tag{17}$$

Let γ_c be the true cohort effects and $\tilde{\gamma}_c$ be the cohort effects estimated under the period view. Then $\tilde{\gamma}_c = \gamma_c + k(c - \bar{c})$ for some number k. Under the period view, $\sum_c \tilde{\gamma}_c (c - \bar{c}) = 0$, which implies

$$k = k^{period} \equiv -\frac{\sum_c (c - \bar{c})\gamma_c}{\sum_c (c - \bar{c})^2}. \tag{18}$$

Now, since the data satisfy (6), the researcher will obtain

$$\tilde{\boldsymbol{\alpha}} = \mathbf{q}(\boldsymbol{\theta}^*) + k^{period}(\mathbf{a} - \bar{a}) + \boldsymbol{\epsilon}. \tag{19}$$

The researcher using the standard method therefore estimates the structural parameters by

$$\tilde{\boldsymbol{\theta}} = \arg\min_{\boldsymbol{\theta}} [\mathbf{q}(\boldsymbol{\theta}) - \mathbf{q}(\boldsymbol{\theta}^*) - k^{period}(\mathbf{a} - \bar{a}) - \boldsymbol{\epsilon}]' \mathbf{W} [\mathbf{q}(\boldsymbol{\theta}) - \mathbf{q}(\boldsymbol{\theta}^*) - k^{period}(\mathbf{a} - \bar{a}) - \boldsymbol{\epsilon}]. \tag{20}$$

Unless $k^{period} = 0$ — that is, unless the normalization imposed in the standard method is correct — $\boldsymbol{\theta}^*$ generally does not solve problem (20).

Multiple age profiles. In many applications, researchers examine the age profiles of two or more variables. There is no reason to use the same normalization on the age, time, and cohort effects for all variables. Hence, a different slope k_j should be estimated for each variable j.

Example. Suppose that the theoretical model makes predictions about the age profiles of both income i and consumption c:

$$
\begin{aligned}
i(a) &= \xi_{0,i} + q_i(a;\boldsymbol{\theta}^*),\\
c(a) &= \xi_{0,c} + q_c(a;\boldsymbol{\theta}^*).
\end{aligned}
\tag{21}
$$

The structural parameters should be estimated as follows. First, estimate (4) separately for income and consumption, obtaining estimated age effects $\hat{\boldsymbol{\alpha}}_i$ and $\hat{\boldsymbol{\alpha}}_c$. Second, estimate the structural parameters by solving

$$
(\hat{\boldsymbol{\theta}}, \hat{k}_i, \hat{k}_c) \in \arg\min_{\boldsymbol{\theta}, k_i, k_c} \begin{bmatrix} \mathbf{q}_i(\boldsymbol{\theta}) - \hat{\boldsymbol{\alpha}}_i - k_i\mathbf{a} + k_i\bar{a} \\ \mathbf{q}_c(\boldsymbol{\theta}) - \hat{\boldsymbol{\alpha}}_c - k_c\mathbf{a} + k_c\bar{a} \end{bmatrix}' \mathbf{W} \begin{bmatrix} \mathbf{q}_i(\boldsymbol{\theta}) - \hat{\boldsymbol{\alpha}}_i - k_i\mathbf{a} + k_i\bar{a} \\ \mathbf{q}_c(\boldsymbol{\theta}) - \hat{\boldsymbol{\alpha}}_c - k_c\mathbf{a} + k_c\bar{a} \end{bmatrix}, \tag{22}
$$

where $\mathbf{W}$ is now a $(2A) \times (2A)$ symmetric, positive definite weighting matrix.

3. Analytic example: consumption inequality over the life cycle

In this section, I exhibit a simple analytic example in which the standard method does not identify the structural parameters of an economic model but my method does.

A. The economic model

An agent i is born in year c and lives for $A+1$ periods, $t = c, c+1, \ldots, c+A$. The agent begins life with assets $x_{i,0,c} > 0$ and receives a stochastic income $y_{i,a,t}$ in each period $t = c, c+1, \ldots, c+A$. Income is independently and identically distributed across agents and dates with mean μ and variance σ^2. Let $C_{i,a,t}(y_i^t)$ be i's consumption in year t, when he is age $a = t - c$, after a history of income shocks $y_i^t \equiv (y_{i,0,c}, \ldots, y_{i,a,t})$. The agent's preferences are represented by

