

## Estimating nonlinear intergenerational income mobility with correlation curves

*William Nilsson*

### Abstract

A correlation curve is introduced as a tool 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 parents' 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 labour earnings are compared to disposable income. The correlation between the parental income and the child's adult disposable income becomes stronger for higher percentiles in the income distribution of the parents. Above the median the correlation is found to be stronger than for labour earnings. Interestingly, the elasticity is higher for labour earnings for most part of the distribution and complementing the elasticity with correlation curves provides a much more complete picture of the intergenerational income mobility.

**JEL** C14 D63 J62

**Keywords** Intergenerational mobility; nonlinear; nonparametric; correlation curve

### Author

*William Nilsson*, Department of Applied Economics, University of the Balearic Islands  
Palma de Mallorca, Spain, [william.nilsson@uib.es](mailto:william.nilsson@uib.es)

**Citation** William Nilsson (2019). Estimating nonlinear intergenerational income mobility with correlation curves. *Economics: The Open-Access, Open-Assessment E-Journal*, 13 (2019-7): 1–20. <http://dx.doi.org/10.5018/economics-ejournal.ja.2019-7>

Received May 25, 2018 Published as Economics Discussion Paper September 4, 2018

Revised December 3, 2018 Accepted December 10, 2018 Published January 17, 2019

© Author(s) 2019. Licensed under the [Creative Commons License - Attribution 4.0 International \(CC BY 4.0\)](https://creativecommons.org/licenses/by/4.0/)

## 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 not using 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 and Solon, 2006; and Grawe, 2006) and the possibility of a nonlinear relation (Corak and Heisz, 1999; Österbacka, 2001; Björklund and Chadwick, 2003; Fertig, 2003; Grawe, 2004; and Bratsberg et al. 2007). Couch and 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 et al. (2005) complement their analyses with transition matrices. Eide and 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. Many recent studies complement the elasticity with Spearman rank correlation. See, for example, Chetty et al. (2014), Corak et al. (2014), Chen et al. (2017), Gregg et al. (2017), Heidrich (2017) and Bratberg et al. (2017). Nybom & Stuhler (2017) show that rank correlation is particularly robust to the problems of measurement that are mentioned above, i.e. life cycle bias and few years of observed income. Adermon et al. (2018) use rank correlation to study multigenerational mobility in wealth for a Swedish sample. This very recent literature, which considers information on at least three generation, is reviewed in Solon (2018).

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 a 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 and Mazumder (2008), where trends in inequality are also related to trends in intergenerational elasticity. In Blanden et al. (2007) the elasticities are scaled, i.e. multiplied by the ratio of standard deviations of the parent's income to the standard deviation of the income of the adult children ( $\sigma_x / \sigma_y$ ),

to obtain partial correlation coefficients. The reason is precisely to cope with different dispersion for the different generations. Recent research on intergenerational income mobility has focused on comparing different measures of economic situation (for example father's or parents' incomes or earnings) or different subgroups (for example intact families or divorced families and trends over time). See for example Nordli Hansen (2010), Lucas and Kerr (2013), Bratberg et al. (2014) and Lefranc et al. (2014). The ratio of  $\sigma_x / \sigma_y$  can be very different for different subgroups or depending on which measure is used and it is important to complement the elasticity with the coefficient of correlation.

The intergenerational *transmission* tries to answer how a hypothetical increased income for the parents would manifest itself in a different income for the child in the next generation, without imposing a causal interpretation. The intergenerational *degree of association* refers to how strongly the income of the two generations is related. 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 elasticity, 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. The correlation curve was in fact mentioned by Björklund and Jäntti (2000) as "an obvious direction in which to move" in a survey of the literature on intergenerational mobility of socio-economic status.

Bjerve and 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. Point wise confidence intervals can be included with a bootstrap technique (Nilsson and del Barrio Castro, 2012). The concept of a local correlation is a natural extension

of the correlation coefficient to nonlinear models and can be applied to any area in social science. Despite that, very few applications can be found and it appears that the method is not very well known. The concept of a local correlation has probably gained most interest in financial economics, but even there only a few articles have used the method (see, for example, Inci et al. (2011) and Støve et al. (2014)).

An advantage of using this method is that intergenerational mobility is measured at each position in the income distribution of the 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. It is important to clarify that the correlation curve is proposed as a complement, and not as an alternative to the nonlinear intergenerational income elasticity.

