Estimating Nonlinear Intergenerational Income Mobility with Correlation Curves

William Nilsson

Department of Applied Economics, University of the Balearic Islands
Ctra Valldemossa Km 7,5, E-07122 Palma de Mallorca, Spain
Phone: +34 971 17 13 77, Fax: +34 971 17 23 89
william.nilsson@uib.es

Abstract: A correlation curve is proposed as an alternative measure to study the degree of intergenerational income mobility, i.e. how income status is related between parents and adult child. The method overcomes the shortcomings of the elasticity of children's income with respect to fathers' income (i.e. its sensitiveness to different dispersion among the generations) and the correlation coefficient (i.e. its inability to capture nonlinearities). The method is particularly suitable for comparative studies and in this study is applied to labour income in comparison to disposable income. Nonlinear correlation curves are found, which in some cases substantially differ from corresponding nonlinear elasticities.

Keywords: Intergenerational mobility; nonlinear; nonparametric; correlation curve

JEL: C14, D63, J62

1 Introduction

The empirical literature on intergenerational income mobility has, during the last 20–25 years, been highly focused on estimating the correlation between a father's income and a child's income in adulthood. Often the estimate of interest has been the elasticity obtained from a regression of the father's log income on the child's adult log income. Compared to the earlier literature, Zimmerman (1992) and Solon (1992) provide important contributions by underlining the importance of avoiding homogenous samples as well as avoiding the use of short-run measures of income, which otherwise would result in downward-biased estimates. Other important methodological concerns when it comes to estimating the intergenerational income elasticity are life cycle bias (Jenkins, 1987; Haider & Solon, 2006; and Grawe, 2006) and the possibility of a nonlinear relation (Corak & Heisz, 1999; Österbacka, 2001; Björklund & Chadwick, 2003; Fertig, 2003; Grawe, 2004; and Bratsberg *et al*. 2007). Couch & Lillard (1998) highlight the sensitivity of the elasticity due to different sample selection rules.

Estimating intergenerational income elasticity has become the main way to study intergenerational income mobility, but in some occasions it is complemented with alternative methods. Dearden *et al*. (1997), Fertig (2003) and Bratberg, Nilsen & Vaage (2005) complement their analyses with transition matrices. Eide & Showalter (1999) use quantile regression to estimate different elasticities for different quantiles of the son's earnings distribution. Österberg (2000), Fertig (2003) and Jäntti *et al*. (2006) estimate the probability that the child will end up at a particular decile (or quintile) given the decile (or quintile) of the parent. However, both transition matrices and conditional probabilities are accompanied by the problem that upward mobility from the top (or downward mobility from the bottom) is impossible, and as a consequence the result tends to indicate low mobility for these groups.

Data on parents' and children's adult income are usually measured at different times. Even if the aim is to collect data for the two generations at more or less the same stage of the life cycle, the distributions of parents' and children's incomes are likely to differ. Different generations grow up in quite different societies and it is not necessary that the distributions are approximately the same. Solon (1992) discusses the assumption that the variance in income for the two generations is the same. If it is not, the elasticity cannot be used as measure of the degree of association. This important observation has in many cases been left aside, and only the elasticity has been reported in many empirical studies (Blanden, 2011). A difference in dispersion would, *ceteris paribus*, affect the intergenerational income elasticity, while the correlation coefficient would not be affected (Jäntti *et al*., 2006). The difference between these two measures of mobility is also discussed by Fertig (2003) and Aaronson & Mazumder (2008), where trends in inequality are also related to trends in intergenerational elasticity. In Blanden *et al*. (2007) the elasticities are scaled to obtain partial correlation coefficients. The reason is precisely to cope with different changes (between the generations) in inequality in different cohorts.

Björklund *et al*. (2012) found an intergenerational elasticity of 0.260 for earnings and 0.168 for income for Swedish data. The difference between the measures was found to be much smaller when the elasticities were standardized to obtain correlation coefficients (0.23 respective 0.19). The difference between an intergenerational *transmission*, i.e. an elasticicity, and a *degree of association* measured with a correlation coefficient can be crucial for comparisons. In the same study, using a linear spline regression across fathers' fractiles, elasticities of 0.896 for incomes and 0.447 for earnings, for the top 0,1% of the income distribution of the fathers were found. Comparing the descriptive statistics for the log earnings of the fathers and the sons reveal a much higher dispersion, at the top tail of the distribution, for the sons. This is in particularly the case when income is analysed. Hence, it is expected that the elasticity will be higher at this position of the income distribution. While the intergenerational income transmission is found to be very strong, the measures cannot provide a conclusion about the *degree* of intergenerational income or earnings association.

The purpose in this study is to introduce a local measure of the degree of association of incomes of two generations. Using a *correlation curve* follows the recommendation in Blanden (2011) to complement the elasticity with a correlation coefficient, and, in addition, the measure is local, which allows a varying degree of association as recommended in Bratsberg *et al*. (2007). The approach avoids the sensitivity to differences in dispersion in the two generations that is accompanied by the (nonlinear) elasticity.

Bjerve & Doksum (1993) and Doksum *et al.* (1994) introduced correlation curves, $\rho(x)$, to measure the strength of a relation locally at different values of a covariate X. Their correlation curve is a universal scale-free measure which shares many properties with the correlation coefficient, but it is, in addition, suitable for nonlinear models. In the same way as the correlation coefficient is a standardized version of the regression slope, the correlation curve is a standardized local regression slope. The correlation curve is invariant to changes in the origin and scale and is $-1 \leq \rho(x) \leq 1$ for all x. Accordingly, the measure is easy to interpret and fairly easy to implement once a nonparametric estimation technique for the local regression slope is specified. Pointwise confidence intervals can be included with a bootstrap technique (Nilsson & del Barrio, 2012).

An advantage of this method is that intergenerational mobility is measured at each position in the income distribution of the fathers/parents, which allows a varying strength of the relation as suggested in previously mentioned literature. In addition, the method is particularly suitable for comparing different populations, for example different countries, or comparisons over time.

In this study, the method is applied to Swedish data and complemented by nonparametric estimates of the elasticity. The purpose is to compare two different aspects of intergenerational income mobility. First, mobility is measured for labour income; i.e., interest is in equality of opportunity in providing a salary for the household. Labour income is measured as labour income before paying taxes and receiving benefits and without taking into account the size of the household. The second aspect studies income mobility in terms of equality of opportunity to enjoy a certain standard of living. In this case, disposable income is used and income is measured after paying taxes and receiving benefits. This measure also contains income acquired by other household members, and disposable income is weighted for the composition of the household.[1] In this case, both parents' income is used instead of only the father's income to better capture the standard of living.

The results indicate important nonlinearities, both for labour income and disposable income, in intergenerational income correlation. Using the correlation coefficient does not give an accurate summary of the correlation. The correlation between disposable income and parental income is higher for higher deciles of the income distributions of the parents.

---

[1] The consumption weights are based on norms defined by the National Board of Health and Welfare in Sweden. A family of one adult implies a weight of 1.16. For two or more adults, each adult is weighted 0.96. Children 0-3, 4-7 and 11-17 years old add, respectively, 0.56, 0.66 and 0.76.

In addition, the correlation is, for higher deciles, stronger for disposable income than for labour income. At the same time, the nonparametric elasticity is clearly higher for most parts of the distribution when labour income is analysed compared to disposable income. The explanation for these seemingly contradictory results is the sensitivity of the elasticity to differences in the dispersion of the children's and parents' incomes. Using the correlation curve for comparisons is clearly justified.

The method is described in section 2. The data are explained in section 3 and the results are presented in section 4. Concluding remarks are offered in section 5.

2 Method

The literature review of intergenerational income mobility is very brief in this paper as important surveys are available (see, for example, Solon [1999] and Blanden [2011]). The focus in this section is on the introduction of a new method to study intergenerational income mobility. The method is based on nonparametric techniques to estimate correlation curves. Bootstrap techniques are used to obtain pointwise confidence intervals.

2.1 Intergenerational income correlation and intergenerational income elasticity

Consider a data set of individuals with income measured at a point several years into adulthood. Several years of data on parental income are available. The incomes for the individual and the parents could be measured at the same years, but then the incomes would be measured at very different times in the life cycle for the two generations. Another possibility is data where the incomes for parents are measured sometime during the childhood or youth of the individuals. In either case it is possible that the distribution of incomes will be different for the two generations. Being in different positions in the life cycle is likely to provide different distributions. The income dispersion could also have increased or decreased over time. Differences in standard deviation of the income distributions affect the intergenerational income elasticity, while the correlation coefficient is unaffected. As a consequence, using the correlation coefficient makes the comparison of intergenerational income mobility much more transparent. The degree of relation is not mixed up with differences in the standard deviation for the incomes in the two generations.

2.2 Intergenerational correlation curves and nonparametric elasticities

The intergenerational income elasticity can easily be allowed to vary over the income distribution of the fathers/parents by estimating the regression function with nonparametric techniques, $m(x) = E(Y \mid X = x)$, where X in this case is the logarithm of income of the fathers/parents and Y is the logarithm of income of the children at adult age. The slope of the regression function, $m'(x)$, corresponds to a local measure of the intergenerational income elasticity. An advantage of a nonparametric estimation technique is, of course, that we do not impose restrictions on the functional form. A disadvantage, at least when the interest is in comparing the results, is, however, that the elasticity is still affected by the distributional differences of the two generations. If we want a more transparent measure of the local *degree* of relation a correlation curve can be estimated,

$$\rho(x) = \frac{\sigma_x m'(x)}{\left[\{\sigma_x m'(x)\}^2 + \sigma^2(x)\right]^{1/2}}, \tag{1}$$

where $\sigma^2(x) = \mathrm{var}(Y \mid x)$ is the residual variance. σ_x is the standard deviation of X. The correlation curve is easily calculated once a nonparametric technique is used to estimate the derivative of the regression function and the residual variance.

Bjerve & Doksum (1993) discuss in detail the properties of the correlation curve. A few of the properties that are relevant for this application are mentioned below. First, the correlation curve is invariant to changes in the origin and scale. The correlation curve is standardized to be $-1 \leq \rho(x) \leq 1$ for all x, and the strength of the association is interpreted in the same way as the correlation coefficient: $\rho(x) = 0$ for all x when X and Y are independent, and $\rho(x) = \pm 1$ for all x when X is a function of Y. For linear models the correlation curve reduces to the correlation coefficient. An important difference compared to the correlation coefficient is that, in general, $\rho_{XY}(\cdot) \neq \rho_{YX}(\cdot)$. In this particular application the main interest is on $\rho_{XY}(x)$, which can be compared to $m'(x)$. $\rho_{XY}(x)$ is clearly affected by $m'(x)$, but also detects a possible heteroscedastic pattern where the association could be locally weaker or stronger. As an additional analysis, $\rho_{YX}(y)$ is also estimated. The idea is to relate the elasticities from a quantile regression approach with a correlation curve, allowing the correlation to vary with respect to different positions of adult children's income distribution.

To implement the correlation curve it is appropriate to use a flexible technique to obtain a local measure of the regression slope and a local measure of the residual variance. In this case, the nonparametric method is local polynomial regression, due to its advantageous properties (Fan, 1992, 1993). The data-driven procedure suggested by Fan & Gijbels (1995a) to find the optimal bandwidth for the second derivative is used. Technical details can be found in Appendix.

2.3 Pointwise confidence intervals for the correlation curve

The correlation curve can easily be estimated with the nonparametric method described in the Appendix. It is necessary to complement this method with confidence intervals to be able to make statistical inferences based on the correlation curve. This is particularly the case if it is of interest to compare correlation curves for different populations. Nilsson & del Barrio (2012) suggest bootstrapping to estimate confidence intervals and the coverage rates are found to be satisfactory. They use a wild bootstrap technique that maintains a possible heteroscedastic pattern in the data. Härdle & Mammen (1993) introduced wild bootstrapping to obtain confidence intervals for nonparametric regressions.