$$
-\frac{1}{2}\mathrm{E}_c \sum_{a=0}^{A} \rho^a [\bar{C} - C_{i,a,t}(y_i^t)]^2, \tag{23}
$$

where ρ is the rate of time preference and $\bar{C}$ is a bliss level of consumption. The agent can borrow or save without limit at the nonstochastic gross interest rate $(1+r) = \rho^{-1}$, except that the agent cannot borrow at age A. Thus, the law of motion of assets x for $a < A$ is

$$x_{i,a+1,t+1}(y_i^t) = (1+r)[x_{i,a,t}(y_i^{t-1}) + y_{i,a,t} - C_{i,a,t}(y_i^t)], \quad a = 0, \dots, A-1. \tag{24}$$

To keep the notation concise, in the remainder of the analysis, I suppress the dependence of x and C on the history y_i^t. The agent maximizes (23) by choice of $\{C_{i,a,t}, x_{i,a+1,t+1}\}_{a=0}^{A}$, subject to (24) and

$$C_{i,A,c+A} = x_{i,A,c+A} + y_{i,A,c+A}, \tag{25}$$

taking r and $x_{i,0,c}$ as given. It can be shown (see, e.g., Krueger, 2007, section 3.2) that the solution to the agent's problem is

$$C_{i,a,c+a} = (1+\phi_a)^{-1}(x_{i,a,c+a} + y_{i,a,c+a} + \mu\phi_a), \tag{26}$$

where

$$\phi_a = \sum_{s=1}^{A-a} \rho^s = \rho \frac{1-\rho^{A-a}}{1-\rho}. \tag{27}$$

It can also be shown that

$$C_{i,a+1,c+a+1} - C_{i,a,c+a} = (1+\phi_{a+1})^{-1}(y_{i,a+1,c+a+1} - \mu). \tag{28}$$

It follows from (28) that the cross-sectional variance of consumption among agents who are age a and born in cohort c is

$$\mathrm{Var}[c_{i,a,c+a}|a,c] = (1+\phi_0)^{-2}\mathrm{Var}[x_{i,0,c}] + \sigma^2 \sum_{s=0}^{a}(1+\phi_s)^{-2}. \tag{29}$$

B. Identification

The parameters of the economic model are A, ρ, and σ^2. I assume A is known. I now show that, under reasonable assumptions on measurement error, my method identifies ρ and σ^2 despite the age-time-cohort identification problem. (In addition, the distribution of $x_{i,0,c}$ is a nuisance parameter; I will not discuss identification of it here.)

Suppose that, as in Deaton and Paxson (1994a), an econometrician observes consumption in repeated cross sections of agents of various ages at various dates. Assume that i's consumption is measured with error: The econometrician observes

$$\hat{C}_{i,a,t} = C_{i,a,t} + \epsilon_{i,a,t}, \tag{30}$$

where the measurement error $\epsilon_{i,a,t}$ is independent of $C_{i,a,t}$, uncorrelated across agents, and has mean $\nu_{a,t}$ and variance η_t^2 at date t. (The bias $\nu_{a,t}$ and measurement error variance η_t^2 could change over time due to, for example, changes in the survey instrument. I show below that the structural parameters can be identified without identifying $\nu_{a,t}$ and η_t^2.)

Since the econometrician has repeated cross sections and not a true panel, he cannot estimate the parameters by looking at the time series of an agent's consumption. However, he can construct moments of consumption for each age and date and create a synthetic panel. The mean of observed consumption is uninformative because of the age- and time-varying bias $\nu_{a,t}$. The variance of observed consumption among people who are age a at date t is

$$\mathrm{Var}[\hat{C}_{i,a,t}|a,t] = \eta_t^2 + \mathrm{Var}[C_{i,a,t}|a,t] = \eta_t^2 + (1+\phi_0)^{-2}\mathrm{Var}[x_{i,0,c}] + \sigma^2 \sum_{s=0}^{a} (1+\phi_s)^{-2}. \tag{31}$$

This is identical to (3) with $\xi_0 = 0$, $\alpha_a = \sigma^2 \sum_{s=0}^{a}(1+\phi_s)^2$, $\beta_t = \eta_t^2$, and $\gamma_c = (1+\phi_0)^{-2}\mathrm{Var}[x_{i,0,c}]$. It follows that my method identifies σ^2 and ρ as long as the following equations have a unique solution $\hat{\sigma^2} = \sigma^2$, $\hat{\rho} = \rho$, $k = 0$:

$$\sigma^2 \sum_{s=0}^{a} \left(1 + \rho \frac{1-\rho^{A-s}}{1-\rho}\right)^{-2} = ka + \hat{\sigma^2} \sum_{s=0}^{a} \left(1 + \hat{\rho} \frac{1-\hat{\rho}^{A-s}}{1-\hat{\rho}}\right)^{-2}, \quad a = 0, \ldots, A. \tag{32}$$