In this study the method is applied to Swedish data and nonparametric estimates of the elasticity are also included. The purpose is to compare two different aspects of intergenerational income mobility. First, mobility is measured for individual labour earnings; i.e., interest is in equality of opportunity in providing a salary for the household. Labour earnings are measured 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.<sup>1</sup> 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 earnings and disposable income, in intergenerational income correlation. Using the correlation coefficient does not give an accurate summary of the degree of association. The correlation for both measures of the son's income and parental income is higher for higher deciles of the income distributions of the parents. For higher deciles the correlation is stronger for disposable income than for labour earnings. At the same time, the nonparametric elasticity is clearly higher for most parts of the distribution when labour earnings are 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.

---

<sup>1</sup> 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.

## 2 Method

The literature review of intergenerational income mobility is very brief as important surveys are available (see, for example, Solon (1999) and Blanden (2011)). The focus in this section is on the introduction of an alternative method to study intergenerational income mobility.

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,  $m(x) = E(Y|X = x)$ , 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. The local *degree* of association can be estimated using a correlation curve,

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

where  $\sigma^2(x) = \text{var}(Y|x)$  is the local residual variance.  $\sigma_x$  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 and 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 in addition, variation in the residual variance, i.e. a possible heteroscedastic pattern, will influence the measure, indicating whether the association is locally weaker or stronger. It is important to clarify that standardizing both variables before estimating the local elasticity is not going to provide a correlation curve. The reason is that variation in residual variance around the regression function would remain, and the first derivative would not take such variation into account. Accordingly, the strength of the local association would not be detected.

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 and Gijbels (1995a) to find the

optimal bandwidth for the first derivative is used. Confidence intervals for the correlation curve is added using wild bootstrapping. The method was introduced for the correlation curve in Nilsson and del Barrio Castro (2012) and the coverage rates are found to be satisfactory. Details concerning both the nonparametric method and the bootstrap procedure are included in Appendix.

### **3 Data**

The empirical analysis in this study is based on Swedish register data administered by Statistics Sweden. The sample consists of 10% random sample of individuals born between 1949 and 1958. The sample is divided into male and female samples that are analyzed separately. Several sample selection criteria are used to obtain the data to analyze. These restrictions include that parents are required to be present in the household and having a positive average income. The details are clarified below.

Parents living in the same household as the individual are found in the Population and Housing Census for the year 1965. By using the censuses in 1960 and 1970, the final sample is restricted to individuals where both the father and the mother are identified at least in two censuses five year apart. 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 and 1977 and come from the Income and Wealth Register. Parents were required to be alive until at least the year 1978 to be included.

The longitudinal database LOUISE is used for the income of the individuals, which is measured for the years 1994 to 1999, and the outcome is, accordingly, measured over six years between the ages of 36 and 50. Labour earnings and disposable income are used in the analysis. Note that an exact counterpart to the measure of parents' income is not available in the registers. Fathers' and mothers' income is aggregated net labour and capital income. 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 are included in Table 1.

The natural logarithm is used for all income measures. The samples are restricted to adult child and parents with positive average income/earnings. This restriction decreases the male sample size by about 6% when labour earnings are used. The sample size is reduced by about 5% when the female sample is used. The restriction to exclude an average income that is zero hardly affects the sample sizes at all when disposable income is the variable measuring adult income.

Table 1: Summary statistics

	Years	Mean age	P10	P30	P50	P70	P90	Mean	Std. dev.	N
Labour income										
Sons	1994-1999	42.9	10.96	12.12	12.33	12.51	12.86	12.04	1.21	34539
Parents	1971, 74, 77.	Father: 52.3 Mother: 49.0	11.94	12.32	12.55	12.74	13.06	12.52	0.48	
Disposable income										
Sons	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	12.47	0.48	
Labour income										
Daughters	1994-1999	42.9	10.82	11.74	12.00	12.18	12.42	11.73	1.03	33797
Parents	1971, 74, 77.	Father: 52.3 Mother: 49.0	11.95	12.33	12.55	12.74	13.06	12.52	0.48	
Disposable income										
Daughters	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: All income variables are expressed after taking the logarithm. 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 sum of the aggregate net labour and capital income for both parents. Note that parents' incomes are after deducting possible deficit in income sources. The incomes are measured in Swedish Crowns (krona) deflated to the price level of 2001. Mean age refers to the age when income is measured.

The standard deviation of labour earnings is substantially higher than the standard deviation for the income of the parents. The larger dispersion for the adult children is also evident when incomes in different percentiles are compared. Note, however, that parents' 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 parents 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. This selection can make the distribution more compressed, while no such selection is applied for the adult children.

