

Earnings and Income Mobility

Cross-National Estimates of the Intergenerational Mobility in Earnings

By Dean R. Lillard*

Summary

This paper examines the similarity in the association between earnings of sons and fathers in Germany and the United States. It relaxes the log-linear functional form imposed in most studies of the intergenerational earnings association. Theory implies the relationship between earnings of fathers and sons could be nonlinear, especially at the tails of the distribution of earnings of fathers. When a more flexible function form is fit to the data, the apparent similarity between Germany and the United States disappears. Relative to mobility in Germany, upward mobility is higher in the United States for sons with the poorest fathers and downward mobility is lower for sons with fathers with high earnings.

1. Introduction and Background

This paper empirically investigates the association between earnings of sons and fathers in the United States and Germany. It uses, as a starting point, evidence that suggests that the degree of association is similar in both countries. To estimate the strength of the association in earnings across generations, most empirical studies assume that earnings are log-linearly related. This assumption imposes the restriction that, at each point in the distribution of earnings of fathers, a percentage change in fathers' earnings results in the same percentage change in earnings of sons. Theory suggests, however, that the association in earnings across generations is likely to vary across the distribution of fathers' earnings (Becker and Tomes 1986; Mulligan 1997).¹ Motivated by these theoretical implications, this paper relaxes the functional form assumption implicit in log-linear models of the intergenerational earnings association (hereafter IGE association).

It estimates the IGE association with linear, quadratic, cubic and log-linear functional forms.

While most researchers acknowledge the potential nonlinear relationship between sons' and fathers' earnings, few actually include higher order terms in parental earnings.² The inattention to the underlying relationship leaves open the possibility that many estimates of the IGE association are potentially biased.

Perhaps because of differences in specifications and variables used to empirically examine the IGE association, no consensus estimate of the intergenerational earnings association has emerged in the literature. The IGE association is measured alternatively by either the gross or partial correlation between earnings of sons and fathers or by the elasticity of sons' earnings with respect to fathers' earnings. In their 1986 survey of the empirical literature, Becker and Tomes (1986) report estimated elasticities between earnings of sons and fathers ranging from .05 to .43 with a mean of about .20. Such an estimate implies a high degree of mobility. More recent estimates are mixed. For example, Peters (1992), Corak and Heisz (1997), and Couch and Dunn (1997) report estimated associations close to the median estimate reported by Becker and Tomes. Altonji and Dunn (1991) report estimates that span the range. Behrman and Taubman (1990), Solon (1992), Lillard and Reville (1997), and Zimmerman (1992) estimate substantially higher IGE associations. There is some evidence that these higher estimates may arise because of the way in which samples were chosen.³

* Department of Policy Analysis and Management, Cornell University, and DIW.

¹ It should be noted that most mobility studies in sociology do not assume that the association between status of fathers and sons is constant across the distribution of status, but they focus on occupational status rather than income (Featherman and Hauser 1978; Goodman 1979).

² Exceptions include Behrman and Taubman (1990), Corak and Heisz (1997), and Peters (1992). All find evidence of nonlinearities. See also footnote 1.

³ Couch and Lillard (1998) show that the estimates of the IGE correlation are quite sensitive to rules used to select the sample from which the correlation is estimated.

The empirical work here uses data on matched pairs of sons and fathers from the National Longitudinal Surveys Old Cohort Databases (NLS) for the United States and from the German Socio-Economic Panel (GSOEP) for Germany. I estimate a log-linear earnings model and models with linear, quadratic, and cubic terms in fathers' earnings. The estimates are then used to calculate the elasticity between earnings of fathers and sons at each quintile of the fathers' earnings distribution. These elasticities are compared to the elasticity from the log-linear model.

The resulting picture of the IGE is sharply different than found in previous studies. In particular, the apparent similarities across Germany and the United States are found to result from averaging very different IGE associations across the distribution of fathers' earnings. The correlation between earnings of poor fathers and their sons is relatively low in the United States compared to Germany. At the other end of the fathers' earnings distribution, sons of fathers with high earnings in the United States have earnings that are more closely associated with their fathers' earnings than in Germany. The results imply that poor youth in the United States are relatively more upwardly mobile than are poor youth in Germany while downward earnings mobility of sons at the upper tail of the fathers' earnings distribution is relatively low in the United States and very high in Germany.

