

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

February 2012

DRAFT: DO NOT CITE OR CIRCULATE
WITHOUT PERMISSION

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 generically produces inconsistent estimators of the structural parameters, and unnecessary, because consistent estimators can be obtained under weaker fewer assumptions. I derive an identification method that avoids the problems of the standard approach and illustrate its benefits in a simple model of consumption inequality.

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

*Schulhofer-Wohl: Federal Reserve Bank of Minneapolis. I thank Greg Kaplan, Yang Yang, and Motohiro Yogo for helpful comments. 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 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. For example, the model may describe how age affects the share of a portfolio allocated to stocks (?), how inequality among a fixed group of people changes as they age (??), or how a household optimally arranges consumption over the course of its life (?). Suppose that, according to the model, an outcome y depends on age a according to

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

where 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 y and a might seek to estimate $\boldsymbol{\theta}^*$ by choosing $\boldsymbol{\theta}^*$ such that the model’s predicted relationship between y and a , as given in (??), is as close as possible to the observed empirical relationship between y and a . ? and ? follow exactly this procedure; ? and ? follow it loosely by comparing models’ qualitative predictions to the observed relationship between y and a . Other papers that follow this procedure include the study of health expenses and saving

among the elderly by ?; the study of wages, hours, and consumption by ?; and the study of household investments by ?.

Equation (??) is oversimplified: In the real world, outcomes y depend not only on age but on a host of other variables. To connect the data to the model, the researcher must first empirically estimate “the effect of age on y , holding all other variables constant” (a quantity I will henceforth denote $\partial y/\partial a$). Unfortunately, it is not possible to estimate $\partial y/\partial a$, even with a controlled experiment: “What is the effect of age on y , all else equal?” is a “fundamentally unidentified question” in the sense of ?, p. 5.¹

“The effect of age on y , all else equal” is fundamentally unidentified because time and birth cohort may also affect y . For example, an investor’s portfolio allocation to stocks may depend not only on her age but also on likely returns this year (time) and on whether she grew up during the Great Depression (birth cohort) and is thus averse to stocks. To find the effect of age on portfolio share, all else equal, we 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 we collected the data at different times. We cannot vary age without varying time or birth cohort.

Recall, though, that estimating $\partial y/\partial a$ was meant to be only an intermediate step toward estimating the structural parameters θ^* . The parameters may be identified even if $\partial y/\partial a$ is not. This paper analyzes methods for identifying θ^* despite the impossibility of identifying $\partial y/\partial a$.

¹I thank Joshua Angrist for colorfully making this problem clear to me.

A. The standard solution

? expressed the identification problem and the standard solution to it as follows. 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 :

$$y_{a,t} = f(a, t, c). \quad (2)$$

Because $a = t - c$, there exist functions $g(a, t) = f(a, t, t - a)$, $h(a, c) = f(a, a + c, c)$ and $j(t, c) = f(t - c, t, c)$ such that (??) can be written as

$$y_{a,t} = g(a, t) \quad \text{— or —} \quad y_{a,t} = h(a, c) \quad \text{— or —} \quad y_{a,t} = j(t, c). \quad (3)$$

A researcher who wants to nonparametrically model the dependence of y on age, time and cohort must drop one of these three explanatory variables and estimate (??) rather than (??).

In practice, instead of nonparametric models, researchers typically employ linear models of the form

$$y_{a,t} = \alpha_a + \beta_t + \gamma_c. \quad (??')$$

where α_a , β_t , and γ_c are the coefficients on dummy variables for age, period, and cohort, respectively. The identification problem arises in (??) just as it does in (??): If equation (??) holds, then for any real number k , the following equation also holds:

$$y_{a,t} = (\alpha_a + ka) + (\beta_t - kt) + (\gamma_c + kc). \quad (??'')$$

The standard method for estimating age effects is, following the idea of (??), to eliminate either time or cohort effects from (??) and instead estimate one of the following equations:

$$y_{a,t} = \alpha_a + \beta_t \quad (??a')$$

$$\text{--- or ---} \quad y_{a,t} = \alpha_a + \gamma_c. \quad (??b')$$

That is, to estimate the structural parameters θ^* in (??), the researcher estimates age effects α_a using either (??) or (??), then chooses the parameters θ^* such that $q(a, \theta^*)$ is as close as possible to the estimated age effects α_a . Sometimes, one of equations (??) or (??) is used to obtain “preferred estimates” and the other equation is used as a robustness check; if similar estimates of θ^* are obtained using both equations, the researcher may argue that the results are not sensitive to the choice of identification strategy.