Disposable income has a smaller standard deviation for the adult children, and this is in fact quite 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. As a departure for the analysis, correlation coefficients and linear elasticities are estimated. These results can be found in Table 2. Pearson's correlation coefficient is higher when disposable income is used compared to when labour earnings are used, while the opposite is the case for the elasticity. Comparing the results for the different measures it is clear that the intergenerational elasticity only captures intergenerational income transmission, and not the degree of association. 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 earnings. 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. Pearson's correlation coefficient indicates that society is more rigid when it comes to disposable income compared to labour earnings. In this case, the results are not affected by a different standard deviation between the income of the individuals and that of the parents. Rank correlation is found to be fairly similar for the two measures. It is even slightly higher for labour earnings compared to disposable income.

The estimated values for the correlation coefficients 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 parents' income distribution.

Figure 1 includes scatter plots, nonparametric elasticities and the correlation curves using labour earnings respectively disposable income for the male sample. 95% confidence intervals are also included. When analysing the second column in the figure for both labour earnings and disposable income it is important to remember that 80% of the sample is found between log income of 11.94 (first decile) and 12.86 (the 9<sup>th</sup> decile) and the measures are relevant for fewer individuals in the tails of the distribution.

To simplify the comparison the figure in the third column in Figure 1 includes the elasticity and correlation curves expressed for different percentiles of the parents' income distributions. The elasticity is found to be substantially higher for labour earnings compared to disposable income for most parts of the distribution. From percentile 35 the elasticity is above 0.4 and from decile 7<sup>th</sup> to almost 9<sup>th</sup> decile the elasticity is even above 0.45 for labour earnings.

When disposable income is analysed, the elasticity is below 0.15 at percentile 35 and reaches its highest value, i.e. about 0.28, at about the 9<sup>th</sup> decile. The intergenerational income transmission is stronger for labour earnings compared to disposable income.

The correlation gradually increases until between the 8<sup>th</sup> and 9<sup>th</sup> decile for both labour earnings and disposable income. For lower deciles, i.e. from the first to the fifth decile, the correlations are fairly similar for labour earnings and disposable income. The correlation is found to be slightly above 0.15 at percentile 35 for labour earnings, but even at the highest position, the curve is only about 0.2. This happens approximately at percentile 85. For dispo

Table 2: Elasticities and correlation coefficients

	Elasticity	Std. error	Pearson correlation	Rank correlation
<b>Sons</b>				
Labour income	0.3488***	0.0134	0.1382***	0.2171***
Disposable income	0.1665***	0.0045	0.1904***	0.1865***
<b>Daughters</b>				
Labour income	0.1810***	0.0178	0.0834***	0.1479***
Disposable income	0.1121***	0.0041	0.1421***	0.1323***

Notes: \*\*\*  $p < 0.01$ .

