Published in final edited form as:

Labour Econ. 2010 June ; 17(3): 592–597. doi:10.1016/j.labeco.2009.04.009.

Estimating Intergenerational Persistence of Lifetime Earnings with Life Course Matching: Evidence from PSID

Elena Gouskova¹, Ngina Chiteji², and Frank Stafford³

¹Economic Behavior Program, Institute for Social Research, University of Michigan, Ann Arbor, MI, 48106, <egouskov@umich.edu>, PH: 734-936-0307; FX: 734-936-3809

²Department of Economics, Skidmore College, <nchiteji@skidmore.edu>

³Frank Stafford, Department of Economics, University of Michigan and Institute for Social Research, University of Michigan, <fstaffor@isr.umich.edu>

Abstract

Why do estimates of the intergenerational persistence in earnings vary so much for the United States? Recent research suggests that life-cycle bias may be a major factor (Haider and Solon 2006; Grawe 2006). In this paper we estimate the intergenerational correlation in lifetime earnings by using sons' and fathers' earnings at similar ages in order to account for lifecycle bias. Our estimate based on earnings measured at 35–44 for both fathers and sons is similar to that for the age range 45–54.

Keywords

Intergenerational Earning Mobility; Income Mobility; Life-Cycle Bias

Introduction

Empirical studies of intergenerational earnings persistence have reported substantially varying estimates ranging from 0.11 to 0.58 (for review of the studies see Solon 1999, Grawe 2006).¹ While some of the variation can be attributed to differences in the degree to which the studies correct for “textbook” measurement error in the independent variable (parental income), the recent literature points to an important role of life-cycle bias (Jenkins 1987, Grawe 2006, Haider and Solon 2006). For example, Grawe (2006) finds that among studies with otherwise similar methodology, 20 percent of the variance in published estimates of earnings persistence is attributable to cross-study differences in the age of responding fathers, with estimated earning persistence negatively related to age of fathers. Other studies point to a positive association between estimated earnings persistence and son's age at observation². While the importance of life-cycle bias varies by a study's objective, it is an issue when we would like to assess persistence in lifetime income between generations.

Correspondence to: Elena Gouskova.

¹We also have estimated the intergenerational earnings elasticity by using different age intervals for fathers and sons—fathers at age 45–54 with sons age 25–34 for example. The estimates are lower when father's income is measured at older age than son's income, and the estimates are generally higher when father's income is measured at younger age than son's income. The pattern is consistent with the effect of the life-cycle bias. It is also consistent with time effect of ever increasing earnings variance.

¹Also see Corak (2004) for comparisons of estimates across North America and Europe.

²For some references see Grawe (2006).

Paucity of data on lifetime earnings of parents and their children led the majority of early studies of the intergenerational correlation in income to use current income as proxy for lifetime income, while assuming an error-in-variables model. However, as found in Haider and Solon (2006) the error-in-variables assumption is likely to be violated in models with heterogeneous income growth, a feature empirically well documented and theoretically explained by differential investment in human capital (Grawe 2003; Ben-Porath 1967; Ryder, Stafford and Stephan 1976). Since current income approximates lifetime income with age-dependant bias, the estimate of intergenerational lifetime income mobility will be biased as well.

The ideal way to avoid life-cycle bias in estimation of persistence in lifetime earnings is to use the present value of full earning histories of fathers and sons in estimation of intergenerational mobility.³ As such earnings data are rarely available, researchers often use fewer years of data, and face the problem of minimizing the bias in the estimation of lifetime earnings persistence. One method suggested in Grawe (2006) and others is to use observations for fathers and sons that come from similar lifecycle points and from periods where current earnings are likely to be a good approximation for lifetime earnings. As Haider and Solon (2006) show in their analysis of the relationship between current income and lifetime income, current income from early thirties to mid forties generally satisfies the error-in-variables model and, thus, provides an unbiased estimate of lifetime income.

While other research has outlined the reasons that an approach that corrects for life-cycle bias should be taken, this paper implements the suggested approach. More specifically, we use data from the Panel Study of Income Dynamics (PSID) and employ the life course matching of fathers and sons to obtain estimates of the intergenerational elasticity in lifetime earnings. The long history of PSID, which started in 1968, allows us to compare fathers' and son's incomes at similar life cycle points over much of the earnings histories. Furthermore, the range of ages feasible for matching includes the mid-life period often considered to be best for approximation of lifetime income (Mincer 1974, Haider and Solon 2006).⁴

The rest of the paper is organized as follows. Section II presents a theoretical model to illustrate the problems life cycle bias poses for analyses. Section III discusses our data and the construction of our dataset. Section IV presents our results. Section V concludes.