It is clear that $\hat{\sigma^2} = \sigma^2$, $\hat{\rho} = \rho$, $k = 0$ is one solution to the equations; therefore, we need to prove only that there is no other solution. Specializing to $a = 0, 1, 2$, we have

$$\sigma^2 \left(1 + \rho \frac{1-\rho^A}{1-\rho}\right)^{-2} = \hat{\sigma^2} \left(1 + \hat{\rho} \frac{1-\hat{\rho}^A}{1-\hat{\rho}}\right)^{-2} \tag{33a}$$

$$\sigma^2 \sum_{s=0}^{1} \left(1 + \rho \frac{1-\rho^{A-s}}{1-\rho}\right)^{-2} = k + \hat{\sigma^2} \sum_{s=0}^{1} \left(1 + \hat{\rho} \frac{1-\hat{\rho}^{A-s}}{1-\hat{\rho}}\right)^{-2} \tag{33b}$$

$$\sigma^2 \sum_{s=0}^{2} \left(1 + \rho \frac{1-\rho^{A-s}}{1-\rho}\right)^{-2} = 2k + \hat{\sigma^2} \sum_{s=0}^{2} \left(1 + \hat{\rho} \frac{1-\hat{\rho}^{A-s}}{1-\hat{\rho}}\right)^{-2}. \tag{33c}$$

Using (33a) to substitute for $\hat{\sigma^2}$ in (33b) and (33c), then using (33b) to eliminate k and simplifying, we have

$$\left(\frac{1-\rho^{A-1}}{1-\rho^{A+1}}\right)^{-2} - \left(\frac{1-\rho^{A}}{1-\rho^{A+1}}\right)^{-2} = \left(\frac{1-\hat{\rho}^{A-1}}{1-\hat{\rho}^{A+1}}\right)^{-2} - \left(\frac{1-\hat{\rho}^{A}}{1-\hat{\rho}^{A+1}}\right)^{-2}. \tag{34}$$

For $\hat{\rho} \in (0, 1)$, the right-hand side of (34) is monotonically increasing in $\hat{\rho}$; therefore, (34) has a unique solution, which is $\hat{\rho} = \rho$. We then obtain $\hat{\sigma^2} = \sigma^2$ from (33a) and $k = 0$ from (33b). Thus, the solution is unique, and my method identifies σ^2 and ρ. By contrast, the standard method will not identify σ^2 and ρ unless the cohort or period effects happen to be orthogonal to a linear trend.

4. Quantitative example: revisiting Gourinchas and Parker (2002)

Gourinchas and Parker (2002) estimate the structural parameters of a life-cycle model in which households receive a stochastic income and decide how much to consume and how much to save. In this section, I investigate how the results change when I use my method instead of the standard method that they employed to estimate the parameters.

A. Model

I briefly review the model here and refer readers to the original paper for details.

Households work for $T = 40$ periods and then retire. Their preferences are given by

$$\mathrm{E}\left[\sum_{t=1}^{T} \beta^t \frac{(C_t/Z_t)^{1-\rho}}{1-\rho} + \beta^{T+1}\kappa(\zeta_{T+1}/Z_{T+1})^{1-\rho}\right], \tag{35}$$

where β is the rate of time preference, ρ is the coefficient of relative risk aversion, Z_t is a deterministic family size adjustment for households of age t, κ is a constant, and ζ_{T+1} is terminal liquid and illiquid wealth. Households choose consumption and savings at each age to maximize utility given an initial liquid wealth level W_1, the constraint that terminal liquid wealth W_{T+1} is non-negative, and the budget constraint

$$W_{t+1} = R(W_t + Y_t - C_t). \tag{36}$$

Income Y_t evolves according to

$$Y_t = P_t U_t, \quad P_t = G_t P_{t-1} N_t, \tag{37}$$

where P_t is the permanent component of income, N_t is an i.i.d. permanent shock, G_t is the deterministic growth rate of permanent income, and U_t is an independent and identically distributed (i.i.d.) transitory shock. The transitory shocks are 0 with probability p and otherwise follow a log-normal distribution with mean 0 and variance σ_u^2. The permanent shocks follow a log-normal distribution with mean 0 and variance σ_n^2.

