DISCUSSION PAPER SERIES

IZA DP No. 5990

Following in Your Father's Footsteps: A Note on the Intergenerational Transmission of Income between Twin Fathers and their Sons

Vikesh Amin
Petter Lundborg
Dan-Olof Rooth

September 2011

Forschungsinstitut
zur Zukunft der Arbeit
Institute for the Study
of Labor

Following in Your Father's Footsteps: A Note on the Intergenerational Transmission of Income between Twin Fathers and their Sons

Vikesh Amin
Binghamton University (SUNY)

Petter Lundborg
Lund University, VU University Amsterdam, Tinbergen Institute, HEP, CED and IZA

Dan-Olof Rooth
Linnaeus University, Lund University, CED, CReAM and IZA

Discussion Paper No. 5990
September 2011

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-180
E-mail: iza@iza.org

IZA Discussion Paper No. 5990
September 2011

ABSTRACT

Following in Your Father's Footsteps: A Note on the Intergenerational Transmission of Income between Twin Fathers and their Sons*

We provide the first twin-based estimates of the intergenerational transmission of income between fathers and sons. Using Swedish register data on the income of monozygotic twin fathers and their sons, we are able to control for unobserved endowments at the twin-pair level when estimating the intergenerational relationship. We find a cross-sectional intergenerational income elasticity of 0.276, while our twin-based intergenerational income elasticity is 0.12. This is close to the estimate of 0.10 found by Björklund et al. (2006) using an adoption design. This suggests that at most half of the income transmission can be given a causal interpretation.

JEL Classification: J0, J1

Keywords: twins, income, intergenerational transmission, intergenerational mobility

Corresponding author:

Dan-Olof Rooth
Linnaeus School of Business and Economics
Linnaeus University
Kalmar Nyckel
SE-39182 Kalmar
Sweden
E-mail: dan-olof.rooth@lnu.se

* A research grant from the Centre for Economic Demography (CED) at Lund University is gratefully acknowledged.

1. Introduction

There are strong intergenerational associations between parents' and children's outcomes, such as schooling and income. These associations do not necessarily reflect causal relationships because of unobserved genetic and child-rearing parental endowments that affect both parents' and children's outcomes. In order to control for such unobserved endowments, two common used identification strategies have been the adoption design and the within-twins (or equivalently twins fixed effects) approach. Recent studies using these methods have established that at least part of the association between parents' schooling and children's schooling appears to have a causal interpretation (see Holmlund et al. 2010).

In contrast, the parallel literature on the intergenerational transmission of income has mainly used the adoption design. This approach relates income of adopted children to income of their adoptive parents. Since adoptees share their rearing parents' environment but not their genes, this approach controls for genetic endowments. Estimates from the adoption approach could still be upward biased if (1) there is selective placement of children to adoptive families (e.g., high ability children being placed in high ability families) and if (2) child-rearing skills, that may directly affect the income of the adopted child, are positively related to the adoptive parent's income.[1] In a novel study, Björklund et al. (2006) relate income of adopted sons to income of their adoptive fathers and their biological fathers (to control for the impact of selective placements) in Sweden. They found the income elasticity for biological and adoptive fathers to be 0.05 and 0.10, respectively, and only the estimate for adoptive fathers was statistically significant, indicating that nurture is more important than nature.[2]

We contribute to the literature by using the within-twins approach to estimate the intergenerational transmission of income between fathers and sons in Sweden. To our knowledge, this identification strategy has not been used before to estimate the

[1] The adoption design also assumes that all children are adopted at birth. If children are placed in institutional care before being placed with adoptive parents, this may lead to an underestimation of causal intergenerational effect.

[2] We are only aware of two additional studies that have used the adoption design to estimate the intergenerational association in income. Liu and Zeng (2009), find little evidence that adoptive parent income matters in the U.S. Sacerdote (2007) finds a log income elasticity for adopted children of 0.19 in the U.S., but it is not statistically significant.

intergenerational transmission of income. Using data on income for 615 DZ (dizygotic, fraternal) and 393 MZ (monozygotic, identical) twin father pairs and their sons, we find cross-sectional associations of 0.20 and 0.27 for DZ and MZ twin fathers, which is similar to previous estimates for Sweden. When we take income differences between sons who are cousins, and relate them to income differences between fathers who are twins, which eliminates much of the influences of unobserved endowments common to both twin fathers, our estimates fall to 0.17 and 0.12 for DZ and MZ twin fathers respectively.[3] The latter estimate is very close to the estimate obtained by Björklund et al. (2006), using the adoption design. Hence, if one believes that using differences in income for MZ twin pairs eliminates much of the influence of unobserved endowments, we arrive at a similar conclusion that at most half of the income transmission has a causal interpretation.

