Dynamic returns to schooling by work experience

Marconi, Gabriele

OECD

29 December 2015

Dynamic returns to schooling by work experience

Gabriele Marconi*^ – Directorate for Education and Skills, OECD – gabriele.marconi@oecd.org

29/12/2015

Keywords: earnings curve – wage – work experience – dynamic panel data model – lagged dependent variable

JEL codes: J31 – I21 – C51

Abstract: This paper investigates the effect of education on the rate of growth of a worker's salary for workers with different levels of work experience. Two alternative models of the earnings curve are presented, the Mincer model and a model with wage premia that change over the course of a workers' career. Then, I derive analytically the coefficients that would be obtained by estimating these two models with a dynamic panel data model in which earnings are a function of schooling and past earnings. Finally, I estimate these coefficients empirically using data from the German Socio-Economic Panel. The results indicate that the coefficients for schooling and lagged wage in a dynamic panel data model of the earnings curve display substantial variation by work experience, which seems more consistent with the dynamic wage premium model than with the Mincer model.

* The opinions expressed and arguments employed herein are solely those of the author and do not necessarily reflect the official views of the OECD or its member countries.

^ The author has no competing interests in the manuscript

1. Introduction

The literature studying the effect of schooling on wages has traditionally assumed that this effect is constant over time, or "parallel". This implies that a worker with an additional year of schooling would earn a certain percentage (say, z%) more than if he had not studied that additional year, after 1 year of work experience as well as after 20 years. This view has been embedded in the so-called Mincer model of the earnings curve.

Since the 1970s, some studies challenged the assumption of a time-constant earnings premium. However, it is only since the 2000s that the literature on the effect of schooling on the slope of the earnings curve has started to develop more substantially. As I discuss in the next section, the theories underlying these studies, their methodologies, and their results are different.

This paper contributes to the literature on schooling and the earnings curve's slope by using a dynamic panel data model to investigate the effect of education on the rate of growth of a worker's salary by years of work experience (Section 3). The data, presented in Section 4, are from the German Socio-Economic Panel and cover 17841 German male workers between 2004 and 2011. Dynamic panel data models have the advantage of providing estimates of relevant parameters while controlling for individual unobserved heterogeneity by including lagged earnings in the statistical model. Dynamic panel data models have recently been estimated in the labour economics literature. The first contribution of this paper is to analyse how the estimated coefficients vary by the level of experience.

The interpretation of the results presented in this paper differs from the one suggested by other papers in the dynamic panel data models of the earnings curve. For example, Andini (see Section 2) argues that if the Mincer model is correct and the economy is competitive, in a regression of log wage on schooling and lagged log wage, it should be expected the coefficient of lagged log wage to be equal to 0. On the contrary, I argue that according to the Mincer model this coefficient should be equal 1, at least if underlying individual ability (i.e., the intrinsic ability of an individual to earn money independently of schooling) is expected to play a role. More precisely, it is shown that, according to the Mincer model, when a time-constant ability component is added, the deviation of an individual's log wage from the average of workers with the same work experience follows a random walk.

After discussing the Mincer model with underlying individual ability, I discuss a model of dynamic returns to schooling and underlying individual ability, where these two variables are able to influence not only the level of individual salary, but also its rate of growth. This is a way of relaxing the assumption of parallelism mentioned in the first paragraph of this section. The second contribution of this paper is the discussion of these two alternative models and of the coefficients of schooling and lagged wage that would be consistent with each of the two models.

The estimations presented in this paper are based on OLS and the instrumental variable (IV) method, whereby the lagged log wage is instrumented with the twice-lagged log wage. This empirical approach differs from the most recent papers on dynamic panel data model of the earnings curve, using the general method of moments (GMM) estimator. The advantage of the estimators used in this paper is that they allow obtaining reasonably simple analytical expression for the estimated coefficients (and related biases), which can be directly related to the alternative models of the earnings curve that I discuss in this paper (Section 3). These analytical expressions are also used for a suggestive estimation of the impact of underlying ability (and of schooling as well) on the wages' level and rate of growth. This estimation suffers from a number of important limitations, but it provides some suggestive evidence to assess the credibility of one of the assumptions of Section 3.

The results, presented in Section 5, suggest the following conclusions. i) The coefficients for lagged log wage and schooling in a dynamic panel data model of the earnings curve change substantially by working experience. This suggests that pooling workers with different levels of experience in the same dynamic panel regression can be misleading. The IV estimation suggests that, for advanced levels of work experience, the coefficient for lagged log wage is close to 1 and the coefficient for schooling is close to 0. ii) The estimated standard deviation and serial correlation of the error term of the earnings equation seems to be much higher and more unstable in the first years of work experience, suggesting that this phase of the workers' careers is structurally different from later phases. iii) The high variability of the coefficients by level of experience is not consistent with the Mincer model with underlying ability; conversely, this variability is not at odd with a model of dynamic returns to schooling and ability. These conclusions are discussed in Section 6.

2. Literature review

Literature on earnings curves

The path followed by individual log wage over time constitutes the well-known earnings curve, which has been intensively studied by economists. The relationship between education and wages has received special attention, generating a very large literature which received a considerable input with the study by Mincer (1974) and which has been recently summarized, for example, in Psacharopoulos and Patrinos (2004), Harmon et al. (2003), Card (1999), and Heckmann et al. (2006). In the original study by Mincer (1974), as in most of the subsequent literature, the earnings curves of individuals with different educational attainment have been assumed to be parallel, i.e., the expected wage growth of more educated individuals is assumed to be equal to that of less educated individuals.