To obtain a bootstrap pointwise confidence interval the following steps are used:

1. Estimate $\hat{m}_h(x)$ using equation (A1) with an optimal bandwidth, h, to calculate the residuals: $\hat{\varepsilon}_i = y_i - \hat{m}_h(x_i)$. A local quadratic regression $(p = 2)$ is used for estimating the regression function to reduce the bias.

2. Define a new random variable ε_i^*, which is $\varepsilon_i^* = \gamma\hat{\varepsilon}_i(1-\sqrt{5})/2 + (1-\gamma)\hat{\varepsilon}_i(1+\sqrt{5})/2$, where $\gamma = 1$ with probability $p = (5+\sqrt{5})/10$ and $\gamma = 0$ with probability $1-p$. This assures that $E\varepsilon^* = 0$ and $E\varepsilon^{*2} = \hat{\varepsilon}_i^2$.

3. Add the resampled residuals to $\hat{m}_g(x)$, where g indicates a larger bandwidth compared to h, i.e. $y_i^* = \hat{m}_g(x_i) + \varepsilon_i^*$, to obtain new observations for the dependent variable. Galindo *et al*. (2001) suggest the bandwidth $g = h^{(2p+5)/(2p+7)}$ when p is even.

4. Estimate $m'(x)$ and $\hat{\sigma}^2(x_0)$ using equations (A1) and (A2) for the new sample using $(x_1, y_1^*),...,(x_n, y_n^*)$. The results are combined as indicated in equation (1) to obtain the

correlation curve. A local quadratic polynomial, $(p = 2)$, is used with a bandwidth, h, optimal for estimating the first derivative.

5. Steps 2–4 are repeated R times to obtain a pointwise confidence interval based on R bootstrap replications. The percentile method is then used to find a $(1-\alpha)$ pointwise confidence interval, where α is the chosen level of significance. Therefore, the graphs show percentile $100\times(\alpha/2)$ and $100\times(1-\alpha/2)$ of all estimated correlation curves at each x_i.

The method to illustrate mobility explained above is, in this study, applied to Swedish data. The data are discussed in the following section.

3 Data

The empirical analysis in this study is based on Swedish register data administered by Statistics Sweden. The first sample consists of the complete cohort of individuals born in 1965. The second sample is a 10% random sample of individuals born between 1949 and 1958. The years of outcome are 1994 to 1999. The first sample is labelled 'young' as the income is measured between the ages of 29 and 34, and the second sample is labelled 'middle aged' as the income is measured over six years between the ages of 36 and 50. These samples are further divided into male and female samples. The Population and Housing Census is used to identify parents. For the young sample, parents are identified if they are found present in the child's household in each of the years 1970, 1975 and 1980. It is not necessary for the parents to be biological parents, although this is the most common scenario. For the middle-aged sample, parents identified in the Population and Housing Census in 1965 are identified. By using the censuses in 1960 and 1970, the final sample is restricted to individuals where both the father and the mother are identified twice within a five-year difference. For example, for individuals born in 1949 the censuses in 1960 and 1965 are used and the parents are identified at age 11 and 16. The reason for restricting the samples to cases where the parents were actually present is to capture both biological and social reasons for an intergenerational income correlation.

The incomes of the parents are available for the years 1971, 1974, 1977, 1980 and 1983 and come from the Income and Wealth Register. For the middle-aged sample only the years 1971, 1977 and 1977 are used to measure fathers'/parents' income. The reason is that the income variable has more missing values for the years 1980 and 1983, for the natural reason of death. Fathers/parents were required to be alive until at least the year 1978 for the middle-aged samples.

The longitudinal database LOUISE is used for the income of the individuals, which is measured for the years 1994 to 1999. Labour income and disposable income are used in the analysis. Note that an exact counterpart to the measure of fathers'/parents' income is not available in the registers. The income is averaged over time for both generations. Individuals who died, or were living outside Sweden for at least one period, have been dropped from the sample. All income variables are measured in Swedish Crowns (krona) deflated to the price level of 2001. Summary statistics for the different samples are included in Table 1.

[place table 1 here]

The natural logarithm is used for all income measures. The samples are restricted to adult children and fathers/parents with positive average income. This restriction decreases the sample size for the young samples (male and female) by approximately 3% when labour income is used. For the male middle-aged sample the restriction decreases the sample size by about 6%, while the sample size decreases by 5% for the corresponding female sample. The restriction hardly changes the sample size at all when disposable income is used.

The standard deviation of labour income is, for both young and middle-aged samples, substantially higher than the standard deviation for the income of the fathers. The larger dispersion for the adult children is also evident when incomes in different percentiles are compared. Note, however, that fathers' average income is based on incomes over several nonconsecutive years and also that the measures of income actually are different. It is also possible that a general tendency for increased inequality in society, as well as life cycle differences, could explain the differences in dispersion. It is also important to remember that the fathers were selected based on being present in the household. Individuals spending shorter periods in the household, for example due to divorce, or individuals never in a partnership are accordingly not included in the income distribution. If individuals in a stable partnership tend to have a more stable position in the labour market it is possible that the distribution is more compressed. This selection is, of course, not applied for the adult children.

Disposable income has a smaller standard deviation for the adult child, and this is in fact very similar to the standard deviation for the combined income of the parents.

4 Results

The main measure for intergenerational income mobility in this study is the correlation curve. The mobility is analysed separately for men and women. As a departure for the analysis, correlation coefficients and linear elasticities are estimated for each sample. Note that the analysis is first carried out for the young samples and later repeated for samples of middle-aged sons or daughters. These results can be found in Table 2.

[place table 2 here]

4.1 Results for the young samples

The elasticities are substantially higher than the correlation coefficients when labour income is used for the male sample. An intergenerational elasticity of 0.34 is more than twice as high as the correlation coefficient. The same pattern is found for the female sample, where the intergenerational elasticity is 0.26 while the correlation coefficient is 0.11. The explanation of these relatively high estimates is, of course, the larger income dispersion for the adult children compared to the fathers.

When the analysis is repeated for disposable income in combination with parental income, the distributions are much more similar and the intergenerational income mobility almost coincides with the correlation coefficient. For the male sample both measures are 0.23, and for the female sample the elasticity is 0.20, while the correlation coefficient is 0.22.

Comparing the results for the different measures it is clear that the intergenerational elasticity is not very suitable for making comparisons of different samples. Without knowing the standard deviations of the samples, the intergenerational income elasticity would have led us to believe the intergenerational mobility to be higher for disposable income than for labour income. It is, however, clear that such a conclusion is only correct if we are willing to include distributional differences, hence different dispersion, as an important part of the ‘intergenerational mobility’concept. The correlation coefficient indicates that society is more rigid when it comes to disposable income than it is to labour income. In this case, the results are not affected by a different standard deviation between the income of the individuals and that of the parents.

Figure 1 includes a scatter plot and the estimated regression curve for the young male sample with labour income. It is evident that even for fathers of average income a child can have a very low labour income at an early stage in the labour market. It is, however, interesting to surmise whether the correlation, i.e. the degree of mobility, varies over the income distribution of the fathers/parents. The correlation coefficient would in that case give an inaccurate summary of the intergenerational mobility in the society. Correlation curves and nonparametric elasticities are estimated as suggested in section 2. Technical details concerning the nonparametric procedure are included in the Appendix.

[place figure 1 here]

The dotted curve is the median of the correlation curves from the bootstrap replications. The dashed curves are 95% pointwise confidence intervals. The solid grey curve is the median among the nonparametric derivates from the same bootstrap replications. To simplify the interpretation the figures show the correlation curves and elasticity at each rank of the distribution. Figure 1 shows an important difference between the correlation curve and the nonparametrically estimated elasticity. The correlation is stronger from the first decile to approximately the sixth decile and gradually decreases in the tails of the distributions. For these deciles the correlation is significantly above the correlation coefficient presented previously, and the use of a correlation curve is justified to capture the nonlinear relation. Figure 2 includes the results for the male sample when disposable income for the adult child is used and parents' income is used instead of only the father's income.

[place figure 2 here]

When disposable income is used the income distribution of the adult children is not as dispersed as when labour income is used. The elasticity is now found to be very similar to the correlation curve. The correlation is again found to vary over the distribution, and the shape is quite different compared to that found in Figure 1. The correlation is fairly low for low-income parents and gradually increases over the distribution. It is interesting to compare Figures 1 and 2. At decile 8 it is clear that the upper confidence interval for Figure 1 is below the lower confidence interval for Figure 2, hence the correlation is statistically significantly higher for disposable income.[2] Again, the correlation coefficient does not give an accurate picture of the correlation for a large part of the distribution.

Figure 3 includes the results for the young female sample when labour income is analysed.

[place figure 3 here]

---

[2] Note that the figures show 95% pointwise confidence intervals, and the correlation curve for the population is below the lower confidence interval with a probability of approximately 0.025. The same probability is applied for above the upper confidence interval for the other population. Hence, only using the figure to tell whether the two populations have statistically significant different correlations at some point implies a very restrictive significance level, i.e. less than 0.1%.

Figure 3 shows an important difference between the nonlinear elasticity and the correlation curve. The point estimate of the correlation curve is slightly below the corresponding curve for the male sample. The confidence intervals do, however, overlap. The shape of the curves is also fairly similar.

[place figure 4 here]

Figure 4 shows the results when disposable income is used to measure the income of the adult child. Parents' income is used instead of only the father's income. The pattern is fairly similar to the results shown in Figure 2 for the male sample. The correlation is lower for low-income parents and gradually increases over the distribution. The highest correlation is found for the eighth decile, and the point estimate is higher than the result for the male sample at the same decile. At the second decile the point estimate is lower than the result for the male sample. The confidence intervals do, however, overlap in both cases. Above the ninth decile the correlation becomes weaker for the female sample, although the confidence interval is also wider. If Figures 3 and 4 are compared, it is clear that at the eighth decile the correlation is statistically significantly higher for disposable income than for labour income. At the second decile the correlation is very similar in both cases.

4.2 Results for the middle-aged samples

When labour income is analysed for the male middle-aged sample both the elasticity and the correlation coefficient are found to be lower than that found for the young sample. This is also the case when disposable income is used together with parental income. A possible reason is that the age of fathers/parents for the middle-aged sample tends to be very high, which has been suggested to induce a negative life cycle bias in estimating the intergenerational income elasticity (Grawe, 2006). A life cycle bias can occur if workers with higher lifetime earnings tend to have higher wage growth due to human capital investments. Observed early-in-life earnings would be negatively correlated with the deviation of observed earnings from lifetime earnings. This correlation turns positive if observed earnings are measured later in life. Accordingly, if fathers are observed at a young age the life cycle bias will be positive. The life cycle bias will be zero at some point in midlife but turns negative if income is measured beyond that age. Another possible reason for the lower estimates is that fewer years are used to measure fathers' and parents' income, and the results could be affected to a greater extent by the use of a short-run measure. These issues are explored further in chapter 4.5.

The correlation is higher when disposable income is used compared to when labour income is used, while the opposite is the case for the elasticity. The same pattern is found for the female samples. The estimated values for the correlation coefficient and the elasticity can be found in Table 2. We should, however, be careful when interpreting the magnitude of these summary measures as the pattern could be different over the fathers' (or parents') income distribution.

[place figure 5 here]

Figure 5 includes the nonparametric elasticity and the correlation curve for the male middle-aged sample when labour income is used together with fathers' income. The nonlinear elasticity is again found to be well above the correlation curve. The correlation gradually increases for higher deciles of fathers' income and the highest correlation is found for deciles 4 to 8. After the eighth decile the correlation drops substantially. As noted earlier, for the young sample the highest correlation is found between deciles 1 and 6. Despite these differences in shape of the point estimate of the correlation curve, the overall impression that the correlation is found to be lowest in the tails of the distribution is repeated in both samples.