It can be shown that if terminal illiquid wealth is $H_{T+1} = hP_{T+1}$, where h is a constant, then the terminal value function $\kappa(\zeta_{T+1}/Z_{T+1})^{1-\rho}$ induces the household to follow a terminal consumption rule that is linear in liquid wealth normalized by permanent income,

$$\frac{C_{T+1}}{P_{T+1}} = \gamma_0 + \gamma_1 \frac{W_{T+1} + Y_{T+1}}{P_{T+1}}, \tag{38}$$

where γ_1 is the marginal propensity to consume out of terminal wealth and $\gamma_0 = h\gamma_1$. In the remainder of the analysis, the model is expressed in terms of γ_0 and γ_1 instead of κ and h.

B. Original estimation procedure

In the first stage of their estimation procedure, Gourinchas and Parker (2002) use external data to estimate the interest rate R, the variances of the income shocks σ_u^2 and σ_n^2, the probability of zero income p, and the mean initial wealth level W_1.

Next, Gourinchas and Parker (2002) use repeated cross sections from the Consumer Expenditure Survey to estimate age profiles of consumption, income, and family size. The age profile of consumption is estimated by an equation analogous to (3), with log consumption as the dependent variable, but dummy variables for family size are added to adjust for differences in family size, and the time effects are replaced by the unemployment rate to solve the identification problem. The age profile of the family size adjustment Z_t is calculated as the mean of the coefficients on the age dummies, weighted by the distribution of family sizes among households of age t; thus, there are assumed to be no period or cohort effects in the family size adjustment. Finally, income is normalized by the estimated family size adjustment, and the age profile of normalized income is estimated from an equation analogous to (3) but with time effects replaced by the unemployment rate.

The remaining parameters of the model are β, ρ, γ_0, and γ_1. They are chosen by the Method of Simulated Moments to fit the age profile of consumption, given the first-stage parameters and the estimated age profiles of income and family size. That is, given a parameter vector $\theta = (\beta, \rho, \gamma_0, \gamma_1)$ and first-stage parameters χ, Gourinchas and Parker (2002) calculate the household's consumption rule in the model, simulate the behavior of a large number of households, and solve

$$\min_{\beta,\rho,\gamma_0,\gamma_1} [\overline{\ln C_t} - \ln \widehat{C_t(\theta, \chi)}]' \mathbf{W} [\overline{\ln C_t} - \ln \widehat{C_t(\theta, \chi)}], \tag{39}$$

where $\overline{\ln C_t}$ is the estimated age profile of log consumption in the data, $\ln \widehat{C_t(\theta, \chi)}$ is the mean of log consumption among simulated households of age t, and $\mathbf{W}$ is a weighting matrix.

C. Replication

Before implementing my method of identifying structural parameters, I replicated the results of Gourinchas and Parker (2002) using their estimation method. Jonathan Parker kindly shared with me the estimated age profiles and the GAUSS code used to estimate the parameters for the original paper. Because so much time has passed since the original code was written, I could not obtain access to a copy of the GAUSS software that was capable of running the original code, so I wrote new code in C++ to reproduce the original code.[6] My code follows as closely as possible all of the decisions made in the original code, such as the

[6] I use the nonlinear optimization package of Johnson (2012), random number generators from the Intel Math Kernel Library, and some utilities from Galassi et al. (2011).

grid and interpolation method used to approximate the consumption rule. These decisions are largely documented in the appendix to Gourinchas and Parker (2002), and my new code is available to interested readers upon request.

The parameters that minimize my implementation of the objective function (39) are close but not identical to the estimates published by Gourinchas and Parker (2002). Table 1 shows the original parameter estimates from Gourinchas and Parker (2002) as well as the parameters that minimize my implementation of the objective function. Following Gourinchas and Parker (2002), I focus on results using a robust weighting matrix based on the variance of the estimated age profile; results using the optimal weighting matrix proved to be unstable due to the need to numerically differentiate the objective function to estimate the optimal weights. The discrepancy between my results and those of Gourinchas and Parker (2002) for identical estimation procedures could be due to differences in the random number draws used for the simulations, differences in the accuracy of the nonlinear equation solver that is used to solve the household's Euler equation,[7] or differences in the numerical accuracy of the calculations. (For example, the numerical gradient estimates used to calculate standard errors depend on a tolerance whose value in GAUSS I could not determine; the standard error estimates are very sensitive to this tolerance, perhaps explaining why my standard errors are quite different from those in the published paper.) In all, though, the discrepancies in the point estimates are small and show that my replication essentially reproduces the published point estimates. If there are economically significant differences in the point estimates when I apply my estimation method, those differences must be due to the change in method — not to differences between my replication code and the original code.