Theoretical background

This section follows the exposition in Haider and Solon (2006) to explain the nature of life-cycle bias and the approach we use to minimize it. Suppose we are interested in estimating intergenerational persistence in lifetime earnings using the regression-to-mean model

$$y^p = \beta x^p + e, \quad (1)$$

Where y^p and x^p represent logs of child's and parent's lifetime earnings, which are the present discounted values of earnings over the life span. Since lifetime earnings are generally not observed, they often have been approximated by current earnings. Haider and Solon (2006) show that with heterogeneous income growth current earnings relate to lifetime earnings as follows⁵:

³Dahl and DeLeire (2008) estimate the intergenerational earnings elasticity using career-long histories of earnings for fathers and earnings at age around 37 for sons.

⁴A similar issue arises in studying other intergenerational relationships, such as obesity. Those who are obese have the most obvious manifestation in the age range, 40–50.

⁵Income variables expressed as deviations from their population means.

$$y_t = \lambda_t^y y^p + \varepsilon_t \text{ and } x_s = \lambda_s^x x^p + v_s, \quad (2)$$

where y^t and x^s are the logs of son's and parent's current income measured at ages t and s respectively and ε_t and v_t are idiosyncratic error terms.

Non-parametric analysis of Social Security records in Haider and Solon (2006) shows the estimated lambda to exhibit a skewed hump shaped pattern over the life cycle. It increases strongly between early twenties and early thirties from about .2 to 1. Then it stays around 1 between early thirties and late forties. And it is approximately flat after the late forties with values about .8. Böhlmark and Lindquist (2006) find a similar pattern in the relationship between current and lifetime income in Swedish data.

The probability limit of $\hat{\beta}$, the regression coefficient in the regression of son's current income on current parental income, is

$$\text{plim} \hat{\beta} = \frac{\text{cov}(y_t, x_s)}{\text{var}(x_s)} = \beta \left[\frac{\lambda_t^y \lambda_s^x \text{var}(y^p)}{(\lambda_s^x)^2 \text{var}(x^p) + \text{var}(v_s)} \right] = \beta \frac{\lambda_t^y}{\lambda_s^x} \frac{1}{1 + \frac{\text{var}(v_s)}{(\lambda_s^x)^2 \text{var}(x^p)}}. \quad (3)$$

In the last part of the above expression, the ratio of λ_t^y and λ_s^x represents the life-cycle bias, and the following term is the attenuation bias⁶. The combined effect of two biases can be to either overstate or understate the elasticity β . Measuring son's income at younger ages, when λ_t^y is less than 1, will result in a lower estimate of intergenerational elasticity. The effect of father's age is less clear. Measuring father's income when the father is young affects lifecycle bias and attenuation bias in different directions. Based on a number of studies Grawe (2006) finds that, in practice, intergenerational income elasticity estimates decrease with father's age at observation.

One approach to minimize the effect of life cycle bias is to use parental and sons' incomes measured at ages where the lambda values are closest. Assuming the relationship between current and lifetime income remains approximately the same for both generations (i.e.

$\lambda_t^y \approx \lambda_s^x$ when $t = s$) this approach is equivalent to using parental and sons' measurements at similar ages. Measurements near mid-life for both father and son, when $\lambda_t^y = \lambda_s^x = 1$, are likely to be the best choice for estimation of intergenerational lifetime income elasticity.

Data and methods

For the estimation we used three age groups from the PSID: 25 to 34, 35 to 44, and 45 to 54. For each of the age groups the dataset is constructed following the same procedure⁷. We start with the sample of all heads males who have ever participated in the PSID, 7938 cases, and restrict it to those whose father also has been a PSID head, 1973 observations.⁸ Then, for the son an income observation from the specified age range is selected at random, as generally there is more than one valid observation available during the 10-year age interval being observed⁹. For fathers in the same age range 5 observations are selected at random

⁶Notice that the expression for attenuation bias in (3) is age dependent; when lambda is 1, the expression becomes familiar expression for attenuation bias. Small λ_s^x will reduce estimate of intergenerational persistence, i.e. will increase attenuation bias.

⁷From wave to wave the reported age does not always increase by the same number of years between surveys. For example, the difference in reported age between 1994 and 1993 surveys is not always 1. This is due to timing of interviews and measurement error. Instead of relying on the reported age variable we use the BIRTH YEAR variable, which is constructed as the median among differences of survey year and reported age.