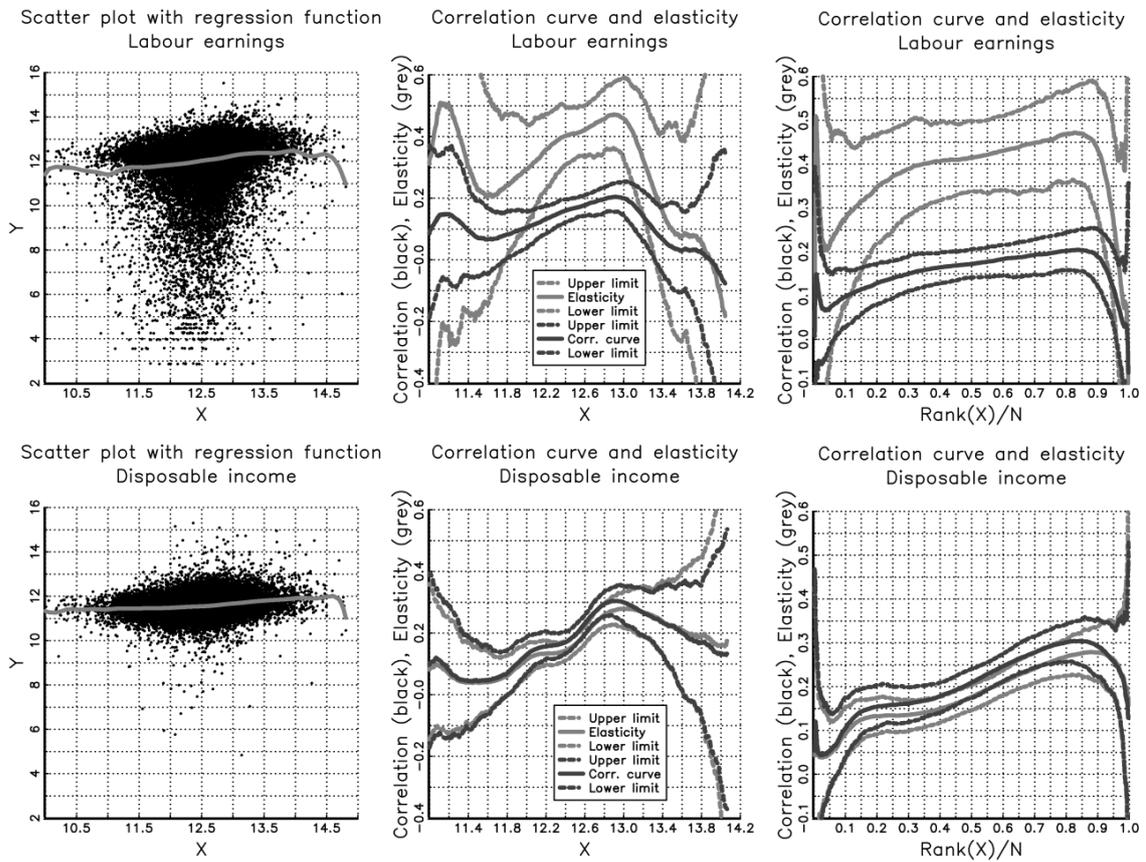


Figure 1: Scatter plot and correlation curves (and elasticity) for sons.

able income the highest correlation is about 0.3 for the same percentile. At the eighth decile the correlation is found to be statistically significantly higher for disposable income compared to labour earnings.<sup>2</sup> The *degree of association* is, accordingly, quite similar below the median parental income, but for higher deciles, the degree of association is much higher for disposable income. The correlation is particularly low for the first deciles.

An elasticity of about 0.45 that is found for labour earnings at the 7<sup>th</sup> decile (to almost 9<sup>th</sup> decile) of the parents' income distribution, implies that an increase of 1% of the parents' income would, on average, mean a 0.45% increase in labour earnings for the son. It is important to remember that it is not possible to make a causal interpretation of the elasticity. An elasticity of 0.45 is comparatively large, in particular, for the Swedish case, but complementing the elasticity with a measure of local correlation of about 0.2 provides an important perspective. If the local correlation is squared, we obtain a local coefficient of determination of about 0.04, and the average transmission that is given by the elasticity is clearly uncertain. Despite an elasticity of 0.45, the child's labour earnings are fairly weakly related to the income of the parents. Inequality of opportunity to acquire labour earnings does not seem to be too severe, as long as parents' income is used as the only relevant circumstance. Factors unrelated to parents' income, could, however, still be important to create inequalities of opportunity to acquire labour earnings.

The corresponding results for the female sample can be found in Figure 2 and the pattern is fairly similar to what was found for the male sample. The difference is relatively large between the correlation curve and the elasticity when labour earnings are analyzed. The elasticity does, however, not reach as high values as for the male sample. For example, at percentile 35 the elasticity is below 0.2 and between decile 6 and 8 the point estimate is slightly above 0.3. For disposable income the elasticity is below 0.1 at percentile 35 and rises as the parents' income increases. At decile 9 the elasticity is above 0.2. The correlation curve, when labour earnings are analyzed, is much below the curve for the elasticity. At its highest position, i.e. about decile 7 and 8, the correlation is still below 0.15. For disposable income the correlation increases as the parents' income increases and the curve reaches its highest value, i.e. slightly above 0.25, at about decile 9. The shapes of both the correlation curves as well as the elasticity are similar to the male sample, but the magnitude is lower.

Using the elasticity to evaluate the *intergenerational income transmission*, and the correlation curve to study the *degree of intergenerational income association*, provides different conclusions concerning if the income mobility is higher or lower for labour earnings or disposable income. The results underline the main argument of the paper, i.e. the intergenerational income transmission should be accompanied with an analysis on the degree of intergenerational income association, and the method should be able to capture nonlinearities.

---

<sup>2</sup> 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%.