B. The flaw in the standard solution

Although the nonparametric equation (??) implies the restricted nonparametric equation (??), the linear version (??) does not imply the restricted linear equations (??) and (??). In particular, only one linear restriction on the parameters is sufficient to achieve identification in (??), but (??) and (??) impose many more restrictions: as many restrictions as there are cohorts or time periods, respectively. Therefore, equations (??) and (??) do not cover the full range of possible estimates of age effects; even if (??) and (??) lead to similar estimates of the structural parameters θ^* , other restrictions on (??) might have led to entirely different estimates of θ^* .

C. An alternative solution

The method proposed in this paper exploits the fact that, as (??) shows, the age effects in (??) 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 θ^* , my method estimates the age effects α_a using any one normalization on (??), then chooses k and θ^* such that $\alpha_a + ka$ is as close as possible to $q(a, \theta^*)$.

There may or may not be a unique pair θ^* that minimizes the distance between $\alpha_a + ka$ and $q(a, \theta^*)$. If the solution is not unique, then θ^* is not identified. My method therefore does not guarantee identification of 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 unique, then the structural parameters are identified even though the age effects themselves are not identified.

My method amounts to identifying θ^* from second and higher derivatives of the age profile $\partial y / \partial a$. Previous papers (e.g., ?) have observed that the second derivative of $\partial y / \partial a$ is identified even though the first derivative is not and have used the second derivative to characterize the relationship between a and y . The innovation here is that I show how to use the second derivative to identify structural parameters of economic models.

The paper proceeds as follows. Section 2 formally describes my proposed method for estimating θ^* and states conditions under which the structural parameters are identified. The section also shows that, in general, the standard 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 concludes.

2. The method

I assume the researcher has data on a variable $y_{a,t}$ for various ages $a = 1, \dots, A$ 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 and cohort effects, y is related to age a and a parameter vector $\boldsymbol{\theta}^*$ according to

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

where the functional form of q is known a priori.

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

1. Estimate the linear model

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

by ordinary least squares using any just-identified set of restrictions on the dummy variable coefficients $\{\alpha_a\}$, $\{\beta_t\}$ and $\{\gamma_c\}$. Four restrictions are needed to identify the equation. For example, these restrictions could be omitting one age dummy, one time dummy and two cohort dummies. Alternatively, one could normalize each set of effects to sum to zero — $\sum_a \alpha_a = \sum_t \beta_t = \sum_c \gamma_c = 0$ — and impose one additional linear restriction, such as that the time effects are orthogonal to a trend. The choice of restrictions does not matter so long as there are exactly four restrictions, the minimum number required for the matrix of regressors in (??) to be nonsingular.

2. Let $\hat{\boldsymbol{\alpha}}$ be the vector of estimated age effects from step 1. Also define the column vectors

$\mathbf{a} = [1, \dots, A]'$ and, for any $\boldsymbol{\theta}$, $\mathbf{q}(\boldsymbol{\theta}) = [q(1, \boldsymbol{\theta}), \dots, q(A, \boldsymbol{\theta})]'$. Choose $\hat{\boldsymbol{\theta}}$ and $\hat{k}$ to solve

$$(\hat{\boldsymbol{\theta}}, \hat{k}) \in \arg \min_{\boldsymbol{\theta}, k} [\mathbf{q}(\boldsymbol{\theta}) - \hat{\boldsymbol{\alpha}} - k\mathbf{a}]' [\mathbf{q}(\boldsymbol{\theta}) - \hat{\boldsymbol{\alpha}} - k\mathbf{a}]. \quad (5)$$

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.