The rest of the paper proceeds as follows. The theory and its implications are summarized in the next section. Section 3 describes the data. Estimation issues and model specification follows in Section 4. Results are presented and discussed in Section 5. I close the paper with a discussion of the implications of the results, shortcomings of the approach used, and suggestions for future research.

2. Theoretical Framework

The theory of Becker and Tomes (1986) suggests that the intergenerational earnings association will be higher among poor fathers when poor families are unable to borrow to finance human capital investments in children. Lillard (1998) introduces other sources of nonlinearities because he includes parental time as an input in production of human capital and assumes that youth may differentially inherit both genetic endowments and what Becker and Tomes (1986) call 'family culture.'⁴

In the basic model, earnings of one generation are linked to earnings of a subsequent generation through four channels, all of which operate through production of human capital. A person earns money by renting his human capital in the labor market. The production of human capital is assumed to depend on a given person's genetic characteristics and family culture, both of which are assumed to be exogenously inherited; on the resources a child and

his parents have to finance his investment in human capital; and the resources a parent directly contributes to production of human capital. Consider how each source potentially leads to a nonlinear IGE association.

Transmission of Endowments

Becker and Tomes (1986) consider endowments to be the set of characteristics which affect human capital production and which are exogenously inherited from parents. Endowments consist of genetically determined characteristics and exogenously transmitted behavioral traits which reflect a particular 'family culture.' Each set of characteristics or behavioral traits collectively affect what a person earns because it determines his optimal level of human capital investment. Family culture might, for example, include traits such as how parents resolve disputes and general attitudes toward risk or education. Family culture can be negative, neutral, or positive in the sense that it can reduce, leave unchanged or increase human capital production. Although there is evidence that environmental factors affect expression of genetic potential (Connor and Streissguth, 1996), I make the simplifying assumption that the transmission of genetic characteristics is independent of family culture. In all that follows, the term 'family culture' refers to exogenously transmitted traits.

Although the process by which characteristics are transmitted is still only poorly understood, rates of inheritance may be nonlinearly related to combinations of the underlying characteristics. For example, it seems plausible that combinations of genetic characteristics which lead to high or low earnings are both rare and unlikely to be replicated in successive generations. This conjecture implies that regression to the mean in genetically-based earnings ability is faster at the tails of that distribution. Similarly, some researchers model family culture to be transmitted at rates which increase as various risk factors accumulate. Sameroff, Seifer and Zax (1982) find strong evidence of lower scores on tests of 'social competence' among children who were exposed to more environmental risks as infants, including among other things multiple factors such as high maternal anxiety and stressful life events. These correlations persisted over the child's life. Less evidence is available about the acquisition of more positive behavioral traits (such as attachment to the workforce and regular work habits). If the cumulative risk models proposed by Sameroff et al. (1982) operate symmetrically across the distribution of family culture, then children of parents with low and high earnings may inherit negative and positive family culture at higher rates, assuming that family culture is positively correlated with parental earnings.

These conjectures and limited evidence imply that earnings may be nonlinearly related even if no other mecha-

⁴ See Lillard (1998) for complete description of the model.

nism links earnings across generations. Children of poor parents with better than average genetic ability may not realize higher earnings because, on average, their advantage of greater ability could be offset by the disadvantage of poorer family culture. A symmetric 'successful family culture' will inhibit regression to the mean at the other end of the parental earnings distribution. Thus, the inheritance of family culture might lead to nonlinear patterns in the IGE association, independently of whether capital markets function well.

Imperfect Capital Markets

Because it is assumed that one must pay money to acquire human capital, the functioning of capital markets potentially raises the IGE association of sons whose fathers have few financial resources. To borrow money to pay for education, a person pays a price equal to the forgone interest rate paid on alternative investments and a risk premium paid to borrow money. Under perfect information this risk premium reflects individual-specific default risks. When information is costly, lenders assign average default risks on the basis of group characteristics. Under imperfect information, the opportunity cost of money resources will reflect capital market imperfections to the extent that an individual's default risk differs from the average risk in his assigned group. In either case, as incomes rise, the risk premium falls.

Capital market imperfections will generally raise the IGE association among poorer families. A child whose ability exceeds that of his poor (and less able) father will pay more to borrow money to finance his education than will a child with similar default risk whose father has more resources. Consequently, the child is unable to fully realize the earnings his ability could command. A less recognized point is that capital market constraints may *lower* the association between earnings of children and parents if they bind on parents with earnings above those generated by average endowments.