This assumption has been criticized at least since the study by Psacharopoulos and Layard (1979). However, the literature estimating the effect of education on the slope of the earnings curve is still relatively limited. Psacharopoulos and Layard (1979) found that the growth in earnings over time tended to be more pronounced for more educated individuals, using the British Household Survey of 1972 (consistent results have been found with similar, more recent data within the study by Chevalier et al., 2004). Heckman et al. (2006) formally test the assumption of "parallelism" between the earnings curves of individuals with different levels of education, and reject it. However, they do not offer estimates of the effect of education on the slope of the earnings curve.

Recently, the role of education on earnings growth has been investigated with panel or longitudinal datasets. The evidence does not always point in the same direction. Altonij and Pierret (2001) use panel data from the 1979 US National Longitudinal Survey of Youth. They find that the effect of education on wage growth is negative, once previous performance on cognitive tests, wage of siblings, and parental education are controlled for. On the contrary, the effect of performance in cognitive tests on wage increases with work experience. They interpret this as evidence of the existence of employer learning about individual traits which are related to productivity.

The results of Altonij and Pierret (2001) motivated a number of studies, which analyzed the hypothesis of employer learning more in depth. For example, Arcidiacono et al. (2008) were able to replicate the basic findings of Altonij and Pierret (2001), but found that the results are different for high school and college graduates, as the effect of cognitive tests on wage increases with experience only for high school graduates. Bauer and Haisken-DeNew (2001) try to replicate the findings of Altonij and Pierret (2001) using the German Socioeconomic Panel (GSOEP) dataset, but find the opposite results (schooling increases wage growth). By using the GSOEP, they cannot control for the performance in cognitive tests, but only for more widely available measures such as parental education. Galindo-Rueda (2003) overcomes this problem by using the British National Child Development Study and British Cohort Study on individuals born in 1958 and 1970, respectively. The results are not easy to interpret, as they tend to change with the specification chosen and depending on whether experience or tenure, and white or blue collar workers are considered. Despite using panel datasets, these studies pool the data of individuals over time, and estimate their models by using a pooled dataset.

A study using a different methodology is the one by Brunello and Comi (2004). They use a pseudo panel of cohorts, with data covering two cohorts (1940-1949 and 1950-1959) in 11 European countries. They group individuals by country, year, period of birth, gender and educational attainment, and look at the effect of educational attainment on the change of average within-group log wage. They find a substantial and significant interaction between education and experience: “10 years of labor market experience increase real earnings by approximately 16%, 30% and 40% for employees with less than upper secondary, upper secondary and tertiary education” (p.79).

A related literature is that on depreciation of human capital (e.g. Neuman and Weiss, 1995, Arrazola and de Hevia, 2004, Weber, 2009). Studies in this field of literature usually model the individual human capital stock as a function of education, on-the-job training, and a depreciation rate. The model is then estimated, often by using a cross-section of individuals or a pooled panel dataset. All the studies reviewed by Weber (2009) find that indeed human capital depreciates at a positive rate, but there is a wide variance in the estimates (which range from 0.2% to 21% in the studies reported by Weber, 2009). This would point to the existence of a negative interaction between education and experience.

Summarizing, the impact of education on the slope of the learning curve has been object of the attention of a number of studies. However, the evidence on the role of education on determining the slope of the earnings curve does not point univocally in one direction. In the next sections, I develop a dynamic panel data model to investigate whether education affects the rate of growth of a worker’s salary. I estimate this model with GSOEP data for 17841 German male workers in the period 2004-2011.

To conclude this section, it is important to note that the coefficient of education on the slope of the earnings curve can be interpreted in different ways: it can be considered the result of employer learning (see e.g. Altonij and Pierret, 2001); of the interaction between education and training (see e.g. Brunello, 2001); of the role of education in fostering learning-by-doing; and of depreciation of human capital (see e.g. Neuman and Weiss, 1995).

Dynamic panel wage models

Most of the labour economics literature has studied the relationship between wages and education by means of the Mincerian equation, which relate schooling to education. However, labour economists have been concerned that individual unobserved factors that have a direct effect on salary could also be related to the amount of schooling undertaken by an individual. This generate the so-called "ability bias" in estimates of the returns to schooling. According to the Mincer model with ability bias, the salary of individual *i* with *x* years of work experience can be described by the following equation:

$$(1)\quad \ln(wage)_{ix} = \pi_0 + \beta s_i + \theta a_i + f(x) + v_{ix}$$

Where π_0, β and θ are parameters of the model; $\ln(wage)_{it}$ is log of earnings of worker i with x years of work experience; s_i is completed schooling; a_i is an unobserved factor with zero mean and unit variance representing unobserved individual ability; $f(x)$ is a function describing returns to experience, and usually assumed to be quadratic in x in the literature; and v_{ix} is a stochastic disturbance.

It is clear that Equation (1) cannot be estimated by using panel data models such as first differencing or fixed effects, because it would not be possible to recover the coefficient of schooling, which typically occurs before work experience. It is possible, however, to use other panel data approaches. One example is using a dynamic panel data model which includes the lagged dependent variable in the set of independent variables.