⁸We use data from the SRC sample only for this paper. This makes the work comparable to Solon (1992)

and averaged to produce an estimate of parental lifetime income. Averaging is a well-known method to reduce attenuation bias resulting from transitory income components and measurement error (Solon 1992).¹⁰ To use available data to fuller extent and to obtain robust estimates we create 100 datasets for each age range by random drawing income observations for fathers and sons. Table 1 reports summary statistics for a typical sample for each of the age groups.

As Table 1 shows the number of father-son pairs decreases with age – 643 in the 25 to 34 age group, to 535 in the 35–44 age group, and to 379 in the 45 to 54 age group. One interesting outcome of our life-cycle matching procedure is that for all three groups, observations originate in approximately the same time periods: the sons' data center around 2000–2002 years, and dads' data mostly come from the mid seventies. This “time fixing” allows us to compare earnings mobility between the three age groups, which turned out also to represent overlapping consecutive cohorts. Table 1 indicates that our age matching is not exact---in all of the three samples dads are about two years older at the time of observation than sons.

Empirical results

Figure 1 shows side-by-side box plots of estimates of lifetime earnings elasticity obtained by regressing son's log of labor income on dad's average of logs of labor incomes. The box plot for each age group is based on estimates obtained from 100 random draws of datasets. Table 2 reports regression results corresponding to the average elasticity estimate for each age group. The estimates are 0.29, 0.41, and 0.42 for 25–34, 35–44, and 45–54 age groups, respectively¹¹. The results suggest that matching son and father ages of observation is helpful in controlling life-cycle bias. In particular, the estimates for the two older groups are close to one another.¹² These estimates are also preferred since current earnings measured at age of 35–44 are likely to be a good proxy for lifecycle earnings assuming lambda to be close to 1.

The estimate of 0.29 in the youngest age group is somewhat lower and statistically different from estimates in two older groups. A number of factors might contribute to this result. First of all, measuring fathers' income at relatively young age is likely to increase the age dependent attenuation bias, the last factor in (3). Second, our assumption that lambdas are a time invariant function of age may be too strong. In particular, flatter earnings profiles at age 25–34 among sons, due to longer period of time at school, for example, may result in lifecycle bias, causing the ratio of lambdas in (3) to be less than 1 and thus resulting in a lower estimate of the intergenerational lifetime income elasticity. Also, the results might indicate a structural change in earning mobility with younger cohorts exhibiting higher mobility, possibly the result of wider access to college education.¹³ Finally, sample selection may play a role. The two older samples, 35–44 and 45–54, are smaller in size and

⁹All nominal values are adjusted for inflation using CPI-U index. The procedure used to effect the matching is the Family Identification Mapping System (FIMS) as explained in the on-line tutorial, <http://psidonline.isr.umich.edu/Guide/tutorials/7/tutorial7.pdf>

¹⁰Like Solon we use 5-year averages, although recent work by Mazumder (2005) suggests that using even more years can be helpful.

¹¹We have also explored how different specifications of income variables affect elasticity estimates. Our results show that when log of average income is used then estimates are slightly higher compared to the results reported in the paper. Opposite is true when zero valued income observations are allowed in the 'log of average' specification. Results are available from authors upon request.

¹²These estimates are within the range of existing estimates for the U.S., and compare favorably to the 0.41 intergenerational income elasticity obtained in Solon (1992).

¹³Our analysis does not delve into any specific mechanism to support an elasticity of .41–.42. We did explore separately the elasticity of labor hours between sons and fathers and find an elasticity of .08, suggesting that most of the earnings carryover across generations is from the correlation of hourly wage rates.

may also be more constrained in types of father-son pairs resulting in higher elasticity estimates.

Discussion

Using PSID data, this paper uses life course matching of fathers and sons in order to estimate 'the' intergenerational lifetime income elasticity. Previous research has suggested that the estimates in the existing literature are likely to be subject to life cycle bias. We use life course matching to resolve this problem. If the relationship between current and lifetime incomes changes little between generations, employing fathers and sons' measurements at close ages should reduce life cycle bias in estimates of intergenerational income elasticity. Our estimates of income elasticity using measurements from 35–44 and 45–54 age ranges are similar and equal to 0.41 and 0.42 correspondingly. Because current income at ages 35–44 is likely to approximate the lifetime earnings better than current income at other ages according to existing research, 0.41 is our preferred estimate of the intergenerational lifetime income elasticity. The similarity of the estimate of elasticity that one obtains when using measurements at ages 45–54 suggests that for these ages matching also helps to reduce lifecycle bias.