This point is easily established. On average, high-ability parents have children whose ability is lower than their own. If capital market constraints bind for high-ability parents, their children will be unable to borrow the money to get the education they need to realize as much earnings as their ability would potentially generate. Therefore, the earnings of these children are less similar to their parents than they otherwise would have been. Since the association in earnings falls, downward mobility is raised above the rate implied by the natural rate of regression to the mean.

Parental Inputs of Time and Money

Finally, when parental time is a necessary input in production of a child's human capital, the rate of regression

to the mean is increased, *ceteris paribus*.⁵ Here, parents with high earnings face a time constraint. As earnings increase, parental time grows more costly and parents devote less time to their child's human capital production. All else equal, the consequent reduction in human capital reduces the association between earnings of highly paid parents and children. Conversely, low costs of time among poorly paid parents increases time investment. *Ceteris paribus*, parental earnings fall, human capital production increases, earnings of children increase, and the IGE association is reduced.

The association between earnings of sons and fathers is likely to vary across the fathers' earnings distribution. Certainly there is no reason to expect a log-linear relationship if capital markets function imperfectly, rates of transmission of genetic characteristics and family culture vary, and the relative cost and substitutability of parental time in human capital production of children varies across the distribution of fathers' earnings. Ultimately, the question of the shape of the association is an empirical one.

3. Data and Model Specification

The empirical analysis uses data from the NLS Older Men, Older Women and Young Men cohorts. Nationally representative samples of 5020 men aged 45–59 and 5225 younger men aged 14–24 were surveyed in 1966 (NLS 1987). 5083 older women aged 30–44 were first surveyed in 1967 (U.S. Department of Labor Statistics 1997). Data from sons and fathers are linked in both of the Older Cohorts. Of 3,219 matched father-son pairs, 1,712 pairs had non-missing data needed to estimate the earnings models. In the GSOEP, a nationally representative sample of just under 6,000 households and approximately 16,000 individuals were first surveyed in 1984. All members of the original households and all members of the households subsequently formed have been resurveyed each year. Data from the most recent survey year, 1998, are included in this analysis. As in the NLS data, sons are matched to fathers. Unlike the NLS data, no distinction is made between biologically related or unrelated fathers and sons.

Sample Selection

The sample was selected by age and labor force status. Men were admitted if they reported earnings between the ages of 18 and 65 in the NLS and between the ages of 18 and 60 in the GSOEP. Earnings were not counted for men who were in school, retired, or out of the labor force. In

⁵ Time should be treated here as a composite parental input which includes inputs like love, attention, help with homework, and other less tangible resources which importantly determine production of human capital.

the NLS, to have earnings counted, men must have worked at least 40 weeks in one of the years for which earnings were measured.

In the literature, many studies include earnings of men who are age 25 or older. IGE associations tend to be larger when one excludes earnings of younger men. When panel data are used, arbitrary age restrictions introduce potential attrition bias because young men who drop out of school also tend to exit earlier. Some authors try to justify the exclusion of younger men (Behrman and Taubman 1990) by noting that the coefficient of variation for earnings is much higher among younger men. However, such estimates of the coefficient of variation are based on samples from which younger men systematically exit. For example, of 1,712 sons in the NLS sample, 292 sons left the sample before they turned 25 years of age. Those who left were less well-educated, had parents who were less well-educated and fathers with lower average earnings than the men who remained in the sample past age 25. Of those who left, 23.3 percent completed 11 or fewer grades versus 11.5 percent who survived in the 25+ sample. Of the parents who left, 68.8 percent of the fathers and 89.9 percent of the mothers completed fewer than 12 years of schooling versus 51.8 percent and 84.4 percent of fathers and mothers respectively in the sample who survived. Further, fathers of men who left earned an average of 18,298 real 1984 dollars versus 21,173 among the survivors. Because there are no compelling theoretical grounds for excluding the younger men, they are used to estimate the IGE association.

Earnings of sons and fathers are calculated as the mean reported wage and salary income over all years in which they met the sample criterion and for which data were available.⁶ Earnings were averaged over from 1 to 18 years for fathers and from 1 to 12 years for sons in the NLS data and over 1 to 14 years for fathers and sons in

the GSOEP data.⁷ In the NLS data, if earnings were missing or reported to be zero in any given year, two steps were taken. First, if business income was reported, then the business income was used and a dummy variable was coded to denote the substitution.⁸ If no business income was reported, the earnings were included as a zero in the average only if it could be established that the respondent was unemployed for all weeks in that year. An observation was dropped if earnings were zero in all observed years. Earnings data in the GSOEP were imputed in a very few cases. Those data were used in the analysis.