To the best of my knowledge, the only author estimating the relationship between schooling and wages by using this simple dynamic panel data model is Andini (2007, 2012, 2013a, 2013b). In these papers, the author estimates a dynamic wage model with and without additional control variables, using a panel of workers from one or more countries, and using OLS and GMM for producing the estimates. He finds that the coefficient for lagged log wage is positive and significant, and that the coefficient for schooling is positive, but lower than when it is estimated with the Mincerian specification (i.e., without including the lagged log wage in the estimation).

Andini interprets these results as evidence that, because of imperfect competition in the labour market, the observed salary is different from potential earnings. The author uses a bargaining model to explain the difference between observed earnings and net potential earnings. In particular, Andini (2013a) notices how estimating the earnings curve by the Mincerian specification suffers from a bias that he calls "persistence bias" (in addition to other biases already identified in the literature). He calls the coefficient of schooling that he obtains by estimating Equation (1) by GMM the "'true' estimate of the schooling coefficients" (p.9, quotation marks are from the author). Hence, Andini (2012, 2013a) considers the positive coefficient of the lagged dependent variable as evidence of a persistence bias. The reason is that, in the author's view, if the traditional Mincer model is correct, then the coefficient for lagged salary should be equal to 0.

However, the coefficient for lagged salary should be equal to 0 only if unobserved, individual ability does not play a role in the determination of wages. If wages follow Equation (1), then the coefficient for lagged wage should be positive, and close to 1. This is elaborated in the next section

3. Models of wages and earnings estimated in this paper

This section presents two alternative models of the earnings curve, and discusses the coefficients that would be obtained by estimating these models with a simple dynamic panel data model in which earnings are a function of schooling and past earnings.

The Mincer model with ability

In the empirical estimations that follow, for each level of experience I will estimate the following model:

$$(2)\quad w_{ix} = c_{wx} w_{ix-1} + c_{sx} s_i + \varepsilon_{ix}$$

Where w_{ix} represents individual i's deviation from the mean log wage of all workers with level of experience x; s_i is completed schooling; c_{wx} and c_{sx} are the parameters to be estimated for every level of experience x. For ease of notation, schooling is standardized, so that its mean is zero and its variance is equal to one. In this section, I derive the values of the estimated parameters c_{wx} and c_{sx} under different assumption on the data-generating process.

First, suppose that the Mincerian model of education and earnings with ability bias, as exemplified in Equation (1), is the true model generating the data. Given that the estimation is carried out for every level of experience in deviations from the mean, Equation (1) can be simplified as:

$$(3)\quad w_{ix} = \beta s_i + \theta a_i + v_{ix}$$

To start with, notice that differencing Equation (3) yields the following equation:

$$(4)\quad w_{ix} = w_{ix-1} + v_{ix} - v_{ix-1}$$

Equation (4) suggests that the coefficient for lagged salary if the Mincer model with ability bias adequately represents reality should be equal to one, and that the coefficient for schooling should be equal to zero. This implies that, according to the Mincer model, when a time-constant ability component is added, the deviation of an individual's log wage from the average of workers with the same work experience follows a random walk. Notice that this conclusion refers to the deviation of the log wage from the average of workers with the same level of experience. In terms of log wage, human capital variables still play a role: schooling determines the initial level of log salary (i.e., the starting point of the random walk) and experience impacts the average salary of the workers with a given level of experience.

However, estimation by OLS would yield to different conclusions, because in Equation (4) the independent variable w_{ix-1} is necessarily correlated with the error term (e.g. Greene, 2008). Hence, the coefficients for schooling and for lagged wage will be derived with a bias.

More precisely, fitting Equation (2) to some data that are generated by Equation (3) yields the following coefficients:

$$(5)\quad c_{sx} = \frac{(\beta+\theta\sigma_{as})\cdot(\sigma^2_{v_x}-\sigma_{v_{x,x-1}})}{\theta^2(1-\sigma_{as})+\sigma^2_{v_{x-1}}}$$

$$(6)\quad c_{wx} = \frac{\theta^2(1-\sigma_{as})+\sigma_{v_{x,x-1}}}{\theta^2(1-\sigma_{as})+\sigma^2_{v_{x-1}}}$$

Where σ_{as} is the correlation between ability and schooling; $\sigma_{vx,x-1}$ is the covariance between the error terms for levels of experience x and $x-1$ in Equation (3); σ^2_{vx} is the variance of the error term of Equation (3). Notice that both coefficients are generally different from zero. In addition, if the variance and covariance of the error term in Equation (3) is constant over different levels of work experience, the two coefficients are also constant for every level of work experience. Finally, something can be said on the size of the two coefficients. The coefficient for schooling is positive as long as β, θ, σ_{as}, and the variance and covariance of the error term are positive. The coefficient for the lagged wage should be close to one, as long as: the ability factor has a substantial impact on salary; schooling and ability are not perfectly correlated; and the error term component in Equation (3) is not large. Hence, the coefficient for schooling is likely biased upwards; the coefficient for lagged wage is likely biased downwards.