[place figure 6 here]

Figure 6 shows the corresponding graphs for the middle-aged sample when disposable income is used with parents' income. The general pattern that was shown in Figure 2, when the young sample was analysed, is repeated, i.e., the correlation is higher for higher parental income. If Figure 6 is compared to Figure 5, when labour income is used, the correlation at the eighth decile is again found to be statistically significantly higher for disposable income. For lower deciles, i.e. from the first to the fifth decile, the results are more similar.

[place figure 7 here]

Figure 7 shows the results for labour income for the middle-aged female sample. Compared to the young sample, the correlation is lower for the first deciles. For higher deciles the correlation increases, but it is nevertheless found to remain at a quite low level.

[place figure 8 here]

Figure 8 includes the results when disposable income is used. The correlation for the female middle-aged sample is higher for higher parental income. The magnitude of the correlation curve is substantially lower than that shown in Figure 4, where the results for the young sample are included. For example, at the median, the confidence bands do not overlap and it is clear that the curves are statistically significantly different. Despite that, at higher deciles the correlation is found to be stronger for disposable income than for labour income.

A few general patterns can be outlined from the previous results. The correlation between disposable income and parental income is higher for higher deciles of the income distributions of the parents. The correlation is particularly low for the first deciles. When labour income is analysed the pattern is a little bit different. Low correlation is found for both tails of the distribution, but the correlation can be already higher at the second decile. These differences are found for both male and female samples, and also for samples of different ages. Even though analysing the reasons for these differences is beyond the scope of this article, a sensitivity analysis has been carried out.

4.3 Using quantile regression

Eide & Showalter (1999) use quantile regression to estimate the intergenerational income elasticity at different quantiles of the conditional earnings distribution, i.e. the earnings distribution of the adult child. Other studies that have used this method is Grawe (2004) and Bratberg *et al*. (2007). Table A1 includes the elasticity for different deciles. For labour income the elasticity is very high for the lowest decile. It gradually decreases until reaching the median. For higher deciles the elasticity is slightly higher than for the median, but very far from what is found for the first deciles. The general pattern of higher elasticity for lower deciles is a common result (Bratberg *et al*., 2007). For example, Eide & Showalter (1999) conclude "These results suggest that family income is less important as an explanatory variable at the upper end of the conditional earnings distribution than it is at the bottom tail of the distribution." Bratberg *et al*. (2007) make a similar conclusion, "implying that earnings persistency among high-earning offspring is lower than among low-earning offspring", when referring to the results. Despite observing a similar pattern we should not make the same conclusion for the present study. When labour income is analysed for the male young sample the elasticity is 0.61 at the first decile and 0.36 at the second decile. It is clear that the lower tail of the distribution is important for the elasticity of 0.3445 presented in table 2. But, should we interpret this elasticity as a sign of high persistence? The answer is no. Table 1 shows that the distributions of fathers' income and sons' income is fairly similar for percentile 30, 50, 70 and 90. The large difference in standard deviation clearly is due to the lower tail of the distribution, where the dispersion is much higher for the adult children. This is the reason for the high elasticity in table 2 and the high elasticity for lower deciles in table A1. A high regression slope does not necessarily mean a high degree of association. It is easy to see, in figure 1, that many adult children below the first decile (i.e. log income < 11.00), actually has a father above the 9th decile, (i.e. log income > 12.75)! In fact, the difference of the average log income of the fathers of children with log income below, respective above 11.00, i.e., decile 1, is only 0.11. The difference of the median log income is only 0.08. To further analyse the degree of the relation, conditioning on the income distribution of the adult child, figure A1 includes the corresponding correlation curve. Note that the local correlation now is expressed at different percentiles of the income distribution of the adult child to compare the correlation with the elasticity found using quantile regression. The difference compared to table A1 is striking. While the elasticity is 0.61, 0.36 and 0.26 for the first three deciles when quantile regression is used, figure A1 shows that the correlation is not even significantly different from zero for the same deciles!

The correlation curve clarifies the result found using quantile regression. While the elasticity is high for lower deciles we cannot draw the conclusion that the income persistence is high. Any result using quantile regression should be accompanied with a careful analysis of the income distributions. If this is not done, the conclusion could be severally incorrect. It is also clear that quantile regression is not an alternative to the approach introduced in the present article.

4.4 Using different income measures

The first sensitivity analysis (A) was to use labour income in combination with parents' income, instead of only fathers' income, to evaluate the importance of using more complete information about the economic situation.

Note that the measure of disposable income contains income from other household members, and also the size and composition of the household are taken into account. Taxes and benefits also affect the income measure. Therefore, several factors, including effects from the welfare state and household composition, are different compared to when labour income is used as the individual outcome. The second sensitivity analysis (B) used another measure of disposable income. The alternative measure only includes individual income after taxes and benefits. In both these cases parental income is, accordingly, used.

The third sensitivity analysis (C) maintained the earlier measure of disposable income but used the income of the father instead of the parents. To reduce the time for computation, the sensitivity analyses were conducted based on 25% random samples of the middle-aged samples used earlier. This implies that about 8400 to 9200 observations were used, but the results were obtained within a few days instead of over a month (for each sample). The disadvantage is, of course, loss of precision and wider confidence intervals.

Figure A2 includes results from sensitivity analysis A) and is included in the Appendix. The results for both the male and female samples are included and these should be compared to Figures 5 and 6, and 7 and 8, respectively. Changing from the father's income to the parents' income, means, for the male sample, that the shape of the correlation curve becomes more linearly increasing over the income distribution of the parents. The magnitude of the correlation, in the upper part of the distribution, is, however, similar to when the father's income is used, as found in Figure 5. The relatively high correlation found in Figure 6, when disposable income is used, does not seem to be due to the use of parental income. This observation is also found for the female sample if the results are compared to Figure 8. The result for the female sample is, in fact, very similar in shape and magnitude to what is found in Figure 7, when the father's income is used.

Figure A3 includes the results for an alternative measure of disposable income and when parental income is maintained in the analysis. The results are very similar to Figures 6 and 8, when disposable income also includes incomes from other household members and the composition of the household is taken into account. The relatively high correlation found for the eighth decile in Figures 6 and 8 does not seem to be due to taking into account the household composition. The point estimate for the correlation is, in fact, found to be even higher for the alternative measure of disposable income.

Figure A4 includes the results for the combination of disposable income and the father's income. Using the father's income seems to make the shape a more inverted U compared to Figure 6, where an increasing correlation is found over the distribution, apart for the last decile. The inverted-U-shaped form is also found in Figure 5, where labour income is used, but there the correlation is almost constant and about 0.2 from decile 3 to decile 8. The magnitude is found to be higher for disposable income and at the seventh decile it is above 0.35. For the female sample the shape is similar to the result in Figure 7, where labour income is used. The highest correlation does, however, reach a higher level when disposable income is used, as shown in Figure A4.

An overview of the results, including the sensitivity analysis, indicates that using fathers' income instead of parental income makes the correlation curve a more inverted U shape, particularly for the male sample. Using parental income implies a more linearly increasing correlation over the income distribution of the parents, i.e. the correlation is strongest for the highest deciles. The correlation is, in general, stronger when disposable income is used compared to labour income. This concerns both individual disposable income and individualized disposable income in the household, where incomes from other household members are also included.

4.5 Life-cycle bias and the measure of permanent income

Table A2, found in appendix, repeats table 2, but for alternative specifications to analyse the differences found for the young samples and the middle-aged samples. For the young sample the fathers' or parents' income is measured in 1971, 1974 and 1977, implying an average of three years instead of five. The idea is to make the results more comparable to the middle-aged sample. The elasticities are reduced, as expected, and are now closer to what is found for the middle aged sample. The correlation is also slightly smaller, but the percentage change is smaller.

The average age of the fathers/parents is quite high for the middle aged sample. Gouskova *et al*. (2010) recommend using a current income around the age of 40 as an approximation of lifetime earnings. The result in table A2 is for the sample when the father's age is restricted to be below 50 years in 1974. For these samples the average age for the individuals is 40.7 and 41.6-41.7 for the fathers. This restriction increases the elasticity, but the correlation coefficient changes only marginally. The reason is that the standard deviation is reduced for the income of the fathers/parents. Note that the sample size is reduced substantially. This way to control for the age of the father is very convenient when the sample size is large. If a small sample is used it is, of course, not possible to simply restrict the sample in this way. An option is to estimate a semiparametric model, following Robinson (1988), assuming that additional covariates are included linearly. The correlation curve can be estimated for the nonparametric part, once the other variables have been conditioned out. Note that the literature on how to measure intergenerational income mobility seldom includes more covariates than age (and its square) of the individuals and the fathers. An exception is of course studies that aim to find the mitigating effects of an observed intergenerational correlation (Blanden *et al*., 2007).

Figures A5 and A6 show the correlation curves for labour income and disposable income for the middle-aged samples with a restriction to only include fathers who were less than 50 years old in 1974. The point estimate of the correlation curve for labour income for the male sample is below the corresponding curves in figure 1 and figure 5 for approximately decile 2 to 5, and the curve is more increasing over the distribution of income of the fathers. Having too young individuals (as in the young sample) or too old fathers (as in the middle aged sample) seems to exaggerate the correlation for these deciles. The results when the age of the father is restricted for the middle-aged sample is, in fact, similar to using parents' incomes instead of only the fathers' income, as showed in figure A2. Imposing the restriction of the age of the fathers for the female middle-aged sample actually makes the correlation curve more similar to the young sample, as shown in figure 3. In this case the correlation seems to be underestimated for lower deciles for the middle-aged sample when the fathers are relatively old.

For the male sample, the effect of restricting the age of the fathers when disposable income is analysed is a slightly weaker correlation compared to figures 2 and 6. For the female sample the restriction reduces the correlation, in particularly for higher deciles, and the tendency is more inverted U-shaped relation. Note, however, that in all these cases a hypothesis of a linear correlation cannot be rejected. Obviously, restricting the age of the fathers has reduced the sample size substantially and the confidence intervals are wider.

5 Concluding remarks

The literature on intergenerational income mobility has been dominated by summary measures such as the correlation and elasticity of adult child income with respect to fathers' income. An advantage with the correlation coefficient is that the dispersion has been standardized. An increased income inequality would (if it increases the standard deviation), for example, result in a higher elasticity, while the correlation would not be affected. A significant problem with the correlation coefficient is, however, that it does not capture different degrees of association over the distribution. In this study, correlation curves are introduced to measure intergenerational income mobility. The results indicate that irrespective of what income measure that is used, the correlation is found to be nonlinear when a large sample is analysed. Using the correlation coefficient is not enough to give a representative measure for the correlation at different parts of the distribution. For example, using the correlation coefficient underestimates the correlation at the eighth decile by about 50% when disposable income is analysed. For lower deciles the correlation is overestimated.

The results indicate fairly high intergenerational income mobility at the lower part of the distribution. This result is similar to what Bratsberg *et al*. (2007) found for Denmark, Finland and Norway. It is possible that the Nordic welfare state, with its highly redistributive educational policies, could have an important role to play in explaining the pattern.

Particularly for higher deciles of the fathers'/parents' income distribution the results show that the degree of intergenerational income correlation is higher for disposable income than for labour income. If nonlinear intergenerational income elasticity were to be used to measure the intergenerational income mobility the opposite conclusion would be reached. The reason is that the dispersion of labour income is substantially higher than the dispersion of disposable income and this inflates its elasticity. Comparing elasticities from nonparametric models can be informative regarding the pattern over the distribution, but the magnitude of different samples is sensitive to differences in the dispersion of the distributions. For example, the elasticity for the middle-aged sample is almost 0.45 at the eighth decile for the male sample when labour income is analysed. The corresponding elasticity for the middle-aged sample, at the same decile, but when disposable income is analysed, is less than 0.3. Despite this, the correlation is stronger in the latter case, with a correlation of 0.3 compared to a correlation of only 0.2 when labour income is studied. Therefore, the elasticity suggests that the labour income is transmitted with much higher persistence among generations than the disposable income. The correlation curve effectively clarifies that this result is due to the different dispersion that inflates the elasticity when labour income is used. The *degree* of relation is, in fact, stronger at higher deciles for disposable income than for labour income. It is important to remember that the elasticity does not measure the *degree* of a relation, and a comparative study would benefit from using the correlation curve to measure intergenerational income mobility. The results are otherwise sensitive to differences in income dispersion of the two generations, which could be due to life cycle difference, changes in society over time, or simply different definitions of the income variables in the two generations.