2. Methodology

Consider a reduced-form intergenerational transmission model where log income of son i born to father j (Y^s_{ij}) is related to his father's log income (Y^f_j), unobserved genetic endowments of the father (h^f_j), unobserved child-rearing endowments of the father (f^f_j) and a son-specific stochastic term (ε^s_{ij})

$$Y^s_{ij} = \beta^f Y^f_j + h^f_j + f^f_j + \varepsilon^c_{ij} \quad (1)$$

An OLS regression of relation (1) provides an estimate of the intergenerational association between father's and son's income, which is a biased estimate of the causal intergenerational effect, since father's income is related to the unobserved genetic and child-rearing endowments that also directly affect son's income. The influence of these unobserved endowments can be controlled for in the within-twins approach, which relates income differences between sons who are cousins to income differences between fathers who are twins and to differences in paternal unobserved genetic and child-rearing endowments. This can be written as:

$$\Delta Y^s_j = \beta^f \Delta Y^f_j + \Delta h^f_j + \Delta f^f_j + \Delta \varepsilon^s_j \quad (2)$$

[3] Like the adoption design, the within-twins approach also has limitations, which prevents identification of causal intergeneration effects. We discuss these limitations in section 2.

Relation (2) can be estimated separately for MZ and DZ twin fathers and their sons. MZ twins are genetically identical, so that unobserved genetic and child-rearing endowments in relation (2) will be differenced out. DZ twins are like non-twin siblings, sharing only half of their genetic endowments, which means unobserved endowments will not be fully controlled for. One would therefore expect within-twins estimates for DZ twin fathers to lie between cross-sectional and within-twins estimates for MZ twin fathers.

Caveats of the within-twins methodology

There are two inherent limitations with the within-twins approach that prevent identification of causal estimates. First, one has to assume that income differences between twin fathers are exogenous (i.e. due to factors that do not directly affect their son's income). This assumption may not always be true. For example, Behrman and Rosenzweig (2004) find that MZ twins with higher birth weight have both more schooling and higher wages. If birth weight differences within MZ twin fathers leads to income differences and positively affect son's income, then our within-twins estimates will be biased upwards. Second, if there is classical measurement error in income, then first differencing exacerbates the downward bias towards zero. As described in the next section, our data comes from administrative sources and not from self-reports. Our measure of fathers' income is also based on several years of income data in order to better proxy lifetime income. We believe that measurement error is less of a problem compared to the exogenous assumption of income differences, so that our estimates are likely to reflect upper bounds.

3. Data

Our empirical analysis is based on a data set constructed by integrating registers from Statistics Sweden (SCB) and the Swedish Twin Registry. The former contains information on the Swedish population being 16-64 years old in 1999. Our data set allow us to link fathers to sons and also includes information on whether the father is a twin. Moreover, we are able to link twin fathers to each other. The data set gives us access to information on twin fathers born 1925-1945 and their sons born 1953-1973. Zygosity has been determined for 82 percent of the individuals in our dataset, based

on survey questions regarding co-twin similarity.[4] Our dataset includes 615 and 393 twin father pairs being classified as DZ and MZ twins, respectively. Our labor income measures are taken from the tax registers where, for son's income, we use labor income in 2003 as defined by Statistics Sweden, and for fathers we use information on labor income for the years 1970, 1975 and 1980.[5] 94 percent of the sons have positive income in 2003 and we top code their income at the 97th percentile.

Given the included birth cohorts, we arrive at analyzing sons that are 30 to 50 years old in 2003, while fathers are between the ages of 25 and 55, depending on the year their income is measured. In the literature on the intergenerational transmission in income, it is often stressed that one should aggregate several income years of fathers, in order to allow for persistent transitory shocks in income.[6] In our data, 89 percent of the fathers have positive income in all the three years considered, which is the data we use in our main specification.[7] We also perform analyses where we relax this restriction and also include twin pairs with less yearly income information, as long as the income information overlap within a twin pair (named *Summed income*). For example, if one twin within a twin pair has income information for all three years while the other has it only for one year, say 1970, only the 1970 income information

[4] The method used has been found to classify twins with an accuracy of 95 percent or more, see Lichtenstein et al. (2002).