If the lagged salary is instrumented with the 2-periods lagged salary and with schooling[1], then the unbiased coefficients can be obtained as long as the error term in not serially correlated. In this case the regression yields the following coefficients:

$$(7)\ \ c_{sx} = \frac{(\beta+\theta\sigma_{as})\cdot(\sigma_{v_{x-1,x-2}}-\sigma_{v_{x,x-2}})}{\theta^2(1-\sigma_{as})+\sigma_{v_{x-1,x-2}}}$$

$$(8)\ \ c_{wx} = \frac{\theta^2(1-\sigma_{as})+\sigma_{v_{x,x-2}}}{\theta^2(1-\sigma_{as})+\sigma_{v_{x-1,x-2}}}$$

The coefficients for schooling and the lagged wage are unbiased (i.e., equal to zero and one, respectively) if the error term is not serially correlated. If it is serially correlated, then there persists an upward bias for the coefficient of schooling, and a downward bias for the coefficient of the lagged wage. However, comparing Equations (5) and (6) with Equations (7) and (8), it appears that: the variance of the error term in Equation (3) is replaced by the first-order serial correlation, which tends to be smaller in size; and the first-order serial covariance is replaced with the second-order serial covariance, which also tends to be smaller in size. Hence, the bias will tend to be smaller for the IV estimation (Equations (7) and (8)) than for OLS (Equations (5) and (6)).

Finally, notice that the Mincer model with ability bias implies that the IV coefficients (as well as the OLS coefficients) for schooling and the lagged wage are constant across levels of experience, as long as the serial covariance of the error term in Equation (3) is constant across levels of experience.

Dynamic returns to schooling and ability

Now, suppose that the salary growth differs between, on the one side, more able or more educated individuals, and on the other side, less able or less educated individuals. This can happen for different reasons: ability and schooling can raise the rate of learning-by-doing, stimulate training, or depreciate. This is captured by the following, equivalent equations which assume that the effect of schooling and ability on log salary changes linearly with experience.

$$(9)\ \ \ln(wage)_{ix} = \pi_0 + (\beta+\gamma x)s_i + (\theta+\rho x)a_i + f(x) + v_{ix}$$

$$(10)\ \mathrm{w}_{ix} = (\beta+\gamma x)s_i + (\theta+\rho x)a_i + v_{ix}$$

[1] This IV approach is a particular case of a GMM-SYS approach where only the level equation is used, and only one lag is used.

Where w_{ix} is the difference between the log wage of individual i and the average for individuals with x years of work experience. This model assumes that the returns to schooling are a linear function of experience (as implicitly assumed by Andini, 2013a, 2013b) and that ρ and γ indicate, respectively, the extent to which ability and education contribute to individual salary growth. In the following, it will be also assumed without loss of generality that θ is positive. In light of the model, this requires defining ability as the set of characteristics of an individual which increase its initial salary, conditional on education.

Notice that the model with dynamic returns to schooling and ability exemplified in Equation (10) is a particular case of the Mincer model with ability, as setting ρ and γ equal to 0 returns Equation (3). Rearranging[2], it is possible to obtain the following equation:

$$(11)\ w_{ix} = \left[1 + \frac{\rho}{\theta+\rho(x-1)}\right] w_{ix-1} + \frac{\gamma\theta-\beta\rho}{\theta+\rho(x-1)} s_i + \left[v_{ix} - \frac{v_{ix-1}}{\theta+\rho(x-1)}\right]$$

The interpretation of the coefficient for schooling in Equation (11) is not straightforward. To start with, suppose that this equation is estimated without bias. If the coefficient of schooling is positive, this means that $\gamma/\beta > \rho/\theta$. In other words, a positive coefficient for schooling means that the relative contribution of schooling to salary growth is bigger than the relative contribution of underlying, unobserved ability to salary growth. Finding a positive coefficient for schooling, however, would not exclude the possibility that human capital depreciates. It remains possible that human capital depreciates, but ability depreciates even faster, so that $\rho<\gamma<0$. In terms of Equation (11), the sign of ρ is implied by the coefficient of w_{it}: if it is bigger than 1, then ρ is positive; otherwise, ρ is negative. So, when estimating Equation (11), both the coefficients of schooling and of lagged salary are important, and they must be interpreted jointly. If both the coefficient for schooling is positive and the coefficient for the lagged wage is larger than one, for example, then both ability and schooling increase the returns to experience, and schooling is relatively more important than ability in doing so. If the coefficient of schooling is negative and the coefficient for lagged wage is smaller than one, then ability and schooling become less important with time in determining salary (human capital depreciation), and this effect is relatively stronger for schooling.

One problem with Equation (11), however, is that (like in any panel data model of this type) w_{ix-1} is necessarily endogenous, because it is correlated with v_{x-1}, which appears in the error term. Estimating Equation (11) by OLS would yield the following coefficients:

$$(12)\ c_{sx} = \frac{[\theta+\rho(x-1)](\gamma\theta-\beta\rho)(1-\sigma_{as}^2)+\sigma_{v_{x-1}}^2[(\beta+\gamma x)+(\theta+\rho x)\sigma_{as}]-\sigma_{v_{x,x-1}}\{[\beta+\gamma(x-1)]+[\theta+\rho(x-1)]\sigma_{as}\}}{[\theta+\rho(x-1)](1-\sigma_{as}^2)+\sigma_{v_{x-1}}^2}$$

$$(13)\ c_{wx} = \frac{[\theta+\rho(x-1)](\theta+\rho x)(1-\sigma_{as}^2)+\sigma_{v_{x,x-1}}}{[\theta+\rho(x-1)]^2(1-\sigma_{as}^2)+\sigma_{v_{x-1}}^2}$$

Equations (12) and (13) show that OLS will estimate the coefficients with a bias. The direction for the bias in the coefficient of schooling is not straightforward. However, if the true coefficient were zero (i.e. the relative role of education and ability in stimulating wage growth is the same, $\gamma/\beta=\rho/\theta$), then the estimated parameter would be positive (implying an upward bias) as long as: β, ρ and σ_{as} are positive; and the variance of the error term of Equation (10) is larger than its first order serial

[2] Scaling by one year and solving for a_i yields: $a_i = \frac{w_{ix-1}}{\theta+\rho(x-1)} + \frac{[\beta+\gamma(x-1)]v_{ix-1}}{\theta+\rho(x-1)} + \frac{v_{ix-1}}{\theta+\rho(x-1)}$. Substituting this expression into Equation (10) yields Equation (11).

covariance. If these conditions hold, it also follows that the coefficient for the lagged wage is downward biased (i.e., it is smaller than one).