Our estimate of income elasticity obtained using earnings when both father and son are in 25–34 age range is lower than other estimates and equal to .29. This result may suggest that while matching observations at these ages reduces lifecycle bias, the estimate might be affected by increased attenuation bias. The result might also point to structural changes in income elasticity between cohorts.

The approach that we have applied in this paper demonstrates that life course matching is a feasible alternative for researchers working with datasets that do not provide full earning histories. This suggests that panels shorter than the PSID may provide sufficient information to estimate the intergenerational lifetime earnings elasticity accurately. In addition, life course matching may be useful in application to other intergenerational relations, such as obesity or wealth holdings, which also have a distinct life cycle pattern and exhibit heterogeneous growth rates.

References

- Ben-Porath, Yoram. The Production of Human Capital and the Life Cycle of Earnings. *Journal of Political Economy*. 1967; Volume 75:352–365.
- Böhlmark, Anders; Lindquist, Matthew J. Life-Cycle Variations in the Association between Current and Lifetime Income: Replication and Extension for Sweden. *Journal of Labor Economics*. 2006; Vol. 24(n.4)
- Corak, Miles, editor. *Generational Income Mobility in North America and Europe*. Cambridge: Cambridge University Press; 2004.
- Dahl, Molly; DeLeire, Thomas. The Association between Children's Earnings and Fathers' Lifetime earnings: Estimates Using Administrative data. Institute for Research on Poverty. 2008 discussion paper. No. 1342-08.
- Davis M, McGonagle KA, Stafford FP, Schoeni RF. Relative contributions of parental and grandparental obesity to risk of child obesity: Implications for primary care practice. *Journal of the American Board of Family Medicine*. 2008 in press.
- Grawe, Nathan D. Lifecycle bias in estimates of intergenerational earnings persistence. *Labour Economics*. 2006; Vol.13:551–570.
- Grawe, Nathan D. Life cycle bias in the estimation of intergenerational earnings persistence. *Statistics Canada. Analytical Studies Branch Research Paper Series*, No. 207, August. 2003
- Haider, Steven J.; Solon, Gary. Life-Cycle Variation in the Association between Current and Lifetime Earnings. *American Economic Review*. 2006; 96(4):1308–1320.

- Jenkins, Stephen. Snapshots Versus Movies: 'Lifecycle Biases' and the estimation of Intergenerational Earnings Inheritance. *European Economic Review*. 1987; Vol. 31:1149–1158.
- Mazumder, Bhashkar. Fortunate Sons: New Estimates of Intergenerational Mobility in the United States Using Social Security Earnings Data. *Review of Economics and Statistics*. 2005; Volume 87(2):235–255.
- Mincer, Jacob. *Schooling Experience and Earnings*. NY: National Bureau of Economic Research; 1974.
- Ryder, Harl E.; Stafford, Frank P.; Stephan, Paula E. Labor, Leisure and Training over the Life Cycle. *International Economic Review*. 1976; Vol. 17(No. 3):651–674.
- Solon, Gary. Intergenerational Income Mobility in the Labor Market. *Handbook of Labor Economics*. 1999; Vol.3:1761–1800.
- Solon, Gary. Intergenerational Income Mobility in the United States. *American Economic Review*. 1992; Volume 82(3):393–408.

Figure 1.

Note: For each age group box plot is based on 100 estimates of intergenerational lifetime earning elasticity.

Table 1

Summary statistics

	Age group		
	25–34	35–44	45–54
N	643	535	379
Mean year of son's observations	2000	2000	2002
Mean year of father's observations	1977	1975	1976
Mean age of son	28	38	48
Mean age of father	30	40	50
Ranges for birth year of son	1956–1979	1948–1969	1942–1959
Ranges for birth year of father	1937–1961	1927–1950	1917–1938
Mean birth year of son	1971	1961	1953
Mean birth year of father	1946	1934	1925
Mean son's log of labor inc	10.7	10.8	10.8
Mean father's log of labor inc	10.5	10.8	10.9
Std son's log of labor inc	0.72	0.8	0.83
Std father's log of labor inc	0.49	0.56	0.64

Table 2

Regression of son's log of labor income

	Age group		
	25–34	35–44	45–54
Intercept	7.42 (0.609)	6.29 (0.645)	6.31 (0.688)
5-year average of father's log of labor income	0.29 (0.057)	0.41 (0.059)	0.42 (0.063)
N	643	535	379
R-squared	0.04	0.08	0.11

Note: standard errors are shown in parentheses.