[5] The income of fathers and sons is defined in a similar way by Statistics Sweden. Both measures include benefits such as parental leave, unemployment and sickness benefits. Fathers' income is measured in 1980 prices.

[6] One ideally want to estimate $y_i^s=\beta y_i^f + \varepsilon_i$ where y_i^s and y_i^f is the log of lifetime income of sons and fathers respectively. If son's current income at age t (y_{it}^s) is issued as an imperfect proxy for lifetime time and assume that $y_{it}^s=\lambda_t y_i^s + u_{it}$, (where λ_t is the association between current and lifetime income) the OLS estimate of the intergenerational income elasticity is $\lambda_t\beta$.Böhlmark and Lindquist (2006) estimate λ_t for Swedish men by regressing current income on lifetime income. They find λ_t to approximately equal 1 after age 30. If on the other hand, father's current income is used to proxy lifetime income ($y_{it}^f=\lambda_t y_i^f + u_{it}$) and there is an approaite measure of son's lifetime income, then the OLS estimate is $\theta_t\beta$, where θ_t equals $\lambda_t var(y_i^f)/\lambda_t^2 var(y_i^f) + var(u_{it})$. Böhlmark and Lindquist estimate θ_t by regressing lifetime income on current income. They find that single year measures of father's income lead to an attenuation bias of about 50 percent irrespective of the age at which income is measured. As we measure income at 3 points in time stretched over a 10 year period, this is expected to reduce the attenuation bias.

[7] This restriction implies using a somewhat smaller number of DZ and MZ twins, 548 and 367 pairs, respectively.

is used when "calculating" the within-twin difference in father's income.[8] We exclude cases in which the income difference within pairs is zero since these twin pairs do not attribute to identification of the intergenerational transmission coefficient in the within-twins estimation.

Description of selection: We start out by selecting men who are cousins and for which we have information on both their fathers who are twins. Restricting the data to fathers and sons who are born in Sweden, who meet the cohort restrictions described above, and for which zygosity has been determined, we arrive at 3,183 father-son pairs. When only including sons having positive incomes in 2003, this decreases to 3,002 father-son pairs. These 3,002 sons are divided among 1,046 twin pairs of which 4 pairs do not differ in income and another 34 pairs show extreme differences in income (>100 percent difference). These 38 pairs are dropped, leaving us with 1,008 twin pairs and 2,772 sons, which is the data that we use for our analyses.

4. Results

Figure 1 displays within-twin percentage differences in log incomes for MZ and DZ twin fathers. From the figure, it is clear that that there is much more variation in income differences for DZ twin fathers than for MZ twin fathers. The brother correlations in income are 0.63 and 0.73 for DZ and MZ twin fathers, respectively.[9] These correlations appear rather high since modern sibling/brother correlations in income are found to be in the order of 0.2 for Sweden and around 0.4 for MZ twins. However, Björklund et al. (2009) found the brother correlation in income to be as high as 0.49 for the 1932-38 birth cohorts, which decreased to 0.36 for the 1938-44

[8] The inclusion of these cases is expected to exacerbate measurement error but to increase precision. In an unreported sensitivity analysis, we used single yearly observations for fathers' income. Both OLS and within-twins estimates were smaller than those obtained when using several yearly observations for father's income. Results are available upon request.

[9] Somewhat surprisingly, the brother correlations in years of schooling are lower, 0.41 and 0.51 for DZ and MZ twin fathers, respectively. Interestingly, these cohorts thus provide a situation in which incomes within siblings are in fact more similar than those found for schooling.

birth cohorts. Hence, there appears to be a decrease in brother correlations in income just around the birth cohorts that we use, that is, the 1925-45 birth cohorts.[10]

Table 1 provides summary statistics for the sample of MZ and DZ twin parents separately. The distribution of birth cohorts (for both twin fathers and sons), as well as labor income for both sons and fathers are very similar across the DZ and MZ samples.

Table 2 presents our results, showing cross-sectional OLS and within-twins estimates of intergenerational elasticities for DZ and MZ twin fathers and their sons. Row 1 makes full use of the data by using father pairs where both have the same income information in at least one year. Row 2 is our preferred specification which only uses father pairs where we have income data in all 3 years 1970, 1975, and 1980. The OLS associations are 0.17 and 0.25 for DZ and MZ twins in row 1 and slightly higher in row 2 at 0.20 and 0.27, respectively. These associations are similar to that of 0.24 found in Björklund et al. (2006).