Applying the correlation curve to studying intergenerational income mobility is particularly useful for making cross-country comparisons. Firstly, the correlation curve captures different mobility in different parts of the distribution. Different countries could have different patterns of the intergenerational income mobility over the income distribution of the parents. Secondly, the correlation curve is not, in contrast to the elasticity from a nonparametric regression, sensitive to different dispersion in the two generations. In addition, if mobility matrices or other rank-based measures of mobility are used, a country with high income dispersion would have a greater (monetary) difference from one income category to the next, for example from the second decile to the third decile. Therefore, the income could be fairly different without resulting in a changed category, and, accordingly, a higher inequality would automatically imply lower mobility. This problem is avoided with the correlation curve, and international comparisons can be made more easily. Note that the correlation curve can be implemented using a linear regression, adding higher-order polynomial terms of the fathers' income, instead of working with a nonparametric regression that requires large data sets (Blyth, 1994). A disadvantage is of course that this imposes a restriction of the functional form that could be fulfilled to different degrees for different samples. The correlation curve is easy to interpret, and since it is a scale-free measure it is, in fact, a highly useful tool for making comparisons, not only for measuring intergenerational income mobility, but for a wide range of empirical topics.

Acknowledgments

The author gratefully acknowledges the financial support from the Spanish Ministry of Science and Innovation (grant #ECO2010-20829).

References:

Aaronson, D., Mazumder, B., 2008. Intergenerational Economic Mobility in the United States, 1940 to 2000. *Journal of Human Resources* 43, 139-172. DOI: 10.3368/jhr.43.1.139

Bjerve, S., Doksum, K., 1993. Correlation Curves: Measures of Association as Functions of Covariate Values. *The Annals of Statistics* 21, 890-902.

Björklund, A., Chadwick, L., 2003. Intergenerational income mobility in permanent and separated Families. *Economics Letters* 80, 239-246. DOI: 10.1016/S0165-1765(03)00084-3

Björklund, A., Roine, J., Waldenström, D., 2012. Intergenerational top income mobility in Sweden: Capitalist dynasties in the land of equal opportunity? *Journal of Public Economics* 96, 474-484. DOI: 10.1016/j.jpubeco.2012.02.003

Blyth S., 1994. Karl Pearson and the correlation curve. *International Statistical Review* 62, 393-403.

Blanden, J., 2011. Cross-country rankings in intergenerational mobility: a comparison of approaches from economics and sociology. *Journal of Economic Surveys*. DOI: 10.1111/j.1467-6419.2011.00690.x

Blanden, J., Gregg, P., Macmillan, L., 2007. Accounting for intergenerational income persistence: noncognitive skills, ability and education. *The Economic Journal* 117, C43-C60.

Bratberg, E., Nilsen, O.A., Vaage, K., 2005. Intergenerational earnings mobility in Norway: Levels and trends. *Scandinavian Journal of Economics* 107, 419-435. DOI: 10.1111/j.1467-9442.2005.00416.x

Bratberg, E., Nilsen, O.A., Vaage, K., 2007. Trends in intergenerational mobility across offspring's earnings distribution in Norway. *Industrial Relations* 46, 112-129. DOI: 10.1111/j.1468-232X.2007.00459.x

Bratsberg, B., Roed, K., Raaum, O., Naylor, R., Jäntti, M., Eriksson, T., Osterbacka, E., 2007. Nonlinearities in intergenerational earnings mobility: Consequences for cross-country comparisons. *The Economic Journal* 117, C72-C92. DOI: 10.1111/j.1468-0297.2007.02036.x

Corak, M., Heisz, A., 1999. The intergenerational Earnings and Income Mobility of Canadian Men. *Journal of Human Resources* 34, 504-533.

Couch, K.A., Lillard, D.R., 1998. Sample selection rules and the intergenerational correlation of earnings. *Labour Economics* 5, 313-329.

Dearden, L., Machin, S., Reed, H., 1997. Intergenerational Mobility in Britain. The *Economic Journal* 107, 47-66.

Doksum, K., Blyth, S., Bradlow, E., Meng, X., Zhao, H., 1994. Correlation Curves as Local Measures of Variance Explained by Regression. *Journal of the American Statistical Association* 89, 571-582.

Eide, E.R., Showalter, M.H., 1999. Factors affecting the transmission of earnings across generations: a quantile regression approach. *The Journal of Human Resources* 34, 253-267.

Fan, J., 1992. Design-adaptive Nonparametric Regression. *Journal of the American Statistical Association* 87, 998-1004.

Fan, J., 1993. Local Linear Regression Smoother and Their Minimax Efficiency. *The Annals of Statistics* 21, 196-216.

Fan, J., Gijbels, I., 1995a. Data-Driven Bandwidth Selection in Local Polynomial Fitting: Variable Bandwidth and Spatial Adaptation. *Journal of the Royal Statistical Society*. Series B (Methodological) 57, 371-394.

Fan, J., Gijbels, I., 1995b. Adaptive Order Polynomial Fitting: Bandwidth Robustification and Bias Reduction. *Journal of Computational and Graphical Statistics* 4, 213-227.

Fertig, A.R., 2003. Trends in intergenerational earnings mobility in the United States. *Journal of Income Distribution* 12(3-4), 108-130.

Galindo, C.D., Liang, H., Kauermann, G., Carrol, R.J., 2001. Bootstrap Confidence Intervals for Local Likelihood, Local Estimating Equations and Varying Coefficient Models. *Statistica Sinica* 11, 121-134.

Gouskova, E., Chiteji, N., Stanford, F., 2010. Estimating the intergenerational persistence of lifetime earnings with life course matching: Evidence from the PSID. *Labour Economics* 17, 592-597. DOI: 10.1016/j.labeco.2009.04.009

Grawe, N.D., 2004. Reconsidering the Use of Nonlinearities in Intergenerational Earnings Mobility as a Test for Credit Constraints. *Journal of Human Resources* 39, 813-827. DOI: 10.3368/jhr.XXXIX.3.813

Grawe, N.D., 2006. Lifecycle bias in estimates of intergenerational earnings persistence.

Labour Economics 13, 551-570. DOI: 10.1016/j.labeco.2005.04.002

Haider, S., Solon, G., 2006. Life-cycle variation in the association between current and lifetime earnings. *American Economic Review* 96, 1308-1320. DOI:10.1257/aer.96.4.1308

Härdle, W., Mammen, E., 1993. Comparing Nonparametric Versus Parametric Regression Fits. The Annals of Statistics 21, 1926-1947. DOI: 10.1214/aos/1176349403

Jenkins, S., 1987. Snapshots versus Movies: 'Lifecycle Biases' and the Estimation of Intergenerational Earnings Inheritance. *European Economic Review* 31, 1149-1158. DOI: 10.1016/S0014-2921(87)80010-7

Jäntti, M., Bratsberg, B., Roed, K., Raaum, O., Naylor, R., Osterbacka, E., Björklund, A., Eriksson, T., 2006. American Exceptionalism in a New Light: A Comparison of Intergenerational Earnings Mobility in the Nordic Countries, the United Kingdom and the United States. IZA DP No. 1938.

McMurry, T., Politis, D., 2004. Nonparametric regression with infinite order flat-top kernels. *Journal of Nonparametric Statistics* 16, 549-562. DOI: 10.1080/10485250310001622596

Newey, W., Hsieh, F., Robins, J., 2004. Twicing Kernels and a small Bias Property of Semiparametric Estimators. *Econometrica* 72, 947-962. DOI: 10.1111/j.1468-0262.2004.00518.x

Nilsson, W., del Barrio Castro, T., 2012. Bootstrap Confidence Interval for a Correlation Curve. *Statistics and Probability Letters* 82, 1-6. DOI: 10.1016/j.spl.2011.09.001

Robinson, P.M., 1988. Root-n-consistent semiparametric regression. *Econometrica* 56, 931-954.

Ruppert, D., 1997. Empirical-Bias Bandwidths for Local Polynomial Nonparametric

Regression and Density Estimation. *Journal of the American Statistical Association* 92, 1049-1062.

Ruppert, D., Wand, M.P., 1994. Multivariate Locally Weighted Least Squares Regression. *The Annals of Statistics* 22, 1346-1370. DOI:10.1214/aos/1176325632

Scott, D.W., 1992. *Multivariate Density Estimation*. Wiley: New York

Solon, G., 1992. Intergenerational income mobility in the United States. *American Economic Review* 82, 393–408.

Solon, G., 1999. Intergenerational Mobility in the Labor Market. In: Ashenfelter, O., Card, D. (eds). *Handbook of Labor Economics*, vol. 3, pp. 1761–1800. Elsevier: Amsterdam.

Zimmerman, D.J., 1992. Regression toward Mediocrity in Economic Stature. *American Economic Review* 82, 409-429.

Österbacka, E., 2001. Family Background and Economic Status in Finland. *Scandinavian Journal of Economics* 103, 467–484.

Österberg, T., 2000. Intergenerational Income Mobility in Sweden: What Do Tax-data Show? *Review of Income and Wealth* 46, 421-436. DOI: 10.1111/j.1475-4991.2000.tb00409.x

Appendix

Nonparametric technique

The nonparametric method used in this study is local polynomial regression. For x in a neighbourhood of x_0, $m(x)$ is approximated locally by a polynomial of order p:

$$m(x) \approx m(x_0) + m'(x_0)(x - x_0) + ... + m^{(p)}(x_0)(x - x_0)^{(p)} / p! \qquad (A1)$$

The bandwidth, h, controls the size of the local neighbourhood and $K(x)$ is a kernel function that weights the data points closer to x_0 more heavily. The solution to the weighted polynomial regression is: $\hat{\boldsymbol{\beta}} = (\mathbf{X}^{\mathrm{T}}\mathbf{W}\mathbf{X})^{-1}\mathbf{X}^{\mathrm{T}}\mathbf{W}\mathbf{y}$. $\boldsymbol{\beta} = (\beta_0,..., \beta_p)^T$ is the nonparametric estimate of the regression function and its derivatives, $v!\hat{\beta}_v = m^{(v)}(x_0), v = 0,..., p.$ $\mathbf{W}$ is a diagonal matrix with $W_i = K((x_i - x_0)/h)$. $\mathbf{X}$ is a design matrix with (l, j)-th element as $(x_l - x_0)^{j-1}$ and $\mathbf{y} = (y_1,..., y_n)^T$. The standardized residual variance is:

$$\hat{\sigma}^2(x_0) = \frac{1}{tr\left\{\mathbf{W} - (\mathbf{X}^{\mathrm{T}}\mathbf{W}\mathbf{X})^{-1}\mathbf{X}^{\mathrm{T}}\mathbf{W}^2\mathbf{X}\right\}} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 K\left(\frac{x_i - x_0}{h}\right) \qquad (A2)$$

where $\hat{\mathbf{y}} = (\hat{y}_1,..., \hat{y}_n)^T = \mathbf{X}\hat{\boldsymbol{\beta}}$. To estimate the model it is necessary to choose which kind of kernel to use, what order of polynomial to use and the size of the bandwidth. Regarding the choice of kernel, there is a wide variety of different kernels to use: Gaussian, Uniform, Epanechnikov, Biweight and Triweight, to mention but a few. The choice of kernel has received less attention in the literature than the choice of bandwidth.