Other Control Variables

Each model includes the same basic set of control variables. This set includes the number of years used to calculate average earnings for both fathers and sons and, in the NLS data, dummy variables that indicate if business income proxied for earnings. In addition, the age and age squared of both father and son are included to account for differences which arise because sons and fathers are being compared in different life-cycle stages. Table 1 presents descriptive sample statistics.

⁶ In the calculation of average income, earnings or assets, mid-point values were used when the NLS data were categorical. A mean value was assigned to top-coded categories by fitting a Pareto distribution to the observed frequency distribution.

⁷ The number of years used in each average is included as a separate control variable. Also note, because earnings of sons from poorer families are averaged over fewer years, estimates of b will be biased downward among sons from poorer families. The resulting bias works against the hypothesis that the IGE association is higher among sons from poorer families.

⁸ In both the younger and older sample, average earnings were lower among sons for whom business income was used as a proxy in the average of sons' earnings. When business income proxied for fathers' earnings, sons had higher average education and earnings.

Table 1

Summary statistics for NLS and GSOEP samples

Variable	NLS		GSOEP	
	Mean	Std. dev.	Mean	Std. dev.
Earnings sons	28,395	(14,368)	39,349	(24,080)
Number of years in average-sons	5.98	(2.73)	6.33	(4.03)
Earnings fathers	31,926	(21,785)	62,456	(37,606)
Number of years in average-fathers	6.16	(3.47)	8.12	(4.38)
Average age — sons	25.28	(2.62)	25.89	(3.91)
Average age — fathers	52.06	(5.79)	52.00	(5.00)
N	1,712		1,061	
Households	1,399		802	

Notes: US dollar values are reported in constant 1997 dollars. German marks are constant 1997 deutschmarks.
Source: Author's calculations

Estimation

Although several methods have been used in the literature to estimate the IGE association, this study uses regression techniques to estimate the partial correlation between earnings of sons and fathers. The method it uses permits control variables to be used in estimating the IGE association. These estimates are then used to calculate elasticities in each quintile of the fathers' earnings distribution.

Models are estimated with OLS regression, correcting standard errors for observations of multiple sons from the same household. As a consistency check to be sure that the OLS estimates are not unduly influenced by outliers in the data, the models are also estimated using median regression and robust regression techniques. Because median and robust regression coefficient estimates are less sensitive to outliers in the data, those estimates serve as a consistency check on the OLS estimates.

I estimate a reduced form equation relating earnings of sons and fathers. The reduced form model is given by: $E_t = b'E_{t-1} + w'Z + u$ where $u \sim N(0, s^2)$ and E_t is the average of earnings of sons, E_{t-1} is a vector which includes the average of absolute dollar earnings of fathers, the square and the cube of the father's average earnings and Z is a vector of basic control variables described above.

4. Results

Table 2 presents selected OLS regression coefficients for four models of earnings estimated on the GSOEP and NLS data. The first column presents the estimated coefficient on father's earnings when a log-linear model is fit to the data. The elasticity of .109 in the German data and .144 in the United States data are quite similar to esti-

mated elasticities reported by Couch and Dunn (1997) who estimated an elasticity of .112 in the GSOEP for Germany and .127 in the Panel Study of Income Dynamics for the United States. The next three columns report estimated coefficients for models with linear, quadratic, and cubic terms in father's earnings. In the German data, the coefficients on both the quadratic and cubic terms in father's earnings are statistically significant at conventional levels of significance. In the data for the United States, the quadratic and cubic terms in model 3 are not statistically significant while the coefficients on both the linear and quadratic terms in model 2 are many times larger than their standard errors.

Figure 1 and Figure 2 present the implied elasticity of sons' earnings with respect to fathers' earnings across the distribution of fathers' earnings in the United States and Germany. The elasticities are computed as the mean earnings elasticity in each quintile of the father's earnings distribution. Figure 1 shows the change in the estimated earnings elasticity in the United States data as one relaxes the log-linear functional form assumption. Relative to the log-linear model, earnings elasticities implied by the quadratic and cubic models are higher at all portions of the distribution of fathers' earnings. The quadratic and cubic models imply much less earnings mobility in the United States than implied by the log-linear model, especially in the upper half of the distribution of fathers' earnings.