D. Estimation without normalizations on the age profiles

Gourinchas and Parker's (2002) estimation procedure uses arbitrary normalizations to estimate the age profiles of consumption, income, and family size. I now use the new estimation method described in section 2 to examine the consequences of these normalizations for the estimates of the structural parameters $\beta, \rho, \gamma_0, \gamma_1$.

[7] Gourinchas and Parker (2002) use a built-in solver in GAUSS. I was unable to determine details of its implementation.

Gourinchas and Parker (2002) actually impose more normalizations than are necessary to solve the age-time-cohort identification problem and identify the age profiles: For consumption and income, they restrict the time effects to move in parallel with the unemployment rate, whereas for family size, they restrict the cohort and time effects to be zero. To maintain comparability with the original results, I do not relax the extra restrictions. Instead, I treat Gourinchas and Parker's (2002) estimated age profiles as if they were estimated using only the minimum required restrictions — i.e., one arbitrary normalization on the first derivative of each age profile — and then apply my estimation method. Because there are three age profiles, I estimate three arbitrary slopes along with the structural parameters.

The model in Gourinchas and Parker (2002) suffers from an additional identification problem that is unrelated to the age-time-cohort problem. Suppose that $R\beta = 1$ and $\rho = 0$. Then the household is indifferent as to the timing of consumption, and *any* observed age profile of consumption that satisfies the budget constraint is consistent with the model. In practice, if $R\beta = 1$ and $\rho = 0$, the simulated age profile of consumption from the model will mirror the initial guess that is used to find a consumption rule that satisfies the household's Euler equation. Thus, when applying the new estimation method, if there is some slope k such that the observed age profile minus this slope is close to the initial guess used to solve the Euler equation, the new estimation method will converge to $R\beta = 1$ and $\rho = 0$. These estimates, of course, are not meaningful. Therefore, I impose $R\beta < 1$ and $\rho > 0$.

Table 2 shows the results. For reference, column 1 repeats my estimates using the standard method. (Results from the new method should be compared with my estimates using the standard method, rather than with the published estimates, because my code produces results slightly different from the published estimates even when applying the standard method.) Column 2 allows an arbitrary trend in the age profile of consumption; this change causes the estimated coefficient of relative risk aversion to more than double — to 1.78 from 0.69 — and decreases γ_1, the marginal propensity to consume out of final wealth, by nearly 20 percent. In columns 3 and 4, I instead allow arbitrary trends in family size or income instead of consumption; these changes have relatively little effect on the structural parameters. Finally, in column 5, I allow arbitrary trends in all three age profiles — consumption, family size, and income. With all three trends allowed, the coefficient of relative risk aversion is similar to that obtained using a consumption trend and the marginal propensity to consume out of final wealth is even lower. Allowing arbitrary trends also significantly improves the model fit as measured by the χ^2 statistic.

Because the consumption profile is the one in which allowing an arbitrary trend has the largest consequences for the structural parameters, it is instructive to examine the age profile of consumption that the model generates under different parameters. Figure 1 shows detrended age profiles of the natural logarithm of consumption — specifically, the figure plots residuals from regressing the age profile of log consumption on a linear trend in age. (The figure shows log consumption rather than the level of consumption because Gourinchas and Parker, 2002, estimate the age profile of log consumption in the data and use it to construct their moment conditions; thus, the age-time-cohort normalization most directly affects the profile of log consumption.) By removing a linear trend in age, I remove the effect of the normalization on the empirical age profile and make it possible to focus on the curvature of the age profile, which is what my method uses to identify the structural parameters. Figure 1 shows that my method brings the curvature of the age profile of consumption in the model closer to the curvature in the data, compared with the standard method, by choosing structural parameters that make the age profile less curved during the first half of the life cycle.