It is possible to remove the bias, under the assumption that the error term of Equation (10) is not serially correlated, by instrumenting the lagged wage with the two-periods lagged wage and schooling. Then, the two coefficients are equal to:

$$(14)\, c_{sx} = \frac{[\theta+\rho(x-1)](\gamma\theta-\beta\rho)(1-\sigma_{as}^2)+\sigma_{v_{x-1,x-2}}[(\beta+\gamma x)+(\theta+\rho x)\sigma_{as}]-\sigma_{v_{x,x-2}}\{[\beta+\gamma(x-1)]+[\theta+\rho(x-1)]\sigma_{as}\}}{[\theta+\rho(x-1)](1-\sigma_{as}^2)+\sigma_{v_{x-1,x-2}}}$$

$$(15)\; c_{wx} = \frac{[\theta+\rho(x-1)](\theta+\rho x)(1-\sigma_{as}^2)+\sigma_{v_{x,x-2}}}{[\theta+\rho(x-1)]^2(1-\sigma_{as}^2)+\sigma_{v_{x-1,x-2}}}$$

If the error term is serially correlated, then a bias persists. The direction of the bias is the same as for the OLS case under similar conditions (i.e., β, ρ and σ_{as} are positive and the first-order serial covariance of the error term is bigger than its second-order serial covariance). In other words, both in the case of OLS and IV, there is an upward bias for the coefficient of schooling, and a downward bias for the coefficient of the lagged wage. However, the bias is more severe for the OLS estimation.

In fact, the error term seems to be serially correlated, so that the estimates of the coefficients of Equation (11) are likely to be biased. The estimated error term of Equation (11) is correlated with its lag in a pooled regression, both if Equation (11) is estimated with OLS (r=–0.27, p-value<1%) and if it is estimated with IV (r=0.25, p-value<1%). However, the approach taken in this paper is to accept the fact that there is some bias in the estimation, provided that in Equations (12)-(15) an expression for the biased coefficients have been derived analytically. This analytical expression allows to assess the direction of the bias and to interpret the meaning of the estimated coefficients despite the bias. This approach provides an alternative to econometric techniques such as GMM (used for example by Andini, 2013b, in an empirical context similar to the one of this paper), where the researcher assume that the estimates are unbiased, instead of deriving analytically the biased, estimated coefficients.

Notice that, if $\rho>0$, the parameters for schooling and lagged wages in Equation (11) will converge towards zero and one (respectively) for increasing levels of work experience. Hence, the model requires estimating coefficients that are not constant over time, but change with experience. Similarly, if the variance and serial covariance of the error term in Equation (11) are approximately constant, the estimated coefficients will converge towards zero and one, although at a slower pace. Of course, work experience cannot increase beyond a certain point. However, it is interesting to notice that the coefficients are time changing, because this is not the case in the standard Mincer model. Hence, evidence on the development of the coefficients for schooling and lagged wages over time can help discriminate between the two models that are considered in this section: the Mincer model and the model with dynamic returns to schooling and ability. In particular, in the Mincer model with underlying ability the two coefficients should be stable over time (unless the variance and serial covariance of the error term are very unstable). Conversely, in the model with dynamic returns: the coefficient for lagged wage should be larger than one, and converge towards one for high levels of work experience; the coefficient of schooling should decrease (assuming that it starts positive) with experience, and become very close to 0 for high levels experience.

Error term's variance and covariance and suggestive model estimation

As discussed in the previous paragraph, if the variance and covariance of the error term is known, the development over time of the estimated coefficients should help to discriminate between the Mincer model with ability and the dynamic returns model (in the first case, the coefficients for schooling and lagged wage would be constant across different levels of experience; in the second case, they would change by experience). Hence, it is very useful to derive some estimates of the variance and serial covariances of the error term in Equation (10).

Under certain assumptions, it is possible to derive such estimates from the variance and serial covariance of the estimated error term in Equation (2). Notice that, if the model represented in Equation (10) is correct, and Equation (2) is estimated by OLS:

$$(16)\ \sigma_{\varepsilon_x}^2 = \sigma_{v_x}^2 + \left[\frac{\theta+\rho x}{\theta+\rho(x-1)}\right]^2 \sigma_{v_{x-1}}^2 - 2\frac{\theta+\rho x}{\theta+\rho(x-1)}\sigma_{v_{x,x-1}}$$

$$(17)\ \sigma_{\varepsilon_{x,x-1}} = \sigma_{v_{x,x-1}} + \frac{\theta+\rho x}{\theta+\rho(x-2)}\sigma_{v_{x-1,x-2}} - \frac{\theta+\rho(x-1)}{\theta+\rho(x-2)}\sigma_{v_{x,x-2}} - \frac{\theta+\rho x}{\theta+\rho(x-1)}\sigma_{v_{x-1}}^2$$