An important branch of the literature has, however, focused on the importance of the order of the kernel. In particular, it has been shown that the bias can be reduced by choosing a kernel of sufficient high order compared to the order of the function that is estimated (see Newey, Hsieh & Robins, 2004 and McMurry & Politis, 2004).

A higher order of polynomial has a few important implications for the estimator. A higher-order polynomial reduces the bias, but this comes at the cost of increased variability. Note that the variability increases when going from an odd-order to an even-order polynomial, while the asymptotic variance is kept constant when going from an even-order polynomial to the consecutive odd-order polynomial. For this reason Fan & Gijbels (1995b) recommend using an odd-order polynomial to estimate the regression function. Ruppert & Wand (1994) show that $p - v$ should be odd, and as a consequence a second-order polynomial ($p = 2$) can be used if the purpose is to estimate the first derivative ($v = 1$) of the regression function. A disadvantage of using a higher-order polynomial is that problems of singularity could occur due to the sparseness of the data points in combination with a too-small bandwidth.

The size of the bandwidth has a similar trade-off between bias reduction and variability. A large bandwidth reduces the variance, but it comes at the cost of increased bias. A small bandwidth reduces the bias while increasing the variance. Several data-driven procedures are available for finding an optimal bandwidth. Fan & Gijbels (1995a) and Ruppert (1997) are two important references. The main idea is to choose a bandwidth that minimizes an estimated mean-squared error (MSE) function. The estimate of the variance is the same, but the estimate of the bias differs for each method. Both these methods can be used to find an optimal bandwidth to estimate the regression function as well as its derivatives. These methods can be used as a guidance to choose the bandwidth when the main objective is the *correlation curve*, but it is no guarantee that the optimal choice has been made. Nilsson & del Barrio (2012) show in a simulation study that using a bandwidth optimal for the first derivative and using the median among bootstrap replications as a point estimate works well for regression functions without abrupt changes in the curves.

Selection of bandwidth

The data-driven procedure suggested by Fan & Gijbels (1995a) to find the optimal bandwidth for the second derivative is used. Before starting the procedure a short algorithm was used to avoid sparse data automatically forcing too large a bandwidth. In the C-code available at Professor J. Fan's Web page, http://orfe.princeton.edu/~jqfan/fan/publications.html, Fan & Gijbels (1995a) include a restriction that counts the number of effective data points. The number of effective data points has to be at least equal to the order of the polynomial used. Therefore, $|x_0 - x| < h$, at least for 'order' observations. When the procedure aims to find the optimal bandwidth for a second-order polynomial regression, the 'order' is actually 4, since it is necessary to estimate a higher-order polynomial to evaluate the bias in the estimate of the MSE. If the effective data points are too few, the loop automatically chooses a larger bandwidth until the restriction is fulfilled for the complete sample. This means that a few sparse observations could force the optimal bandwidth to be the first bandwidth that becomes possible to estimate; hence it could be too large. This is inconvenient, as the bandwidth would be influenced by a few extreme outliers. To avoid this, the same restriction as suggested in the computer code was used before initiating the procedure to find the optimal bandwidth. 'h' was set to 0.1 and observations x_0 that did not fulfil the restrictions were dropped from the sample, and the same restriction was tested again. This sequence was used until the entire sample fulfilled the restriction. Between 9 and 38 observations were dropped before initiating the program. These observations are concentrated at the bottom and top tails of the distributions, and since the method estimates a local measure it is plausible to assume that dropping the observations is harmless.

The C-code suggested the starting bandwidth to be hmin = (xmax - xmin)*(order+2.0)/n, where xmax and xmin are the maximum and minimum values of fathers'/parents' income in the sample. With almost 40,000 observations this would be a very small value, and it would require a lot of time to reach the optimal bandwidth. The starting bandwidth for the procedure was, for this reason, set to 0.04, which, in the case of a second-order polynomial, in practice means an initial bandwidth of 0.08.

Kernel

For the nonparametric estimations, a Gaussian kernel was used:

$$K(x) = K\left(\frac{x_i - x}{h}\right) = \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{1}{2}\left(\frac{x_i - x}{h}\right)^2\right\}$$

Figures

All figures exclude 100 observations in each tail of the income distribution of the fathers/parents. The reason for this is that the derivative can be taken to very extreme values in the extremes of the distributions.

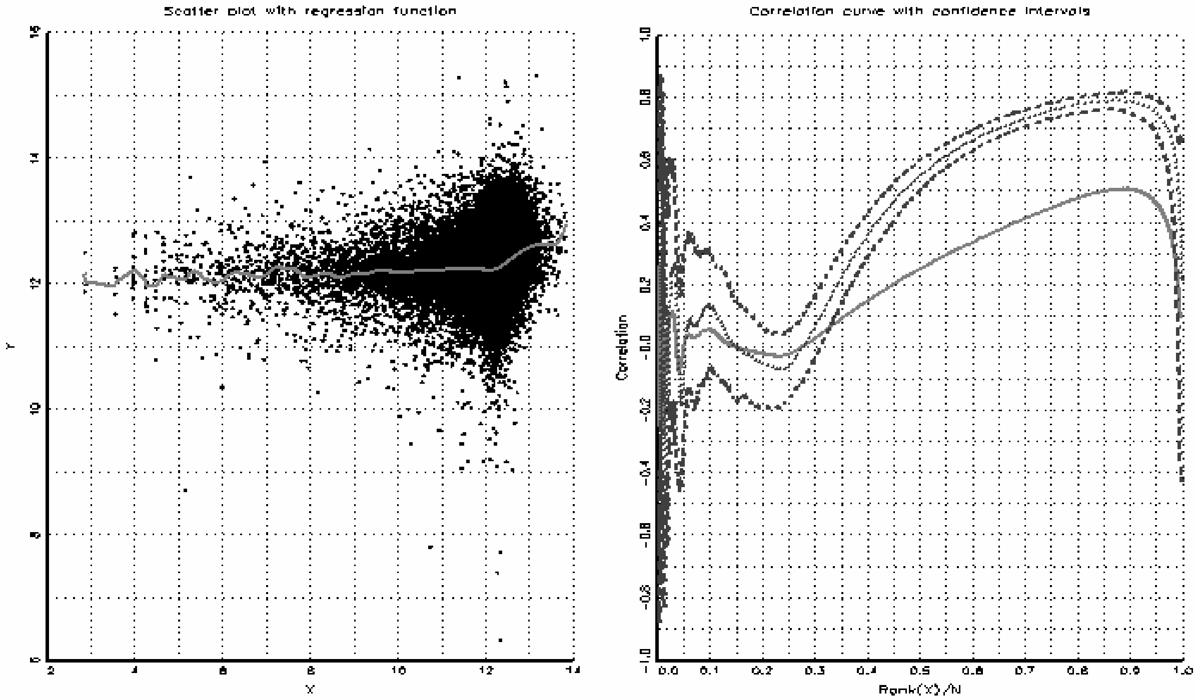

Figure A1. Scatter plot and correlation curve (and elasticity) for male young sample: labor income and father's income. The correlation is expressed conditional on the percentile in the income distribution of the *children*.

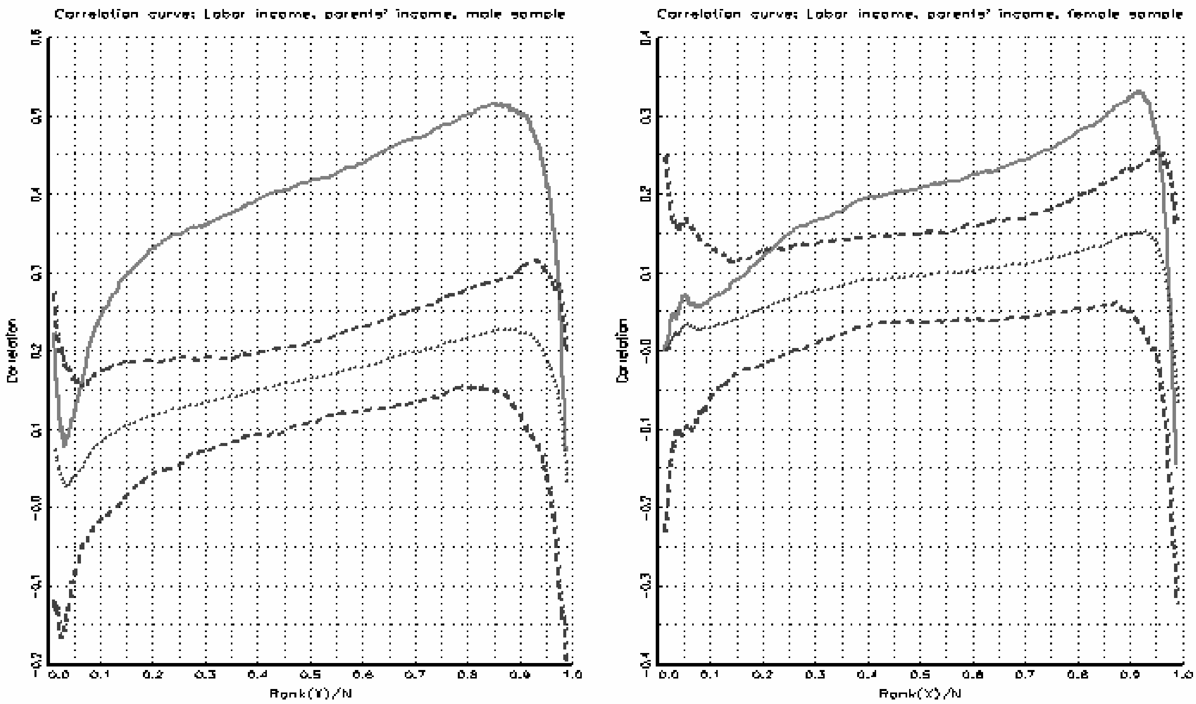

Figure A2. Correlation curves for 25% random middle-aged samples: labor income with parents' income.

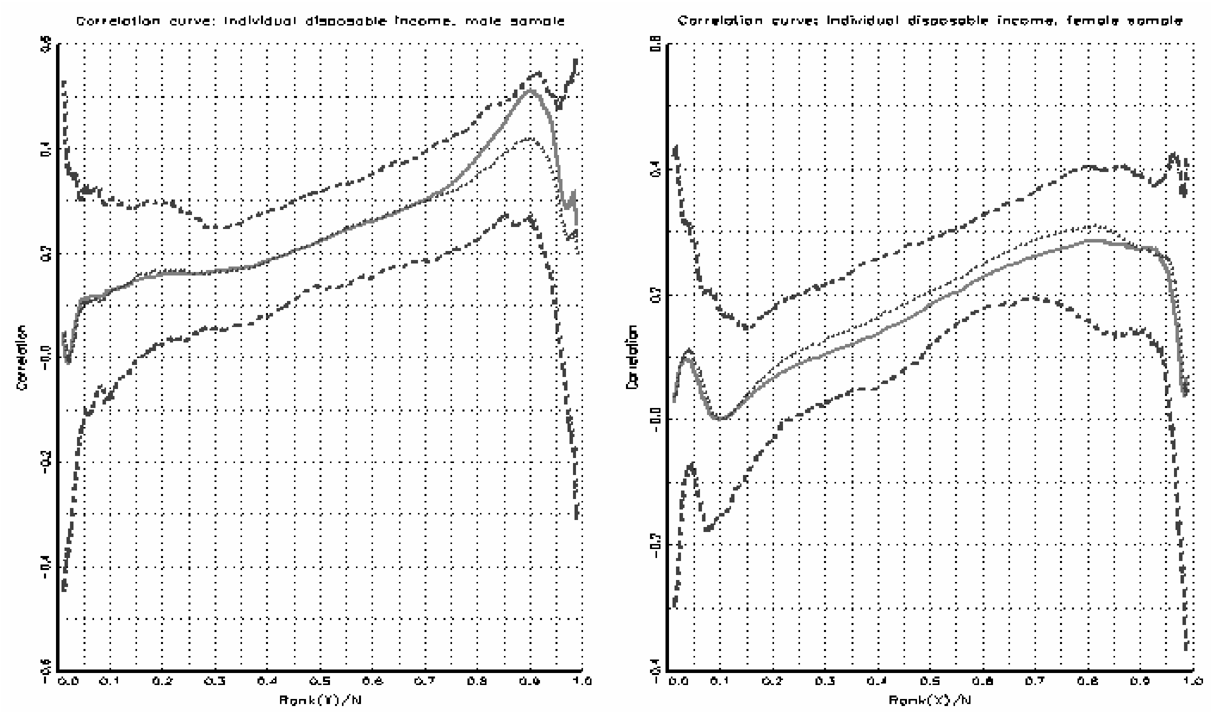

Figure A3. Correlation curves for 25% random middle-aged samples: individual disposable income with parents' income.