As table 2 shows, the structural parameters that best match the curvature of the age profile of consumption include substantially higher risk aversion and a lower propensity to consume from final wealth than the parameters estimated by the standard method. Of course, changing these parameters has implications not only for the curvature of the age profile but also for its first derivative. Figure 2 plots the age profile of the level of consumption as simulated with the parameters estimated by the standard method and by my method. With higher risk aversion, the precautionary motive is stronger, and households save more early in life; thus, consumption rises faster with age. Such a pattern would be grossly inconsistent with the age profile of consumption that Gourinchas and Parker (2002) estimate in the data using the normalization they chose. Therefore, the standard method strongly rejects a high coefficient of relative risk aversion. But the first derivative of the age profile is unidentified and is a function of the normalization, not of the data. My method does not allow this unidentified first derivative to drive inferences about the structural parameters. Instead, my method identifies the structural parameters by matching the curvature, as shown in figure 1. Then, as shown in figure 2, my method allows the age profile in the data to rotate freely by adding or subtracting a linear trend, so that the empirical age profile is made consistent with the age profile that the structural parameters predict.

In general, table 2 shows that allowing arbitrary trends increases the standard errors, but not by very much. The estimates remain relatively precise even after allowing for arbitrary trends. Hence, in the model of Gourinchas and Parker (2002), the structural parameters remain well identified without having to resort to unneeded normalizations on age profiles, but removing those normalizations substantially changes one's conclusions about the true values of the parameters — significantly increasing the coefficient of relative risk aversion and reducing the marginal propensity to consume out of final wealth. One caveat is that, both under the original estimation method and when we remove the normalizations on age profiles, the χ^2 statistics imply that the overidentifying restrictions are strongly rejected. Thus, there may be some doubt as to how well the model describes the data.

5. Conclusion

In estimating structural life-cycle models, the age-time-cohort identification problem arises because researchers must project two-dimensional data — data that vary with both age and time — onto a one-dimensional model that varies only with age. There are many ways to make such projections. The standard approach to estimating structural parameters of life-cycle models assumes a particular projection is correct, then estimates the structural parameters conditional on that assumption. What I show in this paper is that the standard approach's assumption is unnecessary and, in general, leads to incorrect results. I provide an alternative approach that does not have this pitfall. My method demonstrates that the structural parameters can be identified even without imposing enough assumptions to identify the age profile.

As I have discussed, my method identifies the structural parameters from their effect on the curvature and higher derivatives of the age profile, rather than on its slope. If the curvature and higher derivatives are not precisely estimated or if parameters have only weak effects on these derivatives, then confidence intervals for the structural parameters will be large, although in practice, my method did not greatly increase the standard errors on structural parameters in a replication of Gourinchas and Parker (2002). Adding assumptions, as in the standard method, has the potential to make the confidence intervals smaller — but only at the price of potentially producing incorrect estimates. My method allows researchers to determine what they can learn about the structural parameters with only a minimal set of assumptions.

My approach does, however, make some significant assumptions. The additive, linear model (3) assumes that time effects have the same impact on people of all ages and that time effects matter only contemporaneously. Schulhofer-Wohl and Yang (2011) argue that many important economic and social phenomena violate these assumptions and propose a model that avoids them. However, their model requires many years of data and minimal measurement error. In this paper, I have focused on the widely used and easy-to-estimate linear model and asked how best to estimate structural parameters using it. Analysis of more complex models is left for future research.

References

Aguiar, Mark, and Erik Hurst, 2013, "Deconstructing Life Cycle Expenditure," *Journal of Political Economy* 121(3), 437–492.

Ameriks, John, and Stephen P. Zeldes, 2004, "How Do Household Portfolio Shares Vary With Age?" manuscript, Columbia University.

Angrist, Joshua D., and Jörn-Steffen Pischke, 2009, *Mostly Harmless Econometrics: An Empiricist's Companion*, Princeton, NJ: Princeton University Press.

Deaton, Angus, and Christina Paxson, 1994a, "Intertemporal Choice and Inequality," *Journal of Political Economy* 102(3), 437–467.

Deaton, Angus S., and Christina H. Paxson, 1994b, "Saving, Growth, and Aging in Taiwan," in Wise, David A., editor, *Studies in the Economics of Aging*, pp. 331–357, Chicago: University of Chicago Press.

De Nardi, Mariacristina, Eric French, and John B. Jones, 2010, "Why Do the Elderly Save? The Role of Medical Expenses," *Journal of Political Economy* 118(1), 39–75.

Galassi, Mark, Jim Davies, James Theiler, Brian Gough, Gerard Jungman, Patrick Alken, Michael Booth, and Fabrice Rossi, 2011, *GNU Scientific Library Reference Manual*, version 1.15.