3. In many applications, researchers examine the age profiles of two or more variables (for example, income and consumption). 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 .

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, while the standard method may not do so. I now formalize this claim.

Time and cohort effects may enter the data in many ways. My assumption is that they are additive, so that the linear model (??) is appropriate. Specifically, I assume that the observed data satisfy

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

for some parameters β_t and γ_c . Given that (??) holds, when the researcher follows step 1 of my method, his estimates satisfy

$$\hat{\boldsymbol{\alpha}} = \mathbf{q}(\boldsymbol{\theta}^*) - k^* \mathbf{a} \quad (7)$$

for some constant k^* . The researcher's problem in step 2 is then

$$(\hat{\boldsymbol{\theta}}, \hat{k}) \in \arg \min_{\boldsymbol{\theta}, k} [\mathbf{q}(\boldsymbol{\theta}) - \mathbf{q}(\boldsymbol{\theta}^*) + k^* \mathbf{a} - k \mathbf{a}]' [\mathbf{q}(\boldsymbol{\theta}) - \mathbf{q}(\boldsymbol{\theta}^*) + k^* \mathbf{a} - k \mathbf{a}]. \quad (8)$$

One solution to this problem is $\hat{k} = k^*$ and $\hat{\boldsymbol{\theta}} = \boldsymbol{\theta}^*$. If this is the unique solution, then my method correctly identifies the parameters. If this is not the unique solution, my method reports that the parameters are not identified. Is this conclusion correct? If there are multiple solutions, there exist $(\hat{k}, \hat{\boldsymbol{\theta}}) \neq (k^*, \boldsymbol{\theta}^*)$ such that

$$\mathbf{q}(\hat{\boldsymbol{\theta}}) - \mathbf{q}(\boldsymbol{\theta}^*) = (k - k^*) \mathbf{a}. \quad (9)$$

If so, then either there are two parameter vectors that generate the same age profile (so $\mathbf{q}(\hat{\boldsymbol{\theta}}) - \mathbf{q}(\boldsymbol{\theta}^*) = \mathbf{0}$) 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 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.

The standard method may 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 from (??). (Analogous results apply if one uses (??).) The standard method estimates the structural parameters by

$$\tilde{\boldsymbol{\theta}} = \arg \min_{\boldsymbol{\theta}} [\mathbf{q}(\boldsymbol{\theta}) - \tilde{\boldsymbol{\alpha}}]' [\mathbf{q}(\boldsymbol{\theta}) - \tilde{\boldsymbol{\alpha}}]. \quad (10)$$

Equation (??) effectively imposes the normalization that the estimated cohort effects are orthogonal to a linear trend. This normalization is like choosing k in (??) so that $\gamma_c + kc$ is orthogonal to a linear trend; in other words,

$$k = k^{norm} \equiv -\frac{\sum_c \gamma_c}{\sum_c c^2}. \quad (11)$$

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

$$\tilde{\alpha} = \mathbf{q}(\boldsymbol{\theta}^*) + k^{norm} \mathbf{a}. \quad (12)$$

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

$$\tilde{\boldsymbol{\theta}} = \arg \min_{\boldsymbol{\theta}} [\mathbf{q}(\boldsymbol{\theta}) - \mathbf{q}(\boldsymbol{\theta}^*) - k^{norm} \mathbf{a}]' [\mathbf{q}(\boldsymbol{\theta}) - \mathbf{q}(\boldsymbol{\theta}^*) - k^{norm} \mathbf{a}]. \quad (13)$$

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

3. A simple 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 j and lives for $A + 1$ periods, $t = j, j + 1, \dots, j + A$. Let $c_{i,a,t}$ be i 's consumption in year t , when he is age $a = t - j$. The agent's preferences are

represented by

$$-\frac{1}{2}\mathbb{E}_j \sum_{a=0}^A \rho^a [\bar{c} - c_{i,a,t}]^2 \quad (14)$$