These two expressions can be simplified considerably under three assumptions: 1) ρ is small relative to θ or x is large, so that $\theta+\rho x \approx \theta+\rho(x-1)$; 2) the variance and serial covariances of v_{ix} are constant over time; 3) v_{ix} follows an order-one autoregressive process:

$$(18)\ v_{ix} = \delta v_{ix-1} + \xi_{ix}$$

Where δ is a parameter and ξ_{ix} is a random disturbance which is uncorrelated with v_{ix} and v_{ix-1}. Under these three assumptions, the following approximations can be made:

$$(19)\ \sigma_v^2 \approx \frac{1}{\sigma_{\varepsilon\varepsilon}}$$

$$(20)\ \sigma_{vv} \approx \frac{\sigma_\varepsilon^2 - \sigma_{\varepsilon\varepsilon}}{\sigma_\varepsilon^2 \sigma_{\varepsilon\varepsilon}}$$

Where: σ^2_v and σ_{vv} denote the variance and first-order covariance (respectively) of the error term in Equation (10); σ^2_ε and $\sigma_{\varepsilon\varepsilon}$ denote the variance and first-order covariance of the error term in Equation (2). The values defined by Equations (19) and (20) can be computed for every level of experience x, although I prefer to avoid the experience subscript because the assumptions made for the calculation imply that all the parameters defined in these two equations are independent of time. However, computing the estimated value of σ^2_v and σ_{vv} for every level of experience is important empirically, because it provides a way to check whether this assumption is consistent with the empirical results.

To assess whether ρ is small relative to θ (one of the sufficient conditions for the first assumption to hold) a possible way is to estimate all the parameters of Equation (10): ρ, θ, γ and β. To do so, I need the estimate of c_{sx} and c_{wx} for every level of experience, and an estimate for σ^2_v and σ_{vv}. Every level of experience must be considered as a unit of observation. Then, it is possible to estimate the system of equations composed by the two Equations (12) and (13) by non-linear least squares. In principle, this would allow to estimate also the correlation between underlying ability and schooling, σ_{as}. This would be an extremely interesting parameter. In practice, however, changes in σ_{as} lead to absolutely minimal changes in the fit of the estimation of Equations (12) and (13), so I prefer to leave σ_{as} as unknown and exogenous. Hence, I present estimates of ρ, θ, γ and β by assigning exogenously different values to σ_{as}, together with standard errors can be derived by bootstrapping (100 repetitions). This procedure produces some suggestive estimates of all the parameters of the model with dynamic returns to schooling and ability, as exemplified in Equation (10). This provides some suggestive evidence upon which to base an empirical judgement of whether the assumption that ρ is small relative to θ is likely to hold.

This empirical procedure is also, to the best of my knowledge, the first in the labour economics literature to estimate the relative size of the contribution of schooling and overall, underlying individual ability to salary and salary growth. However, they should be taken extremely cautiously, for at least three reasons: the sample size is small (the levels of experience used in the estimation do not exceed 30); using the levels of experience as units of observation leads to what is in fact a cross-sectional estimation, so that a number of unobservable factors (such as cohort effects) impact the coefficients estimates; and, finally, they are induced by imposing a specific theoretical data generation process (Equation (10) on the data). Hence, it is advisable not to rely on them to form an evidence-based empirical judgement on the relative size of the contribution of schooling and underlying individual ability to salary and salary growth.

4. Data

In order to estimate the model that was outlined above, I use data from the German Socioeconomic panel dataset, waves 2004 to 2011. In total, this dataset has 166826 observation of 35956 individuals. I keep observations with available information on years of education, work experience, gross monthly salary, and salary in the previous year. This means that I discard observations if the individual works and the monthly salary is not reported, but also if the individual does not work. I drop outliers with monthly salary lower than 400€ or higher than 15000€, or reporting to have worked less than 12 or more than 60 hours per week. I also drop individuals with more than 35 years of work experience, as they are probably a group of individuals for which there is a strong self-selection in non-retiring. Finally, I keep only male workers in sectors different from education, health and public administration. The choice of retaining only male workers is due to the very different dynamics followed by the careers of female workers, which career is often substantially affected by the birth of a child (e.g. Zhelyazkova, 2015). The sample size reduces to 20895 observations for 4318 individuals. I generate hourly wage by dividing the monthly salary by the reported number of hours worked and deflating at the inflation rate reported by Eurostat (2013). The dependent variable in the empirical model is then log hourly wage.

Since the coefficients are estimated separately for each level of work experience, it is interesting to see how many observations correspond to each level of experience. The number of observations is different for the OLS estimation and the IV estimation, because the latter requires the two-periods lagged salary to be available. The sample size by level of experience for OLS and IV estimations is reported in Table 1 (experience is rounded to the nearest integer). The sample size is small for the first years of work experience, but it is always larger than 140 if work experience exceeds two years.