Figure 2 shows, for Germany, the starkly different earnings elasticities that a log-linear model yields versus the quadratic or cubic models. Here, the pattern is different from that of the United States. In the lowest quintile of the fathers' earnings distribution, the implied earnings elasticity from the cubic model is almost three times larger than that implied by a log-linear model. This sharply higher association is offset by a sharply lower

Table 2

Selected regression coefficients

Variable	Germany			United States				
	Log-linear	1	2	3	Log-linear	1	2	3
Average earnings father*10 ⁻³	0.109 (.024)	0.018 (.014)	0.076 (.029)	0.182 (.057)	0.144 (.030)	0.133 (.029)	0.284 (.032)	0.220 (.069)
Square of father's earnings*10 ⁻⁵			-0.025 (.009)	-0.118 (.044)			-0.017 (.003)	-0.001 (.017)
Cube of father's earnings*10 ⁻⁷				0.018 (.008)				-0.009 (.009)
R ²	0.259	0.265	0.266	0.270	.149	.286	.304	.305
N	1061 sons in 802 households				1712 sons in 1399 households			

Notes: Huber-White standard errors in parentheses are corrected for observations of multiple sons from the same household. All models include linear and quadratic terms for the average age of sons and fathers in the years for which earnings are observed and the number of years over which earnings are averaged. Models for Germany estimated with waves B-O of the GSOEP using robust regression estimation. Models for the United States estimated with data from the NLS Older-Cohort Databases 1966–1991.