The agent begins life with assets $x_{i,0,j} > 0$ and receives a stochastic income $y_{i,a,t}$ in each period $t = j, j+1, \dots, j+A$. Income is independently and identically distributed across agents and dates with mean μ and variance σ^2 . The agents 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 for $a < A$ is

$$x_{i,a+1,t+1} = (1+r)(x_{i,a,t} + y_{i,a,t} - c_{i,a,t}), \quad a = 0, \dots, A-1. \quad (15)$$

The agent maximizes (??) by choice of $\{c_{i,a,t}, x_{i,a+1,t+1}\}_{a=0}^A$, subject to (??) and

$$c_{i,A,j+A} = x_{i,A,j+A} + y_{i,A,j+A}, \quad (16)$$

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

$$c_{i,a,j+a} = (1 + \phi_a)^{-1}(x_{i,a,j+a} + y_{i,a,j+a} + \mu\phi_a) \quad (17)$$

where

$$\phi_a = \sum_{s=1}^{A-a} \rho^s = \rho \frac{1 - \rho^{A-a}}{1 - \rho}. \quad (18)$$

It can also be shown that

$$c_{i,a+1,j+a+1} - c_{i,a,j+a} = (1 + \phi_{a+1})^{-1}(y_{i,a+1,j+a+1} - \mu) \quad (19)$$

It follows from (??) that

$$\text{Var}[c_{i,a,j+a}|a, j] = (1 + \phi_0)^{-2} \text{Var}[x_{i,0,j}] + \sigma^2 \sum_{s=0}^a (1 + \phi_s)^2. \quad (20)$$

B. Identification

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

Suppose that 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

$$\hat{c}_{i,a,t} = c_{i,a,t} + \epsilon_{i,a,t} \quad (21)$$

where the measurement error $\epsilon_{i,a,t}$ is independent of $c_{i,a,t}$, uncorrelated across agents, and has mean $\xi_{a,t}$ and variance η_t^2 at date t . (The bias $\xi_{a,t}$ 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 $\xi_{a,t}$ 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 $\xi_{a,t}$. The variance of observed consumption among people who are age a at date t is

$$\text{Var}[\hat{c}_{i,a,t}|a, t] = \eta_t^2 + \text{Var}[c_{i,a,t}|a, t] = \eta_t^2 + (1 + \phi_0)^{-2} \text{Var}[x_{i,0,j}] + \sigma^2 \sum_{s=0}^a (1 + \phi_s)^2. \quad (22)$$

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

$$\sigma^2 \sum_{s=0}^a \left(1 + \rho \frac{1 - \rho^{A-s}}{1 - \rho}\right)^2 = ka + \hat{\sigma}^2 \sum_{s=0}^a \left(1 + \hat{\rho} \frac{1 - \hat{\rho}^{A-s}}{1 - \hat{\rho}}\right)^2, \quad a = 0, \dots, A. \quad (23)$$

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

$$\sigma^2 \left(1 + \rho \frac{1 - \rho^A}{1 - \rho}\right)^2 = \hat{\sigma}^2 \left(1 + \hat{\rho} \frac{1 - \hat{\rho}^A}{1 - \hat{\rho}}\right)^2 \quad (24a)$$

$$\sigma^2 \sum_{s=0}^1 \left(1 + \rho \frac{1 - \rho^{A-s}}{1 - \rho}\right)^2 = k + \hat{\sigma}^2 \sum_{s=0}^1 \left(1 + \hat{\rho} \frac{1 - \hat{\rho}^{A-s}}{1 - \hat{\rho}}\right)^2 \quad (24b)$$

$$\sigma^2 \sum_{s=0}^2 \left(1 + \rho \frac{1 - \rho^{A-s}}{1 - \rho}\right)^2 = 2k + \hat{\sigma}^2 \sum_{s=0}^2 \left(1 + \hat{\rho} \frac{1 - \hat{\rho}^{A-s}}{1 - \hat{\rho}}\right)^2 \quad (24c)$$