Table 1 Number of observations by level of work experience

	OLS			IV		
Experience	Freq.	Percent	Cum.	Freq.	Percent	Cum.
1	24	0.1	0.1	0	0	0
2	245	1.5	1.6	0	0	0

3	281	1.7	3.3	140	1.2	1.2
4	314	1.9	5.2	170	1.5	2.8
5	340	2.1	7.3	197	1.8	4.5
6	377	2.3	9.5	213	1.9	6.4
7	374	2.3	11.8	246	2.2	8.6
8	404	2.4	14.2	265	2.4	10.9
9	422	2.6	16.8	274	2.4	13.4
10	452	2.7	19.5	302	2.7	16
11	468	2.8	22.3	315	2.8	18.8
12	470	2.8	25.2	329	2.9	21.7
13	514	3.1	28.3	347	3.1	24.8
14	531	3.2	31.5	359	3.2	28
15	546	3.3	34.8	365	3.2	31.3
16	574	3.5	38.2	387	3.4	34.7
17	583	3.5	41.7	415	3.7	38.4
18	585	3.5	45.3	416	3.7	42.1
19	584	3.5	48.8	410	3.6	45.7
20	576	3.5	52.3	413	3.7	49.4
21	607	3.7	55.9	421	3.7	53.1
22	628	3.8	59.7	442	3.9	57
23	619	3.7	63.5	450	4	61
24	607	3.7	67.1	433	3.8	64.8
25	599	3.6	70.7	434	3.9	68.7
26	583	3.5	74.2	432	3.8	72.5
27	572	3.5	77.7	409	3.6	76.2
28	544	3.3	81	397	3.5	79.7
29	512	3.1	84.1	376	3.3	83
30	497	3	87.1	350	3.1	86.1
31	477	2.9	89.9	347	3.1	89.2
32	443	2.7	92.6	336	3	92.2
33	438	2.6	95.3	312	2.8	94.9
34	411	2.5	97.7	284	2.5	97.5
35	376	2.3	100	286	2.5	100
Total	16,577	100		11,272	100	

5. Results

Dynamic panel data models estimates

In order to give a first idea of the magnitude of the parameters, I run an OLS regression which resembles the classical Mincerian regression of log wage on years of schooling, work experience, and squared work experience, with the difference that the lagged salary of the worker is also included. This empirical model is somewhat comparable to the OLS estimations run by Andini (2013b) on a sample of Belgian workers for the years 1994-2001, with the difference that in this model schooling is standardised and wage is expressed in deviations from the mean for a given level of experience. The estimated coefficients (with standard errors in brackets) are reported in the first column of Table 2. The coefficient for schooling is positive and significant suggesting that on average, given an equal salary at time t, a more educated worker can expect a higher salary at time t+1, compared to a less educated worker. The coefficient is 0.036, but the variable is standardized; the un-standardised coefficient is 0.014, close to Andini's (2013b) estimate of 0.016. The coefficient of lagged salary, or w_{ix-1}, is 0.85, a decimal point higher than Andini's (2013b) estimate, but still substantially and significantly smaller than 1. The coefficients for work experience are not significantly different from 0 at the 5% confidence level, which was to be expected given that the log wage of a worker is expressed as a difference with the average log wage across workers with the same level of work experience.

The second column of Table 2 reports results from an IV regression, where lagged salary is instrumented with schooling and two-periods lagged salary. The coefficient for schooling decreases substantially, but it remains positive and significant. It is equal to 0.012 (un-standardised: 0.005). The coefficient for lagged salary is still significantly different from one, although it is very close to one. Experience and its square are never significant at a 5% level, neither in the OLS, nor in the IV regression. Notice that the instrument in the IV estimation is not exogenous, as the two-periods lagged salary is significantly correlated to the current salary conditional on schooling, work experience and the one-year lagged salary. Hence, the coefficients for the IV estimation must suffer from a bias. Although this bias is not discussed specifically for the estimations reported in Table 2 (where observations with all levels of experience are pooled in the same sample), Section 3 shows that when the estimate is carried out at each level of experience there is, under reasonable conditions, an upward bias for the coefficient of schooling, and a downward bias for the coefficient of the lagged wage. The bias is likely to be present both for OLS and for IV, but to be more severe for the OLS estimation.

Table 2 Results from an OLS and IV estimation of a dynamic wage model for all levels of experience

	OLS	IV
w_{x-1}	0.85*** (0.007)	0.97*** (0.010)
S	0.036*** (0.002)	0.012*** (0.003)
X	0.0012 (0.009)	-0.002* (0.001)
x^2	-0.00004* (0.00002)	0.00002 (0.00003)
const.	0.41*** (0.014)	0.12*** (0.028)
R^2	0.81	0.80
Obs.	16577	11272

*** Significant at the 1% confidence level
** Significant at the 5% confidence level
* Significant at the 10% confidence level

Equation (11) suggests that the coefficients for schooling and lagged wage may not be the same for all levels of experience. Because of this reason, I estimated the coefficients separately for each level of experience. The results are reported in Figure 1, Figure 2, Figure 3 and Figure 4. In each of these figures, the coefficient is reported on the vertical axis, and experience is on the horizontal axis.

The coefficient for schooling decreases steadily with work experience. Whereas it is always positive when estimated with OLS, the coefficient is negative for workers with a high level of work experience, when estimated with IV. In contrast, the coefficient for lagged wage increases with work experience. The increase in the coefficient is very sharp for the first six years of work experience, and less pronounced afterwards. The coefficient is always lower than one when estimated by OLS, but it exceeds one for high levels of work experience when estimated by IV.