Figure 1

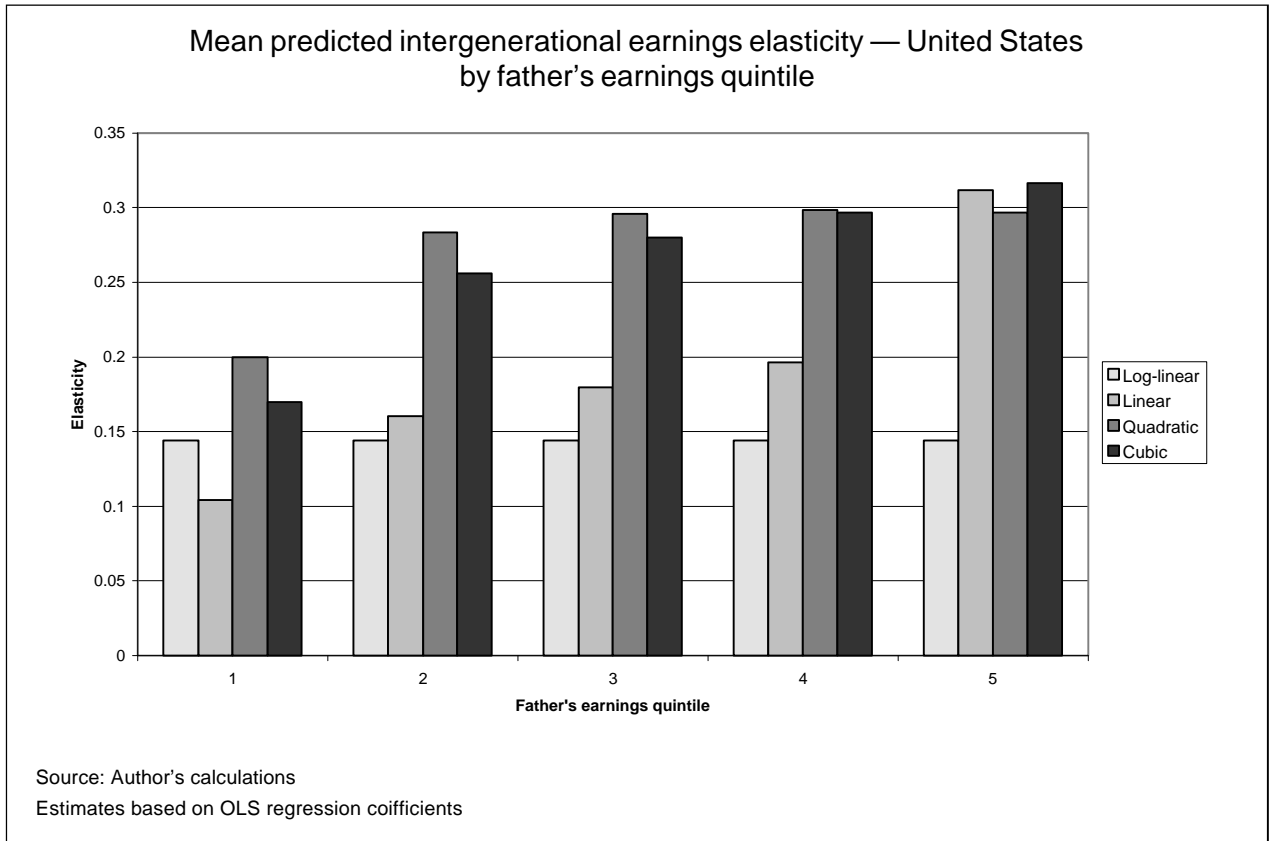
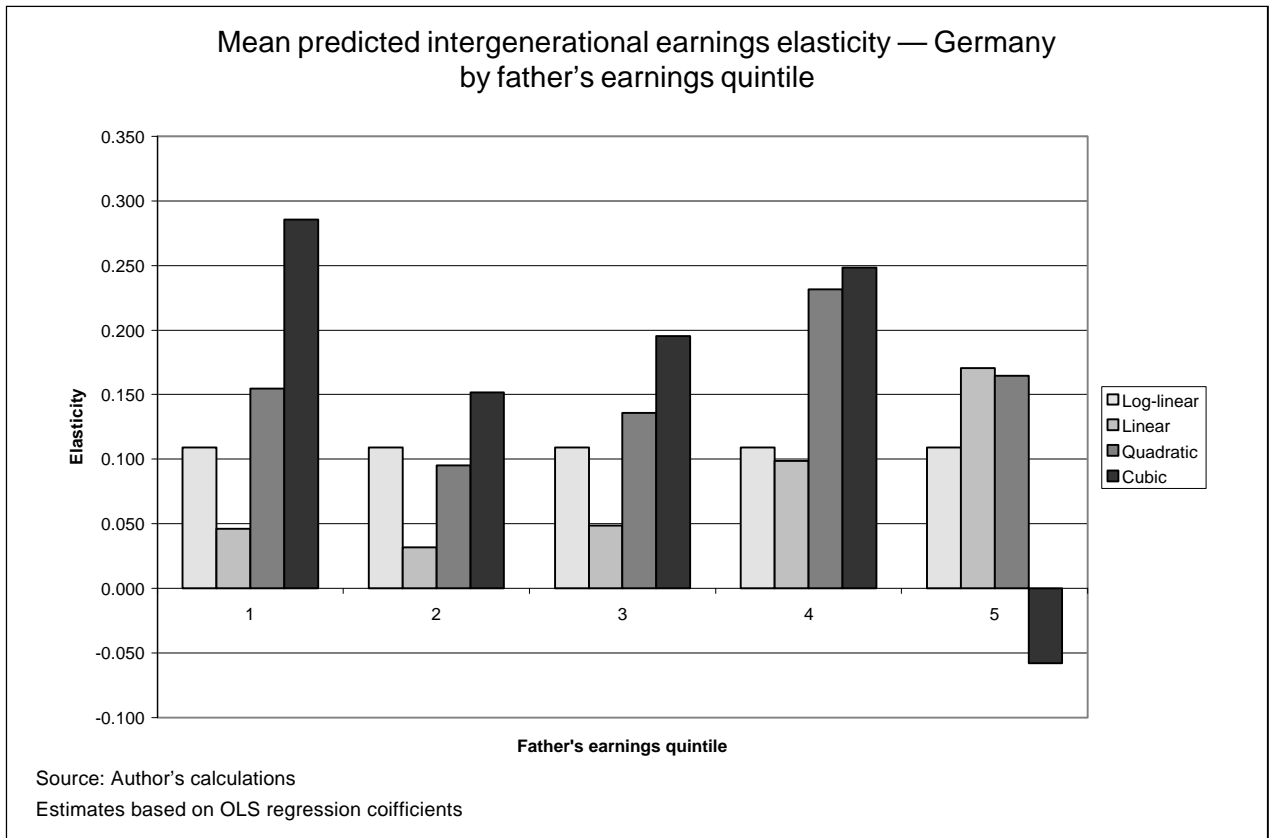


Figure 2



intergenerational earnings association at the other tail of the distribution. In the highest quintile, a percentage increase in fathers' earnings is associated with almost no change in sons' earnings (a slightly negative change is implied).

Figure 3 draws a sharper contrast between the earnings elasticities implied by the cubic models in the United States and Germany. Here we see that sons with poor fathers are much more mobile in the United States compared to Germany but that mobility is relatively higher in Germany for sons from all other portions of the fathers' earnings distribution. Indeed, at the very upper tail, it appears that there is no association between earnings of fathers and sons.