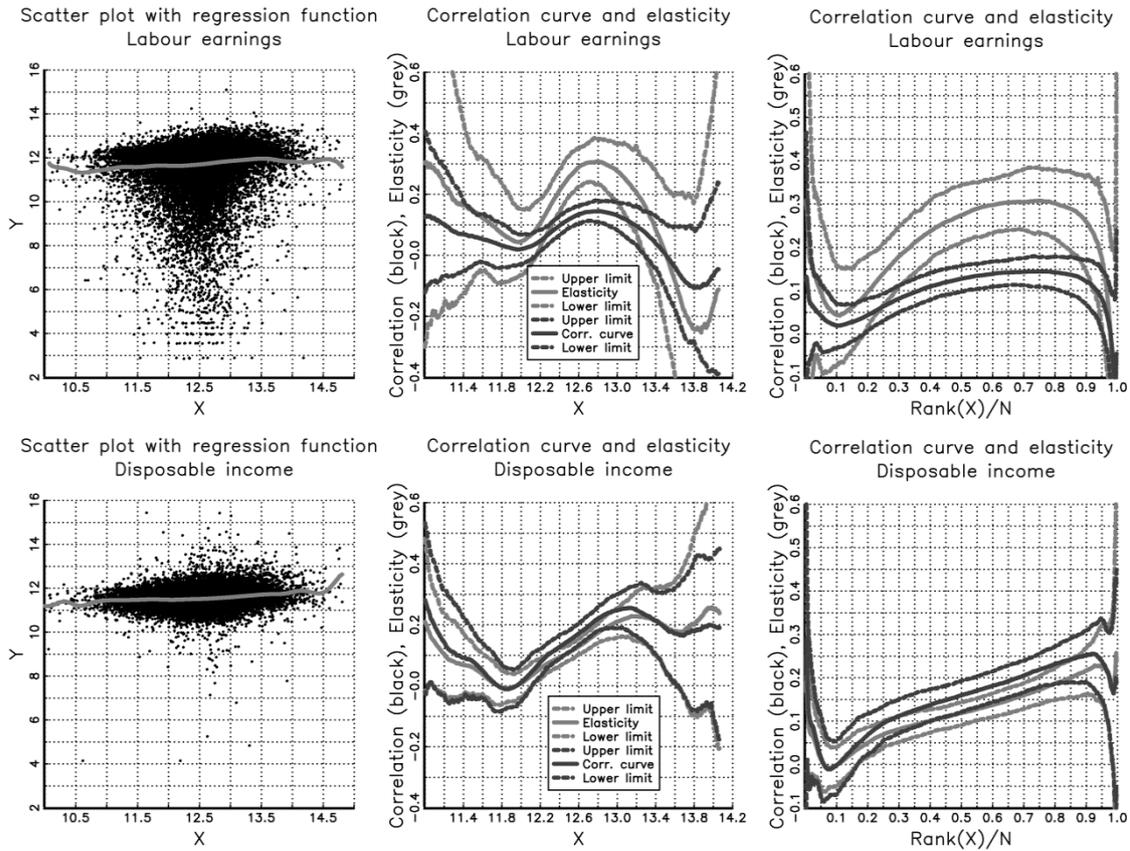


Figure 2: Scatter plot and correlation curves (and elasticity) for daughters.

## 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 father's 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. An important 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. 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 parents' income distribution the results show that the degree of intergenerational income correlation is higher for disposable income than for labour earnings. 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 earnings 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 is about 0.45 at the eighth decile when labour earnings are analysed for a male sample. The corresponding elasticity, 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 earnings are studied. Therefore, the elasticity suggests that the labour earnings are transmitted with much higher persistence among generations than disposable income. The correlation curve effectively clarifies that this result is due to the different dispersion that inflates the elasticity when labour earnings are used. The *degree* of relation is, in fact, stronger at higher deciles for disposable income than for labour earnings. 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.

The data used in this study covers six years of income for the adult child, measured around age 43, and three years of *both* father's and mother's income, measured in 1971, 1974 and 1977. Despite that it is a fairly good situation to avoid large life cycle bias and measurement error due to transitory variation it cannot be ruled out that these concerns are present. In particular it would be interesting to evaluate how the correlation curve would perform in different parts of the distribution going from using a complete working life history to a less favorable situation.

It is also important to remember that father's income, which often has been used in the literature, is not really a "noisy" measure of parents' income. If our preferred measure is parents' combined income, not having mother's income implies missing-out an important component. This component should not be seen as a constant with random noise, because the incomes of parents are (usually) related due to assortative mating or labour market decision within the household. Accordingly, the difference of using father's income instead of parents' income can provide important differences, in particular for local measures.

The key argument in this paper is that the correlation curve provides an important complement to nonlinear elasticities. The correlation curve is not proposed as a substitute to the traditional elasticity, but it is important to be clear about that different measures answer different questions.

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 similar forms of discrete classification of the income distributions are used, a society 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 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 parents' 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 #ECO2014-59055-R). The author thanks seminar participants at the 5<sup>th</sup> ECINEQ conference held in Bari, Italy, 2013, as well as two referees for providing helpful comments.