Using (??) to substitute for $\hat{\sigma}^2$ in (??) and (??) and simplifying, we have

$$\sigma^2 \left(1 + \rho \frac{1 - \rho^{A-1}}{1 - \rho}\right)^2 = k + \sigma^2 \frac{\left(1 + \rho \frac{1 - \rho^A}{1 - \rho}\right)^2}{\left(1 + \hat{\rho} \frac{1 - \hat{\rho}^A}{1 - \hat{\rho}}\right)^2} \left(1 + \hat{\rho} \frac{1 - \hat{\rho}^{A-1}}{1 - \hat{\rho}}\right)^2 \quad (25a)$$

$$\sigma^2 \sum_{s=1}^2 \left(1 + \rho \frac{1 - \rho^{A-s}}{1 - \rho}\right)^2 = 2k + \sigma^2 \frac{\left(1 + \rho \frac{1 - \rho^A}{1 - \rho}\right)^2}{\left(1 + \hat{\rho} \frac{1 - \hat{\rho}^A}{1 - \hat{\rho}}\right)^2} \sum_{s=1}^2 \left(1 + \hat{\rho} \frac{1 - \hat{\rho}^{A-s}}{1 - \hat{\rho}}\right)^2 \quad (25b)$$

Using (??) to eliminate k and simplifying, we obtain

$$\left(\frac{1-\rho^{A-1}}{1-\rho^{A+1}}\right)^2 - \left(\frac{1-\rho^A}{1-\rho^{A+1}}\right)^2 = \left(\frac{1-\hat{\rho}^{A-1}}{1-\hat{\rho}^{A+1}}\right)^2 - \left(\frac{1-\hat{\rho}^A}{1-\hat{\rho}^{A+1}}\right)^2 \quad (26)$$

For $\hat{\rho} \in (0, 1)$, the right-hand side of (??) is monotonically decreasing in $\hat{\rho}$; therefore, (??) has a unique solution, which is $\hat{\rho} = \rho$. We then obtain $k = 0$ from (??) and $\hat{\sigma}^2 = \sigma^2$ from (??). 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 effects happen to be orthogonal to a linear trend.

4. 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 parameters from their effect on the curvature of the age profile, rather than on its slope. If the curvature is not precisely estimated

or if parameters have only weak effects on the curvature, then confidence intervals for the structural parameters will be large. 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. First, (??) assumes that time effects have the same impact on people of all ages. Second, both (??) and the non-parametric version (??) assume that time effects matter only contemporaneously. ? argue that many important economic 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 ask how best to estimate structural parameters using it. Analysis of more complex models is left for future research.

References

- Ameriks, John, and Stephen P. Zeldes, 2004, “How Do Household Portfolio Shares Vary With Age?”, manuscript, Columbia University.
- Angrist, Joshua D., and Jorn-Steffen Pischke, 2009, *Mostly Harmless Econometrics: An Empiricist’s Companion*, Princeton, N.J.: Princeton University Press.
- Deaton, Angus, and Christina Paxson, 1994, “Intertemporal Choice and Inequality,” *Journal of Political Economy* 102(3), 437–467.
- 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.
- Huggett, Mark, Gustavo Ventura, and Amir Yaron, 2011, “Sources of Lifetime Inequality,” *American Economic Review* 101, 2923–2954.
- Gourinchas, Pierre-Olivier, and Jonathan A. Parker, 2002, “Consumption Over the Life Cycle,” *Econometrica* 70(1), 47–89.
- Kaplan, Greg, 2011, “Inequality and the Lifecycle,” manuscript, University of Pennsylvania.
- 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,” Federal Reserve Bank of Minneapolis Staff Report 461.
- Wachter, Jessica A., and Motohiro Yogo, 2010, “Why Do Household Portfolio Shares Rise in Wealth?,” *Review of Financial Studies* 23(11): 3929–3965.