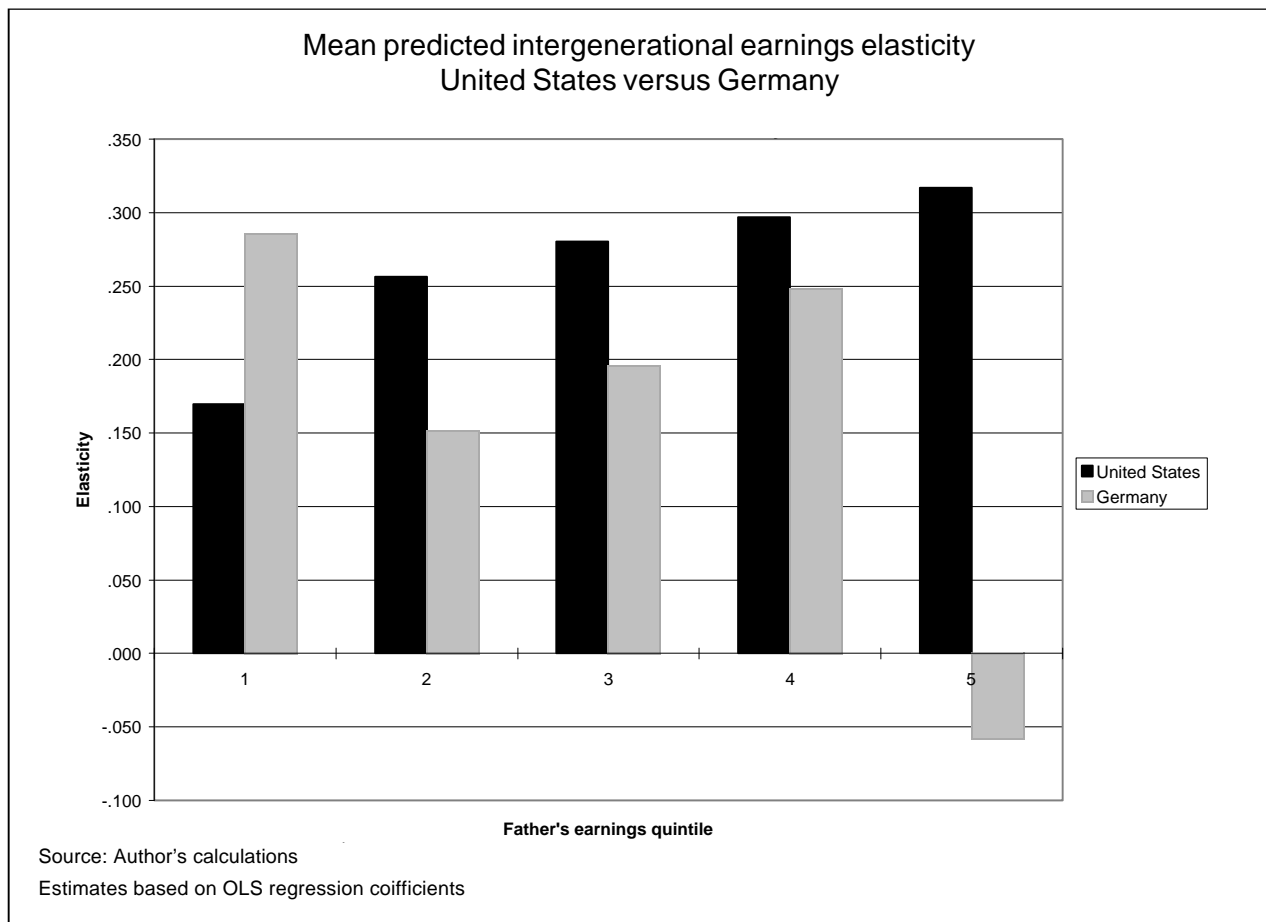
These patterns of intergenerational earnings associations shown in Figure 1 for the United States are consistent with the patterns reported in Behrman and Taubman (1990) and Corak and Heisz (1997). Using different data sets, both studies find the same pattern of higher elasticities among parents with higher earnings/income. Corak and Heisz (1997) draw data from Canadian income tax records on roughly 440,000 father-son pairs. They esti-

mate the elasticity of sons' incomes with respect to fathers' incomes to rise from close to zero in the lowest tail of the fathers' income distribution to a peak of .3 at around the median before falling again to about .2 in the upper tail of the income distribution.