## References

- Aaronson D, Mazumder B (2008) Intergenerational Economic Mobility in the United States, 1940 to 2000. *The Journal of Human Resources* 43(1):139–172.  
<http://jhr.uwpress.org/content/43/1/139.refs>
- Adermon A, Lindahl M, Wadenström, D (2018) Intergenerational Wealth Mobility and the Role of Inheritance: Evidence from Multiple Generations. *The Economic Journal* 128(612):F482–F513.  
<https://onlinelibrary.wiley.com/doi/abs/10.1111/ecoj.12535>
- Bjerve S, Doksum K (1993) Correlation Curves: Measures of Association as Functions of Covariate Values. *The Annals of Statistics* 21(2):890–902. <http://www.jstor.org/stable/2242266>
- Björklund A, Chadwick L (2003) Intergenerational income mobility in permanent and separated Families. *Economics Letters* 80(2):239–246. [http://dx.doi.org/10.1016/S0165-1765\(03\)00084-3](http://dx.doi.org/10.1016/S0165-1765(03)00084-3)
- Björklund A, Jäntti M (2000) Intergenerational mobility of socio-economic status in comparative perspective. *Nordic Journal of Political Economy* 26:3–32.  
[http://www.nopecjournal.org/NOPEC\\_2000\\_a01.pdf](http://www.nopecjournal.org/NOPEC_2000_a01.pdf)
- 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(5-6): 474–484. <http://dx.doi.org/10.1016/j.jpubeco.2012.02.003>
- Blyth S (1994) Karl Pearson and the correlation curve. *International Statistical Review* 62(3):393–403.  
<http://www.jstor.org/stable/1403769>
- Blanden J (2011) Cross-country rankings in intergenerational mobility: a comparison of approaches from economics and sociology. *Journal of Economic Surveys* 27(1): 38–73.  
<https://doi.org/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(519):C43–C60.  
<https://doi.org/10.1111/j.1468-0297.2007.02034.x>
- Bratberg E, Davis J, Mazumder B, Nybom M, Schnitzlein DD & Vaage K (2017) A comparison of intergenerational mobility curves in Germany, Norway, Sweden, and the US. *The Scandinavian Journal of Economics* 119(1):72-101. <https://doi.org/10.1111/sjoe.12197>
- Bratberg E, Nilsen OA, Vaage K (2005) Intergenerational earnings mobility in Norway: Levels and trends. *The Scandinavian Journal of Economics* 107(3):419–435.  
<http://www.jstor.org/stable/3440966>
- Bratberg E, Rieck KME, Vaage K (2014) Intergenerational earnings mobility and divorce. *Journal of Population Economics* 27(4):1107–1126. <https://www.jstor.org/stable/44289697>
- 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(519):C72–C92. <https://doi.org/10.1111/j.1468-0297.2007.02036.x>

- Chen W H, Ostrovsky Y & Piraino P (2017) Lifecycle variation, error-in-variables bias and nonlinearities in intergenerational income transmission: new evidence from Canada. *Labour Economics* 44:1-12. <https://doi.org/10.1016/j.labeco.2016.09.008>
- Chetty R, Hendren N, Kline P & Saez E (2014) Where is the land of opportunity? The geography of intergenerational mobility in the United States. *The Quarterly Journal of Economics* 129(4):1553–1623. <https://doi.org/10.1093/qje/qju022>
- Corak M, Heisz A (1999) The intergenerational earnings and income mobility of Canadian men. *The Journal of Human Resources* 34(3):504–533. <https://www.jstor.org/stable/146378>
- Corak M, Lindquist MJ & Mazumder B (2014) A comparison of upward and downward intergenerational mobility in Canada, Sweden and the US. *Labour Economics* 30:185–200. <https://doi.org/10.1016/j.labeco.2014.03.013>
- Couch KA, Lillard DR (1998) Sample selection rules and the intergenerational correlation of earnings. *Labour Economics* 5(3):313–329. [http://dx.doi.org/10.1016/S0927-5371\(98\)00009-8](http://dx.doi.org/10.1016/S0927-5371(98)00009-8)
- Dearden L, Machin S, Reed H (1997) Intergenerational Mobility in Britain. *The Economic Journal* 107:47–66. [http://dx.doi.org/10.1016/S0927-5371\(98\)00009-8](http://dx.doi.org/10.1016/S0927-5371(98)00009-8)
- 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(426): 571–582. <http://www.jstor.org/stable/2290860>
- Eide ER, Showalter MH (1999) Factors affecting the transmission of earnings across generations: a quantile regression approach. *The Journal of Human Resources* 34(2):253-267. <http://www.jstor.org/stable/146345>
- Fan J (1992) Design-adaptive Nonparametric Regression. *Journal of the American Statistical Association* 87(420):998-1004. <http://www.jstor.org/stable/2290637>
- Fan J (1993) Local Linear Regression Smoother and Their Minimax Efficiency. *The Annals Statistics* 21(1):196–216. <http://www.jstor.org/stable/3035587>
- 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(2):371–394. <http://www.jstor.org/stable/2345968>
- Fan J, Gijbels I (1995b) Adaptive Order Polynomial Fitting: Bandwidth Robustification and Bias Reduction. *Journal of Computational and Graphical Statistics* 4(3):213–227. <http://www.jstor.org/stable/1390848>
- Fertig AR (2003) Trends in intergenerational earnings mobility in the United States. *Journal of Income Distribution* 12(3-4):108–130. <http://jid.journals.yorku.ca/index.php/jid/article/view/1271>
- Galindo CD, Liang H, Kauermann G, Carrol RJ (2001) Bootstrap Confidence Intervals for Local Likelihood, Local Estimating Equations and Varying Coefficient Models. *Statistica Sinica* 11(1):121–134. <http://www.jstor.org/stable/24306813>
- Grawe ND (2004) Reconsidering the Use of Nonlinearities in Intergenerational Earnings Mobility as a Test for Credit Constraints. *The Journal of Human Resources* 39(3):813-827. <http://www.jstor.org/stable/3558998>