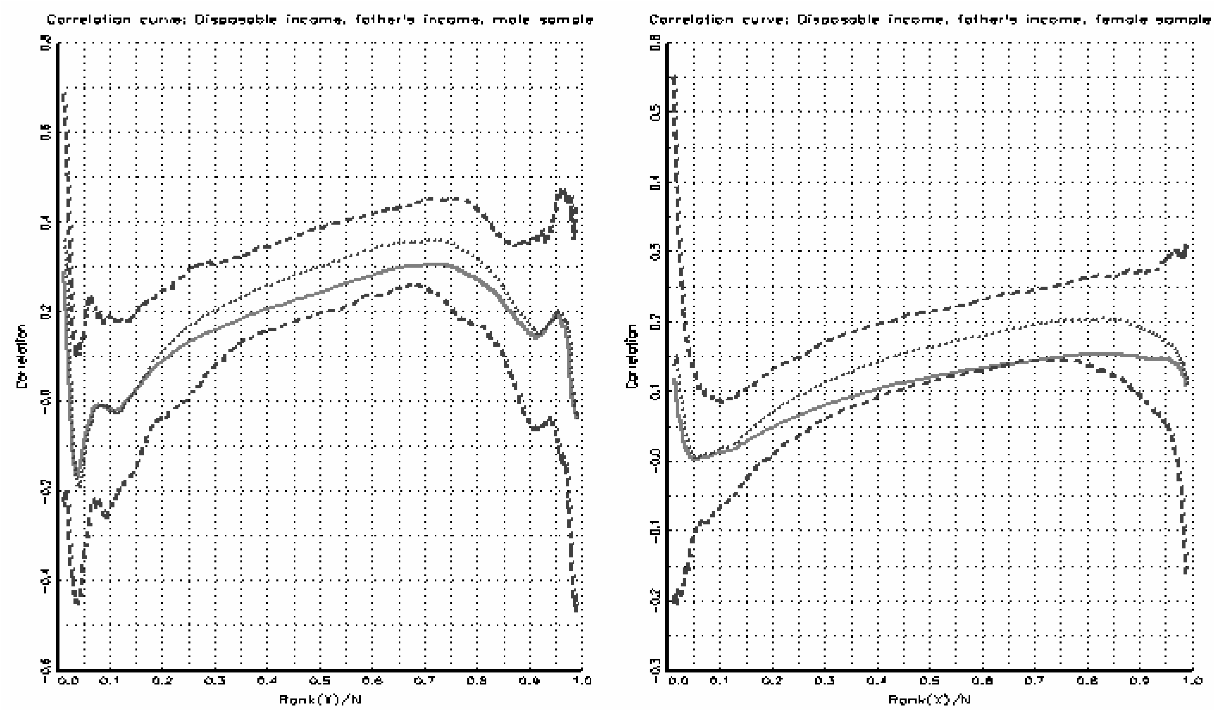

Figure A4. Correlation curves for 25% random middle-aged samples: disposable income with father's income.

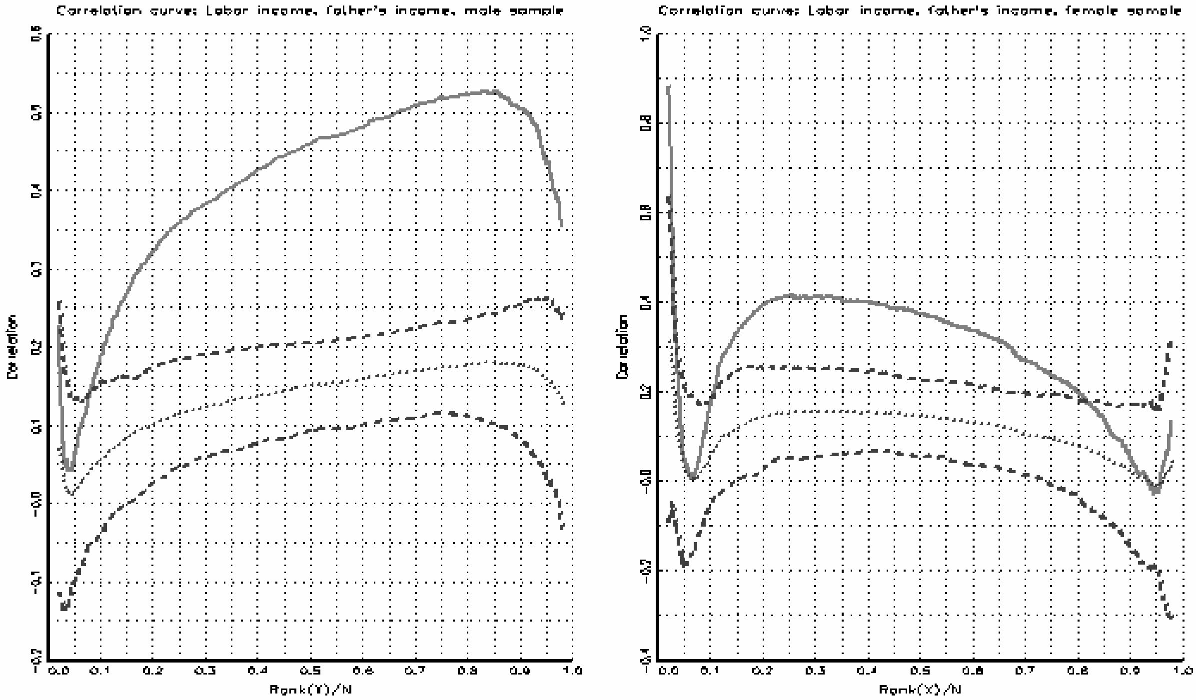

Figure A5. Correlation curves for middle-aged samples: labor income with father's income. Only fathers with age below 50 year are included.

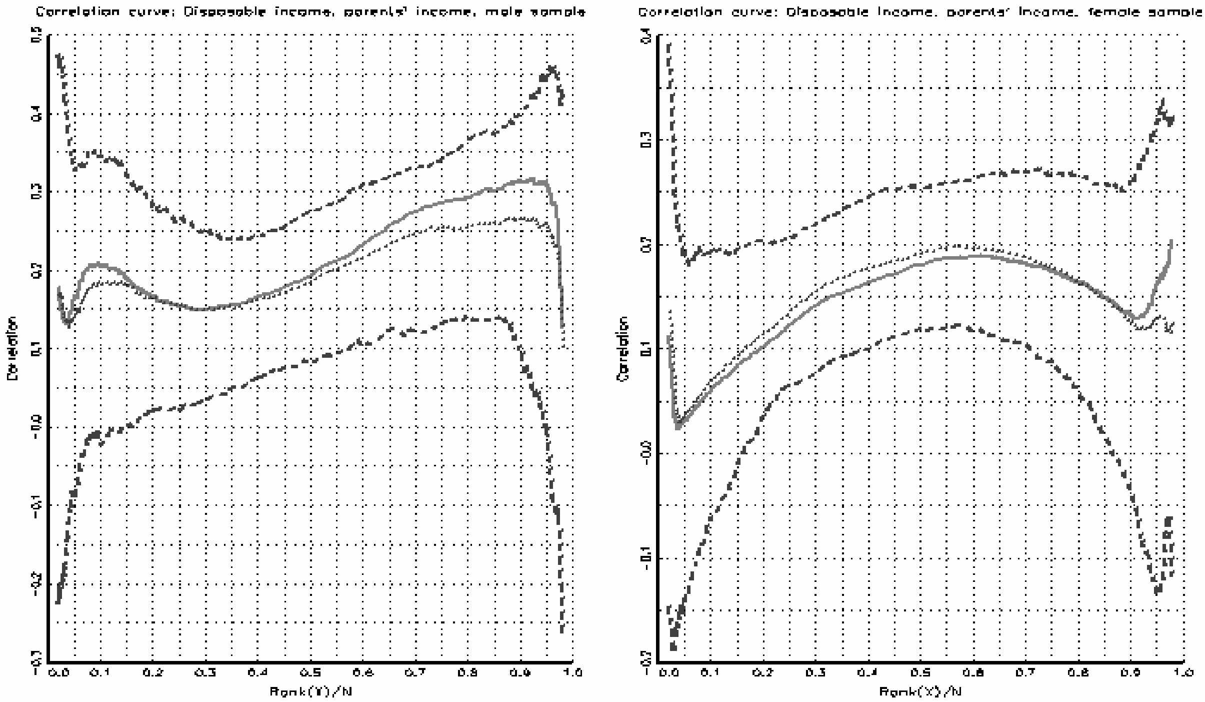

Figure A6. Correlation curves for middle-aged samples: disposable income with parents' income. Only fathers with age below 50 year are included.

Table A1. Quantile regression, regression slope

	P10	P20	P30	P40	P50	P60	P70	P80	P90
Income	Young Sample: Men								
labor	0.6112	0.3578	0.2564	0.2147	0.2051	0.2072	0.2166	0.2292	0.2372
disposable	0.1735	0.1808	0.1953	0.2096	0.2232	0.2333	0.2432	0.2478	0.2532
	Young Sample: Women								
labor	0.4528	0.3020	0.2225	0.1931	0.1762	0.1787	0.1845	0.1859	0.1958
disposable	0.1392	0.1511	0.1564	0.1739	0.1907	0.2073	0.2231	0.2342	0.2514
	Middle aged sample: Men								
labor	0.5319	0.3291	0.2051	0.1777	0.1732	0.1829	0.1941	0.2160	0.2385
disposable	0.1466	0.1328	0.1326	0.1391	0.1421	0.1492	0.1587	0.1663	0.1999
	Middle aged sample: Women								
labor	0.2270	0.1603	0.1139	0.1033	0.0997	0.0970	0.1042	0.1092	0.1163
disposable	0.0823	0.0755	0.0760	0.0783	0.0860	0.0951	0.1057	0.1186	0.1526

Table A2. Summary statistics, elasticities and correlation coefficients for alternative specifications

	Mean (Std. dev.) Individual income	Mean (Std. dev.) Father's/Parents' income	Elasticity	Std. error	Correlation	N
Young sample:	*Using years 71, 74, 77*					
Men, labor income	11.96 (1.00)	12.28 (0.43)	0.2996***	0.0113	0.1302***	40625
Men, disposable income	11.62 (0.38)	12.44 (0.41)	0.1933***	0.0044	0.2094***	41929
Women, labor income	11.42 (1.02)	12.28 (0.43)	0.2392***	0.0120	0.1017***	38098
Women, disposable income	11.49 (0.35)	12.45 (0.42)	0.1801***	0.0041	0.2146***	39253
Middle aged sample:	*Fathers' age <= 50*					
Men, labor income	11.96 (1.19)	12.16 (0.42)	0.3601***	0.0397	0.1285***	4906
Men, disposable income	11.52 (0.39)	12.59 (0.37)	0.1823***	0.0144	0.1731***	5224
Women, labor income	11.63 (1.08)	12.17 (0.43)	0.2483***	0.0358	0.0995***	4816
Women, disposable income	11.47 (0.36)	12.60 (0.38)	0.1433***	0.0135	0.1484***	5020

Notes: Restricting the fathers' age to 50 years or below implies and average age for the individuals of about 40.7 and for the fathers 41.6-41.7 for all of the samples.