Gourinchas, Pierre-Olivier, and Jonathan A. Parker, 2002, "Consumption Over the Life Cycle," *Econometrica* 70(1), 47–89.

Heathcote, Jonathan, Kjetil Storesletten, and Giovanni L. Violante, 2005, "Two Views of Inequality Over the Life Cycle," *Journal of the European Economic Association* 3(2–3), 765–775.

Huggett, Mark, Gustavo Ventura, and Amir Yaron, 2011, "Sources of Lifetime Inequality," *American Economic Review* 101(7), 2923–2954.

Johnson, Steven G., 2012, "The NLopt Nonlinear-Optimization Package," version 2.3, available at `http://ab-initio.mit.edu/nlopt`.

Kaplan, Greg, 2012, "Inequality and the Lifecycle," *Quantitative Economics* 3(3), 471–525.

Krueger, Dirk, 2007, "Consumption and Saving: Theory and Evidence," manuscript, University of Pennsylvania.

McKenzie, David J., 2006, "Disentangling Age, Cohort and Time Effects in the Additive Model," *Oxford Bulletin of Economics and Statistics* 68(4), 473–495.

Schulhofer-Wohl, Sam, and Yang Yang, 2011, "Modeling the Evolution of Age and Cohort Effects in Social Research," Research Department Staff Report 461, Federal Reserve Bank of Minneapolis.

Wachter, Jessica A., and Motohiro Yogo, 2010, "Why Do Household Portfolio Shares Rise in Wealth?" *Review of Financial Studies* 23(11): 3929–3965.

Table 1: Replication of Gourinchas and Parker (2002).

	Published	Replication
discount factor (β)	0.9598	0.9547
	(0.0179)	(0.0026)
risk aversion (ρ)	0.5140	0.6870
	(0.1707)	(0.0091)
γ_0	0.0015	8.87×10^{-6}
	(3.85)	(0.0091)
γ_1	0.0710	0.0672
	(0.1244)	(0.0039)
$\chi^2(36)$	174.10	155.95

Estimates using robust weighting matrix. Standard errors and χ^2 statistics corrected for first-stage estimation.

Table 2: Comparison of estimation methods.

	Standard	New method			
	(1)	(2)	(3)	(4)	(5)
Structural parameters:					
discount factor (β)	0.9547	0.9579	0.9547	0.9553	0.9558
	(0.0026)	(0.0035)	(0.0028)	(0.0018)	(0.0041)
risk aversion (ρ)	0.6870	1.7809	0.6878	0.6855	1.7308
	(0.0091)	(0.0177)	(0.0432)	(0.0289)	(0.1397)
γ_0	8.87×10^{-6}	5.78×10^{-6}	8.93×10^{-6}	8.30×10^{-6}	1.33×10^{-5}
	(0.0091)	(0.0024)	(0.0086)	(0.0055)	(0.0053)
γ_1	0.0672	0.0547	0.0672	0.0660	0.0487
	(0.0039)	(0.0032)	(0.0042)	(0.0052)	(0.0060)
Slope nuisance parameters:					
$k_{consumption}$	0	0.0154	0	0	0.0136
	-	(0.0021)	-	-	(0.0032)
$k_{family\ size}$	0	0	8.38×10^{-6}	0	-0.0042
	-	-	(0.0025)	-	(0.0035)
k_{income}	0	0	0	-0.0010	-0.0003
	-	-	-	(0.0006)	(0.0048)
χ^2	155.95	109.45	152.79	128.33	101.99
d.f.	36	35	35	35	33

Standard errors and χ^2 statistics corrected for first-stage estimation. "d.f." indicates degrees of freedom.