5. Conclusion

This paper relaxed the functional form assumption used in most studies of the intergenerational earnings association. Theory suggests that it is likely that the association will vary across the distribution of fathers' earnings. Using a log-linear specification, others have found that the degree of earnings mobility is similar in Germany and the United States. This study investigates whether the similarity in mobility persists when a more flexible function form is fit to the data. The principal finding of the study is that the pattern of the intergenerational earnings association differs quite dramatically in the United States and Germany. Sons from poor families are much more upwardly mobile in the United States than in Germany. Sons from the highest quintile in Germany are much more

Figure 3



downwardly mobile than are sons in the United States. In the middle of the distribution, mobility is greater in Germany than in the United State.

These findings should be interpreted cautiously. Much work needs to be done to check whether the results persist when other controls are added to the analysis. They

suggest, however, that inferences about intergenerational earnings mobility should not be based on single measures of the average correlation across the whole distribution of earnings. Instead, attention needs to be paid to how the intergenerational association in earnings varies across families of different resource levels.

References

- Altonji*, Joseph G. and Thomas A. *Dunn* 1991. "Relationships Among the Family Incomes and Labor Market Outcomes of Relatives." *Research in Labor Economics* Vol. 12: 269–310.
- Becker*, Gary S. 1991. *A Treatise on the Family*. Enlarged Edition. Cambridge, MA: Harvard University Press.
- Becker*, Gary S. and Nigel *Tomes* 1986. "Human Capital and the Rise and Fall of Families." *Journal of Labor Economics* 4(3, part 2, July): S1–S39.
- Behrman*, Jere and Paul *Taubman* 1990. "The Intergenerational Correlation Between Children's Adult Earnings and Their Parents' Income: Results from the Michigan Panel Survey of Income Dynamics." *Review of Income and Wealth* 36(2): 115–27.
- Connor*, Paul D. and Ann P. *Streissguth* 1996. "Effects of Prenatal Exposure to Alcohol Across the Life Span." *Alcohol Health and Research World* 20(3): 170–74.
- Corak*, Miles and Andrew *Heisz* 1997. "Unto the Sons: The Intergenerational Income Mobility of Canadian Men." Mimeo. Statistics Canada, Ottawa.
- Couch*, Kenneth A. and Thomas *Dunn* 1997. "Intergenerational Correlations in Labor Market Status." *Journal of Human Resources* 32(1): 210–32.
- Couch*, Kenneth. A. and Dean R. *Lillard* 1998. "Sample Selection Rules and the Intergenerational Correlation of Earnings." *Labour Economics* 5: 313–329.
- Featherman*, David and Robert *Hauser*. 1978. *Opportunity and Change*. Academic Press: New York.
- Goodman*, Leo. 1979. "Simple Models for the Analysis of Association in Cross-Classifications Having Ordered Categories." *Journal of the American Statistical Association* 74: 537–552.
- Lillard*, Dean R. 1998. "Nonlinearities in Intergenerational Mobility in Education and Earnings." Manuscript. Cornell University
- Lillard*, Lee A. and Robert T. *Reville* 1997. "Intergenerational Mobility in Earnings and Occupational Status." Manuscript. Rand Corporation, Santa Monica, CA.
- Mulligan*, Casey B. 1997. *Parental Priorities and Economic Inequality*. University of Chicago Press: Chicago.
- National Longitudinal Surveys. 1987. *Surveys of Young Men 14–24*. Codebook Supplement, Appendix 9, Public Release Version 1966–1981, Center for Human Resource Research, The Ohio State University.
- Peters*, H. Elizabeth. 1992. "Patterns of Intergenerational Mobility in Income and Earnings." *Review of Economics and Statistics* 74(3): 456–66.
- Sameroff*, A.J., R. *Seifer*, R. and M. *Zax* 1982. "Early Development of Children at Risk for Emotional Disorder." In: *Monographs of the Society for Research in Child Development*, 47, (No. 7, Serial No. 199).
- Solon*, Gary. 1992. "Intergenerational Income Mobility in the United States." *American Economic Review* 82(3): 393–408.
- U.S. Department of Labor Statistics. (1997). *NLS Handbook*. Bureau of Labor Statistics, Washington, DC.
- Zimmerman*, David J. (1992). "Regression Toward Mediocrity in Economic Stature." *American Economic Review* 82(3): 409–429.