Tables and figures to be inserted in main text:

Table 1. Summary statistics

	Years	Mean age	P10	P30	P50	P70	P90	Mean	Std. dev.	N
	Young Sample: Men, labor income									
Individual	1994-1999	31.5	11.00	12.00	12.23	12.39	12.65	11.96	1.00	40285
Father	1971, 74, 77, 80, 83.	42.2	11.87	12.11	12.24	12.41	12.75	12.27	0.41	
	Young Sample: Men, disposable income									
Individual	1994-1999	31.5	11.26	11.42	11.60	11.81	12.08	11.62	0.38	41371
Parents	1971, 74, 77, 80, 83.	Father: 42.3 Mother: 39.1	12.06	12.34	12.50	12.66	12.96	11.50	0.38	
	Young Sample: Women, labor income									
Individual	1994-1999	31.5	10.33	11.33	11.67	11.93	12.24	11.42	1.02	37706
Father	1971, 74, 77, 80, 83.	42.3	11.87	12.11	12.25	12.41	12.75	12.28	0.41	
	Young Sample: Women, disposable income									
Individual	1994-1999	31.5	11.12	11.32	11.46	11.63	11.91	11.49	0.35	38656
Parents	1971, 74, 77, 80, 83.	Father: 42.3 Mother: 39.1	12.06	12.35	12.51	12.67	12.69	12.51	0.39	
	Years	Mean age	P10	P30	P50	P70	P90	Mean	Std. dev.	N
	Middle aged Sample: Men, labor income									
Individual	1994-1999	42.9	10.96	12.12	12.33	12.51	12.86	12.04	1.21	34552
Father	1971, 74, 77.	52.3	11.56	11.98	12.16	12.34	12.72	12.14	0.53	
	Middle aged Sample: Men, disposable income									
Individual	1994-1999	42.9	11.12	11.37	11.54	11.75	12.07	11.57	0.42	36886
Parents	1971, 74, 77.	Father: 52.4 Mother: 49.0	11.92	12.31	12.54	12.74	13.05	12.51	0.49	
	Middle aged Sample: Women, labor income									
Individual	1994-1999	42.9	10.82	11.74	12.00	12.18	12.42	11.73	1.03	33776
Father	1971, 74, 77.	52.3	11.55	11.98	12.16	12.33	12.72	12.14	0.53	
	Middle aged Sample: Women, disposable income									
Individual	1994-1999	42.9	11.14	11.36	11.52	11.69	11.95	11.53	0.38	35404
Parents	1971, 74, 77.	Father: 52.3 Mother: 49.0	11.94	12.33	12.55	12.74	13.06	12.52	0.48	

Notes: When individual income is log labour income, fathers' income is combined labour and capital log income. When individual income is disposable income it is based on the disposable family income, where an equivalence scale has been used to individualize the income. 'Parents' income' refers to the combined labour and capital income for both parents. The incomes are measured in Swedish Crowns (krona) deflated to the price level of 2001. The exchange rate observed on 31st December 2001 can be used to get approximate numbers in euros (1 euro = 9.3029 SEK). Mean age refers to the age when income is measured.

Table 2. Summary statistics, elasticities and correlation coefficients

	Mean (Std. dev.) Individual income	Mean (Std. dev.) Father's/Parents' income	Elasticity	Std. error	Correlation	N
Young sample:						
Men, labor income	11.96 (1.00)	12.27 (0.41)	0.3445***	0.0120	0.1414***	40285
Men, disposable income	11.62 (0.38)	12.50 (0.38)	0.2317***	0.0047	0.2334***	41371
Women, labor income	11.42 (1.02)	12.28 (0.41)	0.2647***	0.0126	0.1077***	37704
Women, disposable income	11.49 (0.35)	12.51 (0.39)	0.2019***	0.0045	0.2229***	38656
Middle aged sample:						
Men, labor income	12.04 (1.21)	12.14 (0.53)	0.2930***	0.0122	0.1278***	34552
Men, disposable income	11.57 (0.42)	12.51 (0.49)	0.1665***	0.0045	0.1904***	36886
Women, labor income	11.73 (1.03)	12.14 (0.53)	0.1419***	0.0106	0.0725***	33776
Women, disposable income	11.53 (0.38)	12.52 (0.48)	0.1121***	0.0041	0.1429***	35404

Notes: Labor income is used together with the father's income and disposable income is used in combination with the parents' income. *** $p < 0.01$.

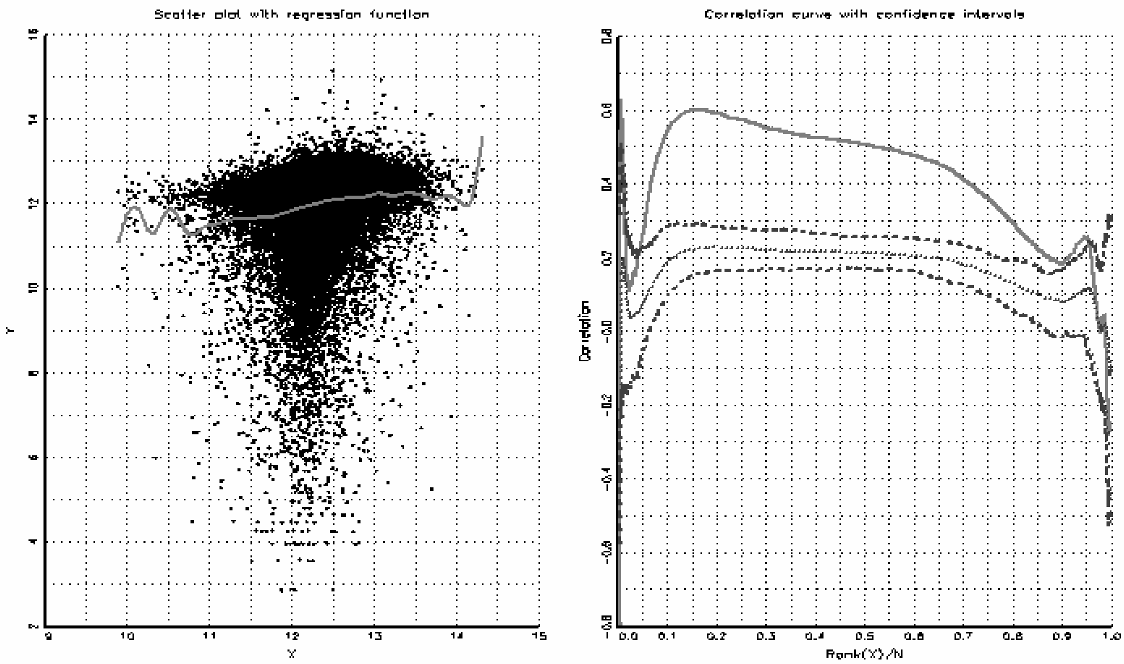

Fig. 1. Scatter plot and correlation curve (and elasticity) for male young sample: labor income and father's income

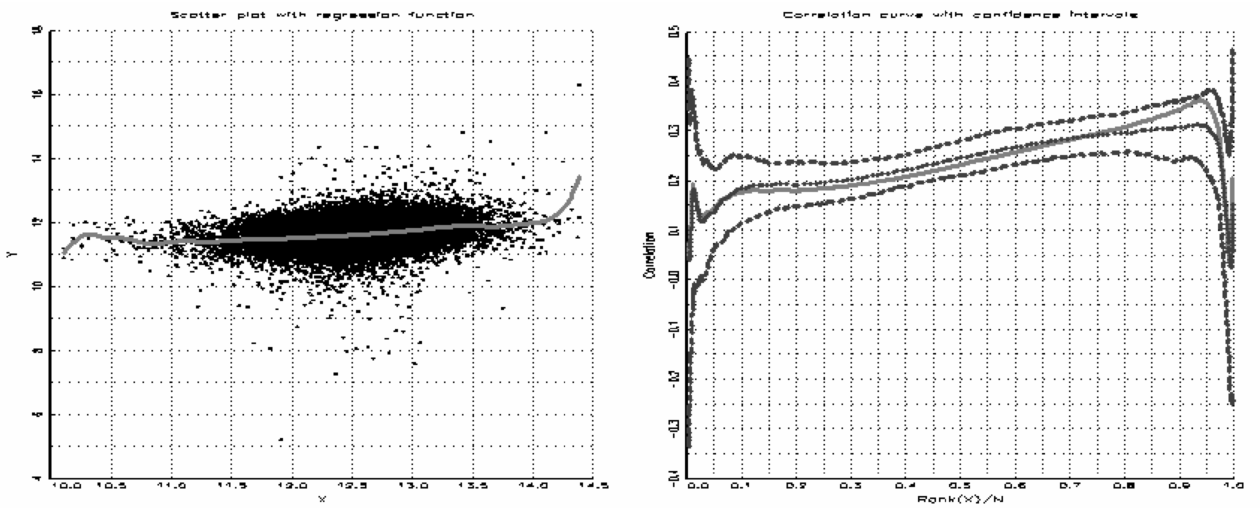

Fig. 2. Scatter plot and correlation curve (and elasticity) for male young sample: disposable income and parents' income

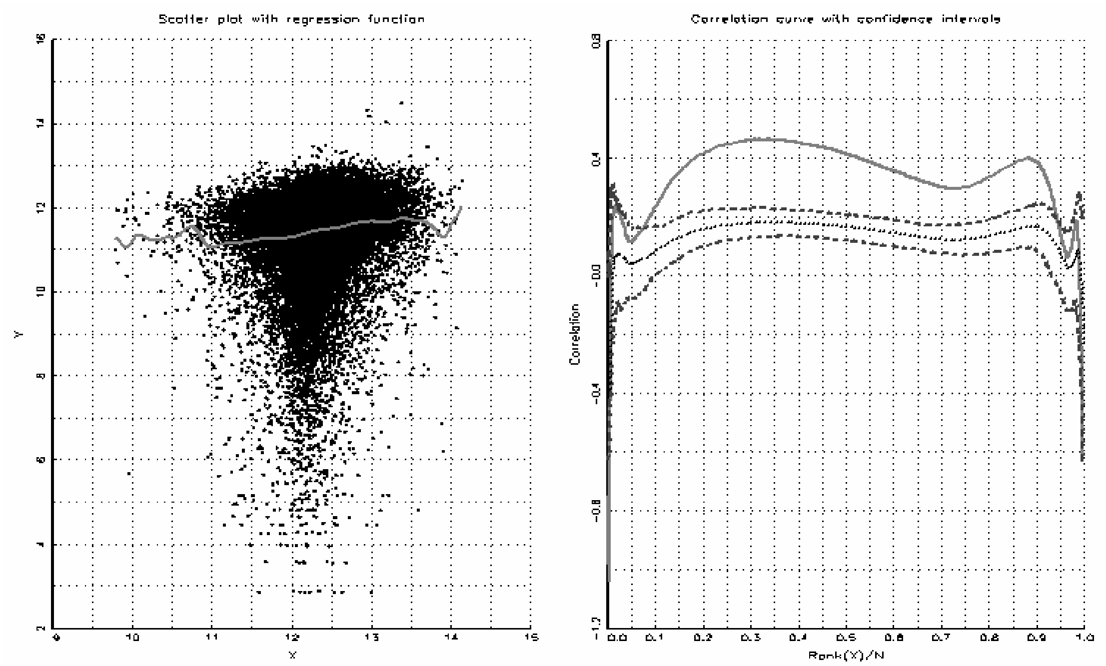

Fig. 3. Scatter plot and correlation curve (and elasticity) for female young sample: labor income and father's income