- Grawe ND (2006) Lifecycle bias in estimates of intergenerational earnings persistence. *Labour Economics* 13(5):551–570. <http://dx.doi.org/10.1016/j.labeco.2005.04.002>
- Gregg P, Macmillan L & Vittori C (2017) Moving towards estimating son's lifetime intergenerational economic mobility in the UK. *Oxford bulletin of economics and statistics* 79(1):79–100. <https://doi.org/10.1111/obes.12146>
- Haider S, Solon G (2006) Life-cycle variation in the association between current and lifetime earnings. *The American Economic Review* 96(4):1308–1320. <http://www.jstor.org/stable/30034342>
- Heidrich S (2017) Intergenerational mobility in Sweden: a regional perspective. *Journal of Population Economics* 30(4), 1241–1280. <https://link.springer.com/article/10.1007/s00148-017-0648-x>
- Härdle W, Mammen E (1993) Comparing Nonparametric Versus Parametric Regression Fits. *The Annals of Statistics* 21(4):1926–1947. <http://www.jstor.org/stable/2242322>
- Inci AC, Li HC, McCarthy J (2011) Financial contagion: A local correlation analysis. *Research in International Business and Finance* 25(1), 11–25. <https://doi.org/10.1016/j.ribaf.2010.05.002>
- Jenkins S (1987) Snapshots versus Movies: 'Lifecycle Biases' and the Estimation of Intergenerational Earnings Inheritance. *European Economic Review* 31(5):1149–1158. [http://dx.doi.org/10.1016/S0014-2921\(87\)80010-7](http://dx.doi.org/10.1016/S0014-2921(87)80010-7)
- Jääntti 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. <http://ftp.iza.org/dp1938.pdf>
- Lefranc A, Ojima F, Yoshida T (2014) Intergenerational earnings mobility in Japan among sons and daughters: levels and trends. *Journal of Population Economics* 27(1):91–134. <https://link.springer.com/article/10.1007/s00148-012-0464-2>
- Lucas REB, Kerr SP (2013) Intergenerational income immobility in Finland: contrasting roles for parental earnings and family income. *Journal of Population Economics* 26(1):1057–1094. <https://link.springer.com/article/10.1007/s00148-012-0442-8>
- McMurry T, Politis D (2004) Nonparametric regression with infinite order flat-top kernels. *Journal of Nonparametric Statistics* 16(3-4):549–562. <http://dx.doi.org/10.1080/10485250310001622596>
- Newey W, Hsieh F, Robins J (2004) Twicing Kernels and a small Bias Property of Semiparametric Estimators. *Econometrica* 72(3):947–962. <http://www.jstor.org/stable/3598841>
- Nilsson W, del Barrio Castro T (2012) Bootstrap Confidence Interval for a Correlation Curve. *Statistics & Probability Letters* 82(1):1–6. doi:10.1016/j.spl.2011.09.001 <http://dx.doi.org/10.1016/j.spl.2011.09.001>
- Nordli Hansen M (2010) Changes in intergenerational economic mobility in Norway: conventional versus joint classifications of economic origin. *The Journal of Economic Inequality* 8(2):133–151. <https://link.springer.com/article/10.1007/s10888-009-9109-4>
- Nybom M & Stuhler J (2017) Biases in standard measures of intergenerational income dependence. *The Journal of Human Resources* 52(4):800–825. <http://jhr.uwpress.org/content/early/2016/12/05/jhr.52.3.0715-7290R.abstract>