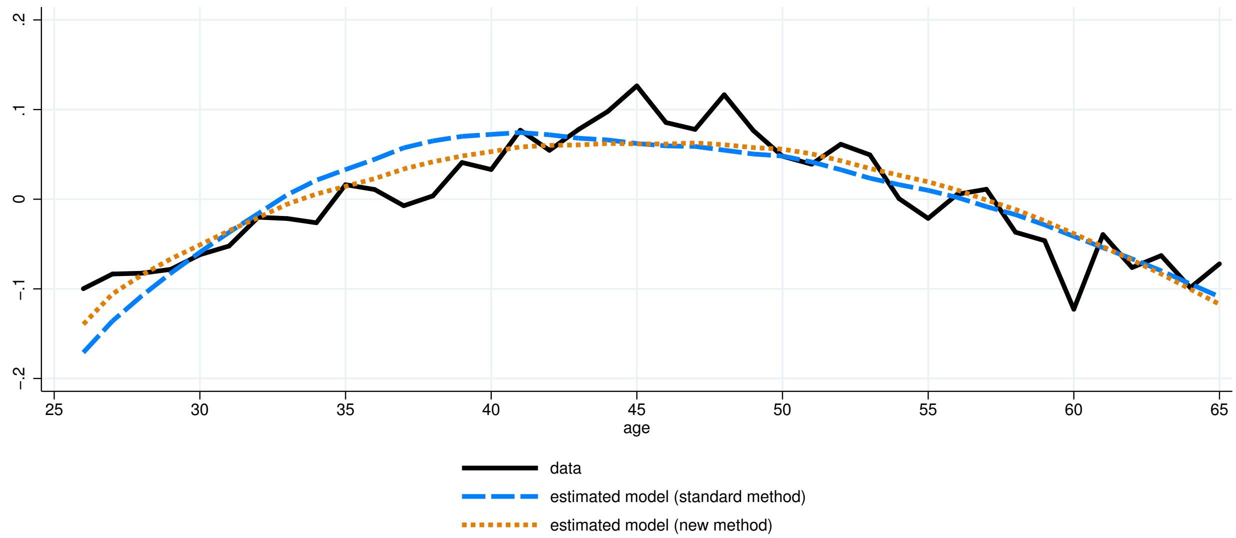

Figure 1: Detrended age profiles of ln(consumption).

Graph shows residuals from regressing age profiles of the natural logarithm of consumption on a linear trend in age. Lines labeled "estimated model (standard method)" and "estimated model (new method)" are simulated from the model using parameter values in table 2, columns 1 and 5, respectively.

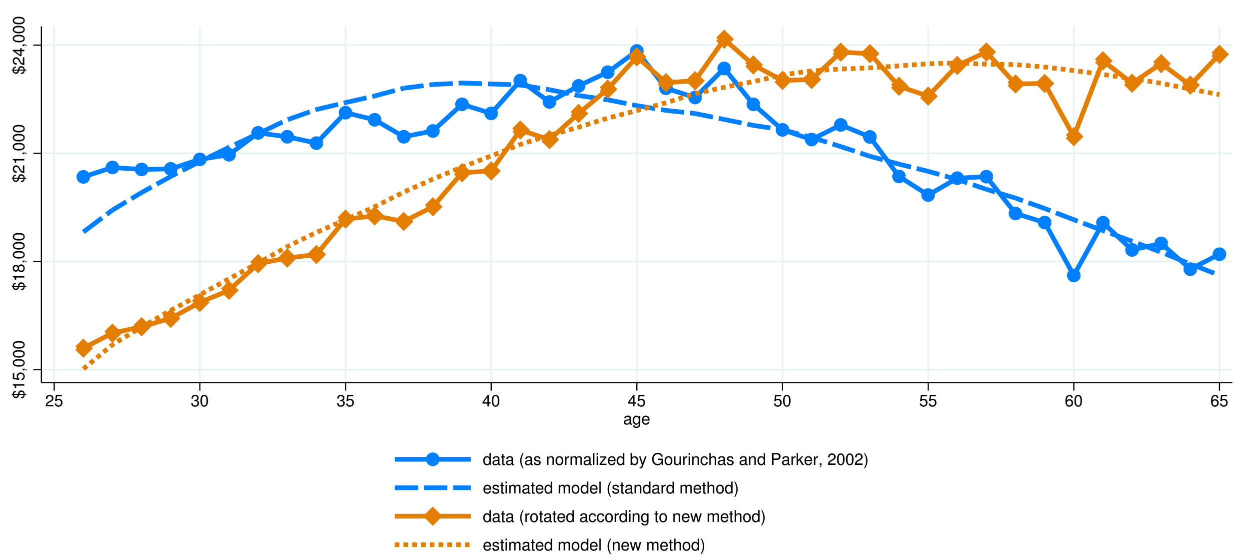

Figure 2: Age profiles of consumption.

Lines labeled "estimated model (standard method)" and "estimated model (new method)" are simulated from the model using parameter values in table 2, columns 1 and 5, respectively. The line labeled "data (rotated according to new method)" is the age profile in the data, rotated by the estimated consumption trend shown in table 2, column 5.