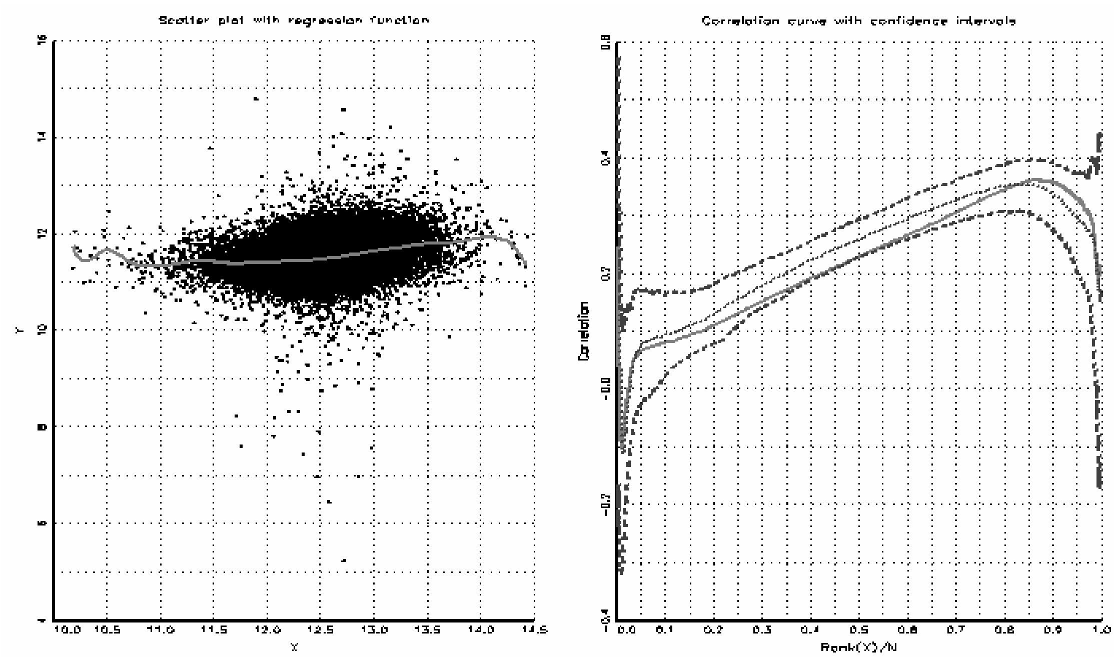

Fig. 4. Scatter plot and correlation curve (and elasticity) for female young sample: disposable income and parents' income

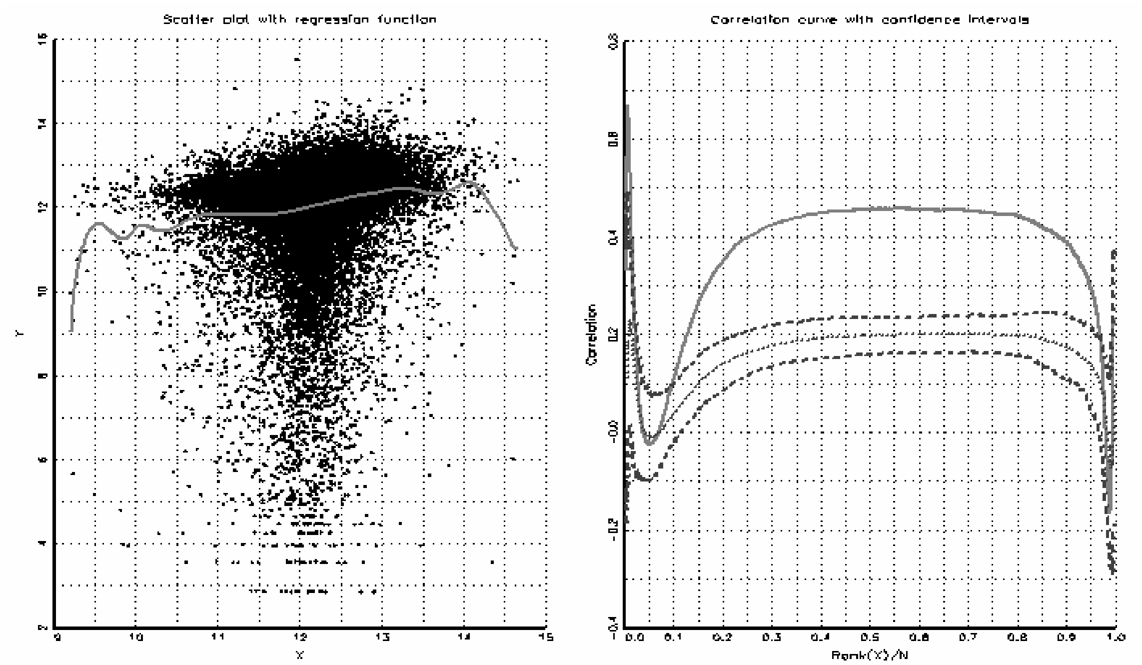

Fig. 5. Scatter plot and correlation curve (and elasticity) for male middle aged sample: labor income and father's income

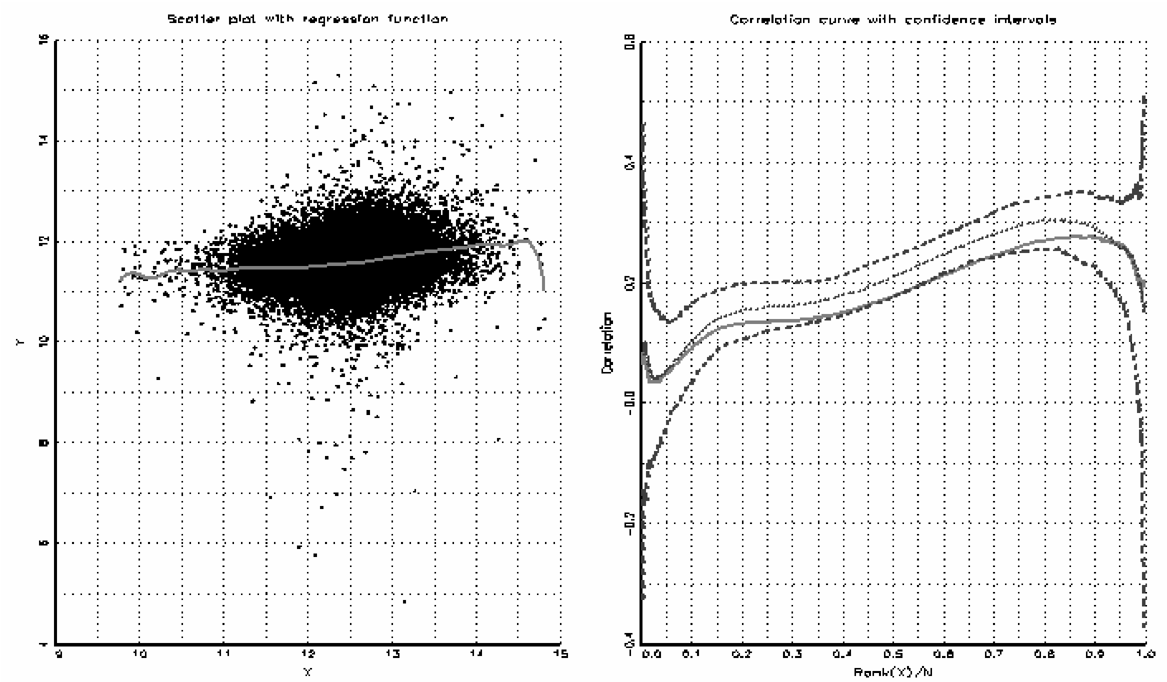

Fig. 6. Scatter plot and correlation curve (and elasticity) for male middle aged sample: disposable income and parents' income

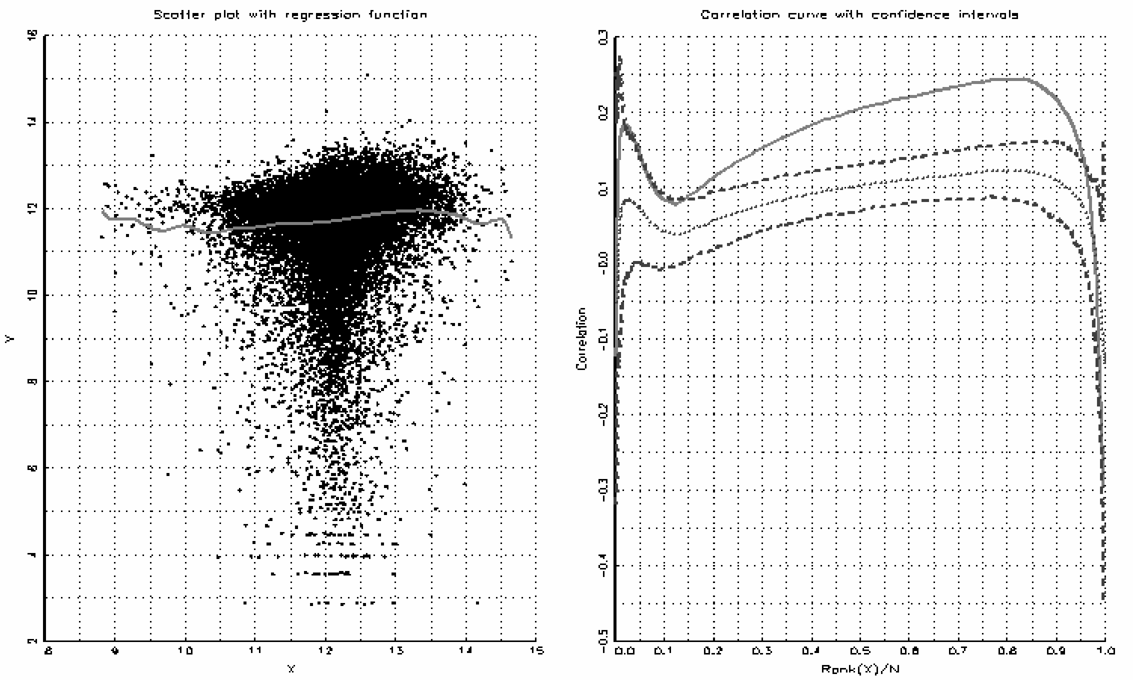

Fig. 7. Scatter plot and correlation curve (and elasticity) for female middle aged sample: labor income and father's income

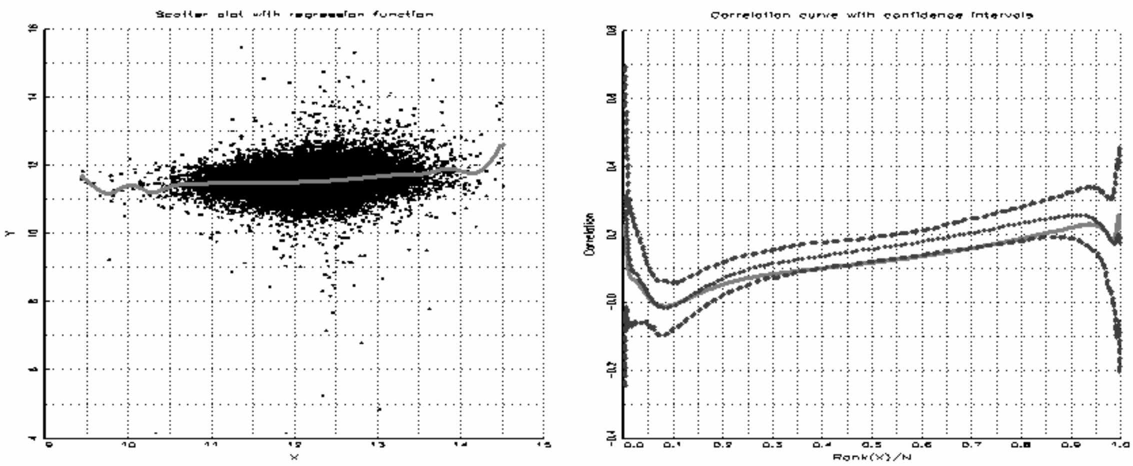

Fig. 8. Scatter plot and correlation curve (and elasticity) for female middle aged sample: disposable income and parents' income