- Ruppert D (1997) Empirical-Bias Bandwidths for Local Polynomial Nonparametric Regression and Density Estimation. *Journal of the American Statistical Association* 92(439):1049–1062. <http://www.jstor.org/stable/2965570>
- Ruppert D, Wand MP (1994) Multivariate Locally Weighted Least Squares Regression. *The Annals Statistics* 22(3):1346–1370. <http://www.jstor.org/stable/2242229>
- Solon G (1992) Intergenerational income mobility in the United States. *The American Economic Review* 82(3):393–408. <http://www.jstor.org/stable/2117312>
- Solon G (1999) Intergenerational Mobility in the Labor Market. In: Ashenfelter O, Card D (eds) *Handbook of Labor Economics* vol 3. Elsevier, Amsterdam, pp 1761–1800
- Solon G (2018) What Do We Know So Far about Multigenerational Mobility? *The Economic Journal* 128(612):F340–F352. <https://doi.org/10.1111/econj.12495>
- Støve B, Tjøstheim D, Hufthammer KO (2014), Using local Gaussian correlation in a nonlinear re-examination of financial contagion, *Journal of Empirical Finance* 25, 62–82. <https://doi.org/10.1016/j.jempfin.2013.11.006>
- Zimmerman DJ (1992) Regression toward Mediocrity in Economic Stature. *The American Economic Review* 82(3):409–429. <http://www.jstor.org/stable/2117313>
- Österbacka E (2001) Family Background and Economic Status in Finland. *The Scandinavian Journal of Economics* 103(3):467–484. doi:10.1111/1467-9442.00255 <http://www.jstor.org/stable/3441083>
- Österberg T (2000) Intergenerational Income Mobility in Sweden: What Do Tax-data Show? *The Review of Income and Wealth* 46(4):421–436. <https://doi.org/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  $p$ :

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

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:  $\hat{\beta} = (X^T W X)^{-1} X^T W y$ .  $\beta = (\beta_0, \dots, \beta_p)^T$  is the nonparametric estimate of the regression function and its derivatives,  $v! \hat{\beta}_v = m^{(v)}(x_0)$ ,  $v = 0, \dots, p$ .  $W$  is a diagonal matrix with  $W_i = K((x_i - x_0)/h)$ .  $X$  is a design matrix with  $(l, j)$ -th element as  $(x_l - x_0)^{j-1}$  and  $y = (y_1, \dots, y_n)^T$ . The standardized residual variance is:

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

where  $\hat{y} = (\hat{y}_1, \dots, \hat{y}_n)^T = X \hat{\beta}$ . 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. For the nonparametric estimations, a Gaussian kernel was used:

$$K(x) = X \left( \frac{x_i - x}{h} \right) = \frac{1}{\sqrt{2\pi}} \exp \left\{ -\frac{1}{2} \left( \frac{x_i - x}{h} \right)^2 \right\} \quad (A3)$$

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 et al. 2004 and McMurry and 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 and Gijbels (1995b) recommend using an odd-order polynomial to estimate the regression function. Ruppert and 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 and 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 and del Barrio Castro (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 and 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 and 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  $h_{min} = (x_{max} - x_{min}) * (order + 2.0) / n$ , where  $x_{max}$  and  $x_{min}$  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.

### Point wise confidence intervals for the correlation curve

The correlation curve can easily be estimated with the nonparametric method described above. 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 and del Barrio Castro (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 and Mammen (1993) introduced wild bootstrapping to obtain confidence intervals for nonparametric regressions.

To obtain a bootstrap point wise 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  $\varepsilon_i^*$ , 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  $(x_1, y_1^*), \dots, (x_n, y_n^*)$ . 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  $(1 - \alpha)$ -point wise confidence interval, where  $\alpha$  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  $x_i$ . All figures exclude 100 observations in each tail of the income distribution of the parents. The reason for this is that the derivative can be taken to very extreme values in the extremes of the distributions.

Please note:

You are most sincerely encouraged to participate in the open assessment of this article. You can do so by either recommending the article or by posting your comments.

Please go to:

<http://dx.doi.org/10.5018/economics-ejournal.ja.2019-7>

The Editor