As expected, the within-twins estimates for DZ and MZ twin fathers in columns 2 and 4 are smaller than the corresponding OLS estimates in columns 1 and 3. Moreover, the within-twins estimates for MZ twin fathers decrease by a larger proportion compared to DZ twin fathers. This is also expected as MZ twin fathers are more similar in their unobserved endowments than DZ twin fathers. The estimated within-twin elasticity for DZ and MZ twin fathers in our preferred specification is 0.17 and 0.12 respectively, and only the estimate for DZ twin fathers is statistically significant. Although the estimate for MZ twin fathers is insignificant (imprecisely estimated due to the smaller sample size), it is very close to the intergenerational income elasticity of 0.10 found by Björklund et al. (2006) using the adoption design. Hence, if one believes that using differences in income within MZ twin pairs eliminates much (but not all) of the influence of unobserved endowments, then at most half of the income transmission has a causal interpretation.

[10] For the 1925-35 birth cohorts we find a correlation for DZ twin father's of 0.67, while the one for MZ twin fathers is 0.76. The corresponding estimates for the 1936-45 birth cohorts is 0.55 for DZ twin fathers and 0.67 for MZ twin fathers. To some extent these differences in correlations across birth cohorts may be due to the fact that the older cohort are on average ten years older when their income is measured.

5. Summary

We provide the first twin-based estimates of the intergenerational transmission of income between fathers and sons. Using the within-twins approach, we are able to control for unobserved endowments at the twin-pair level that may otherwise bias the intergenerational estimates. Based on Swedish MZ twin fathers and their sons, we find the cross-sectional intergenerational elasticity to be 0.27, which is similar to previous estimates for Sweden. Relating income differences between MZ twin fathers to the corresponding differences between their sons, we find that the intergenerational income elasticity decreases to 0.12. This is close to the estimate of 0.10 found by Björklund et al. (2006) using an adoption design and suggests that at most half of the income transmission can be given a causal interpretation.

References

Behrman, Jere and Mark Rosenzweig (2004), "Returns to Birthweight", Review of Economics and Statistics, 86(2): 586-601.

Björklund, Anders, Mikael Lindahl and Erik Plug (2006), "The Origins of Intergenerational Associations: Lessons from Swedish Adoption Data", Quarterly Journal of Economics**,** 121(3): 999-1028.

Björklund, Anders, Markus Jäntti and Matthew J. Lindquist (2009), "Family background and income during the rise of the welfare state: Brother correlations in income for Swedish men born 1932-1968", Journal of Public Economics, 93: 671-680.

Böhlmark, A. och M. Lindquist (2006). "Life-Cycle Variations in the Association between Current and Lifetime Income: Replication and Extension for Sweden", Journal of Labor Economics, 24:879-896.

Holmlund, Helena, Mikael Lindahl and Erik Plug (2010), "The Causal Effect of Parents' Schooling on Children's Schooling: A Comparison of Estimation Methods", Journal of Economic Literature, forthcoming.

Lichtenstein, P., De Faire, U., Floderus, B., Svartengren, M., Svedberg, P., & Pedersen, N. L. (2002). The Swedish 2007-04-27 21 Twin Registry: A unique resource for clinical, epidemiological and genetic studies. Journal of Internal Medicine, 252, 184-205.

Liu, Haoming and Jinli Zeng (2009), "Genetic ability and intergenerational earnings mobility", Journal of Population Economics 22:75-95.

Sacerdote, Bruce (2007), "How large are the effects from changes in family environment? A study of Korean American adoptees", Quarterly Journal of Economics 122: 119-157.

Figure 1: Frequency table of within-twin differences in fathers' log income (percent)

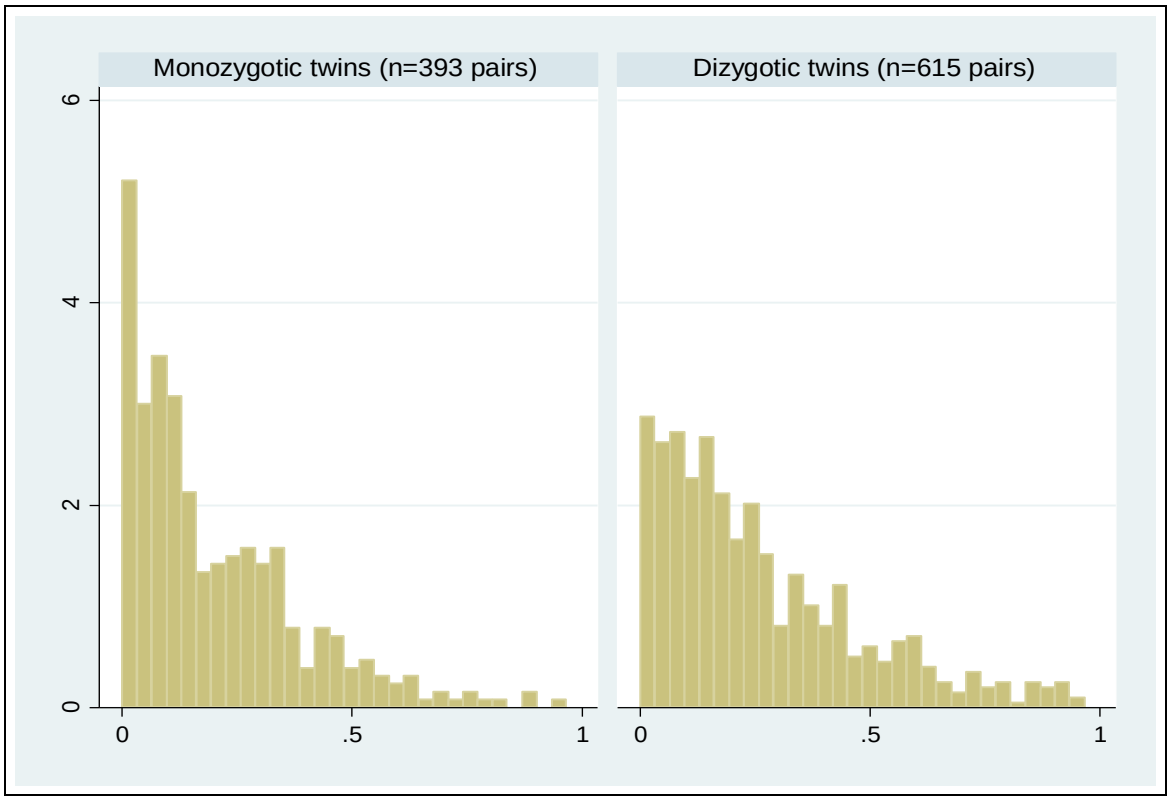

Table 1: Summary Statistics

	DZ (1)	MZ (2)
Sons		
Year of birth:		
1953-1957	11.4	11.1
1958-1962	23.1	24.3
1963-1967	34.2	35.1
1968-1973	31.3	29.5
Log income in 2003	12.4	12.4
	(0.8)	(0.6)
Fathers		
Year of birth:		
1925-1934	43.6	40.2
1935-1945	56.4	59.8
Summed Log income	12.7	12.8
	(0.4)	(0.4)
Log income in 1970, 1975 and 1980	12.8	12.8
	(0.3)	(0.4)

Notes: Standard deviations in parentheses.

Table 2: OLS and Within-Twins Estimates of the Intergenerational Income Elasticity

	DZ Twin Fathers		MZ Twin Fathers	
	OLS (1)	**Within-Twins (2)**	**OLS (3)**	**Within-Twins (4)**
Measure of Fathers' Income				
Fathers' summed income	0.173	0.118	0.247	0.113
	(0.042)***	(0.106)	(0.054)***	(0.147)
N	1,711	[1,711, 615]	1,061	[1,061, 393]
Fathers' income in 1970, 1975 and 1980	0.201	0.167	0.269	0.119
	(0.056)***	(0.103)*	(0.062)***	(0.145)
N	1,526	[1,526, 548]	986	[986, 367]

Notes: Sons are born 1953-1973 with income measured in 2003.Fathers are born 1925-1945All OLS regressions control for son's year of birth fixed effects, fathers' year of birth fixed effects and son's region of residence in 2003 fixed effects. Within-twins regressions control for son's year of birth fixed effects, and son's region of residence in 2003 fixed effects. The first and second number in [.] refers to the number of twins and number of twin pairs with differences in income, respectively. Standard errors in (.). ***significant at 1%, **significant at 5%, *significant at 10%