These patterns suggest that the model with dynamic returns to ability and schooling approximates reality better than the classic Mincer model. The reason is that the classic Mincer model predicts that both coefficients are constant across different levels of work experience, as long as the variance and first-order covariance of the error term in Equation (3) does not change dramatically over time. The data suggest that variance and covariance of the error term do not vary much over time. This can be seen from the fact that the variance for the error term derived from estimating Equation (2) is quite stable after the first years of work experience. This is shown for OLS in Figure 5, but the pattern is similar for the IV results. Figure 5 shows how the standard deviation of the error term is larger for the first 10 years of work experience, but then stabilizes around an average of 0.19. The first-order serial covariance of the error term in Equation (2), shown in Figure 6, stabilizes after just a few years of work experience around a mean of –0.011.

The coefficient of previous wage is increasing with time, and it is most often significantly lower than 1 when using OLS (middle panel of Figure 1). However, the pattern changes when using IV (middle panel of Figure 2). In this case, the coefficients become close to 1 after 10-13 years of experience, and they slightly bigger (but never significantly) than 1 thereafter.

One explanation for the significant difference of the estimated coefficient for lagged wage from 1 in the early stages of the workers' careers could be that the auto-correlation of the error term of Equation (11) induces a strong bias in the estimated coefficients. In the beginning of the working career, the salary growth of a worker could be strongly affected by his initial position in the company, or by a wrong individual-job match. This would imply a strong serial correlation of the error term and, as a consequence, a big downward bias in the coefficient of previous salary (also when estimated with IV). In terms of the notation previously used, $\sigma_{vt,t-1}$ would be substantially larger than 0 for workers in the first years of their careers. When the career path of the worker is established, on the contrary, the error term could be mainly reflecting factors like the exact time at which a promotion occurs, or at which the worker accept an offer from a competing company. Hence, the bias would gradually disappear with increasing work experience. This would also explain why the estimated coefficients reported in Figure 2 and Figure 4 are increasing with work experience, rather than declining (as predicted by Equation (11)).[3]

[3] The alternative explanation is that there are some ability traits that would be very influential in determining the initial salary of the worker, but that quickly deteriorate after the individual starts to work.

Figure 1 Estimated schooling coefficients for different levels of work experience, OLS method

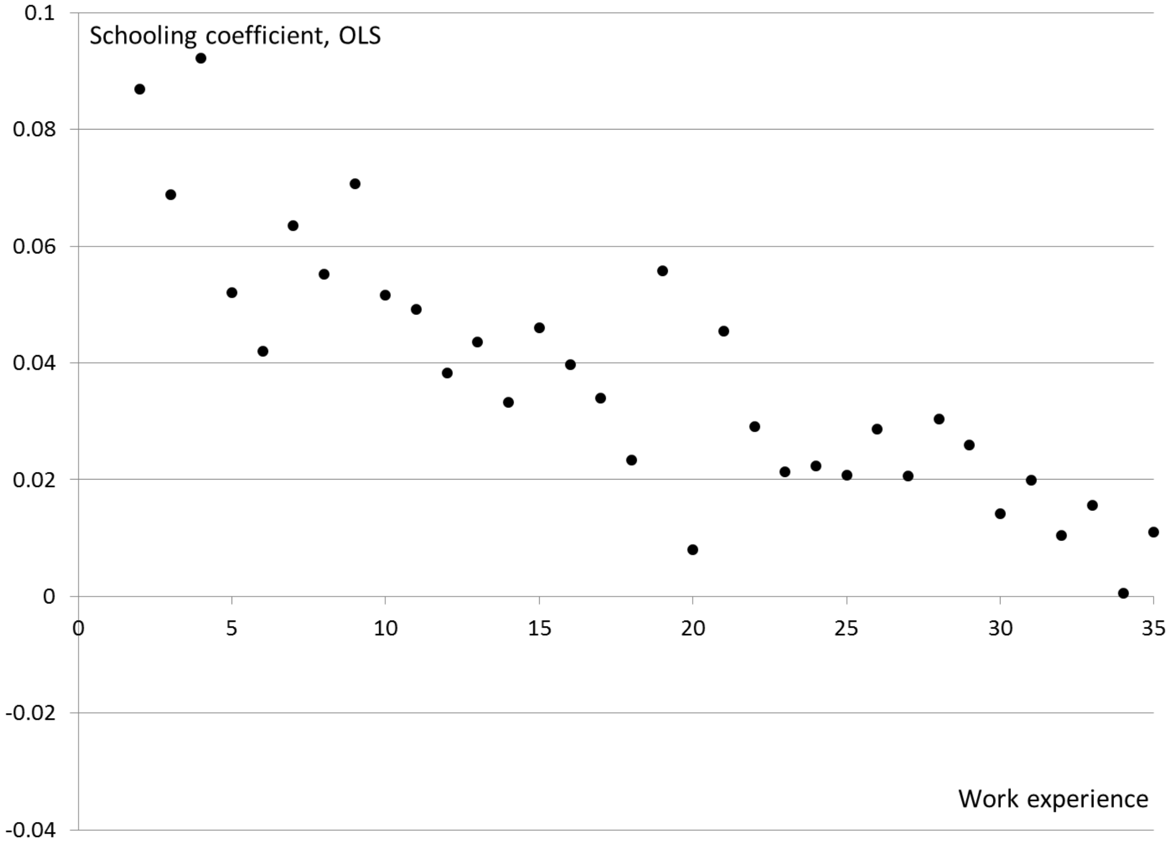


Figure 2 Estimated lagged wage coefficients for different levels of work experience, OLS method

Figure 3 Estimated schooling coefficients for different levels of work experience, IV method

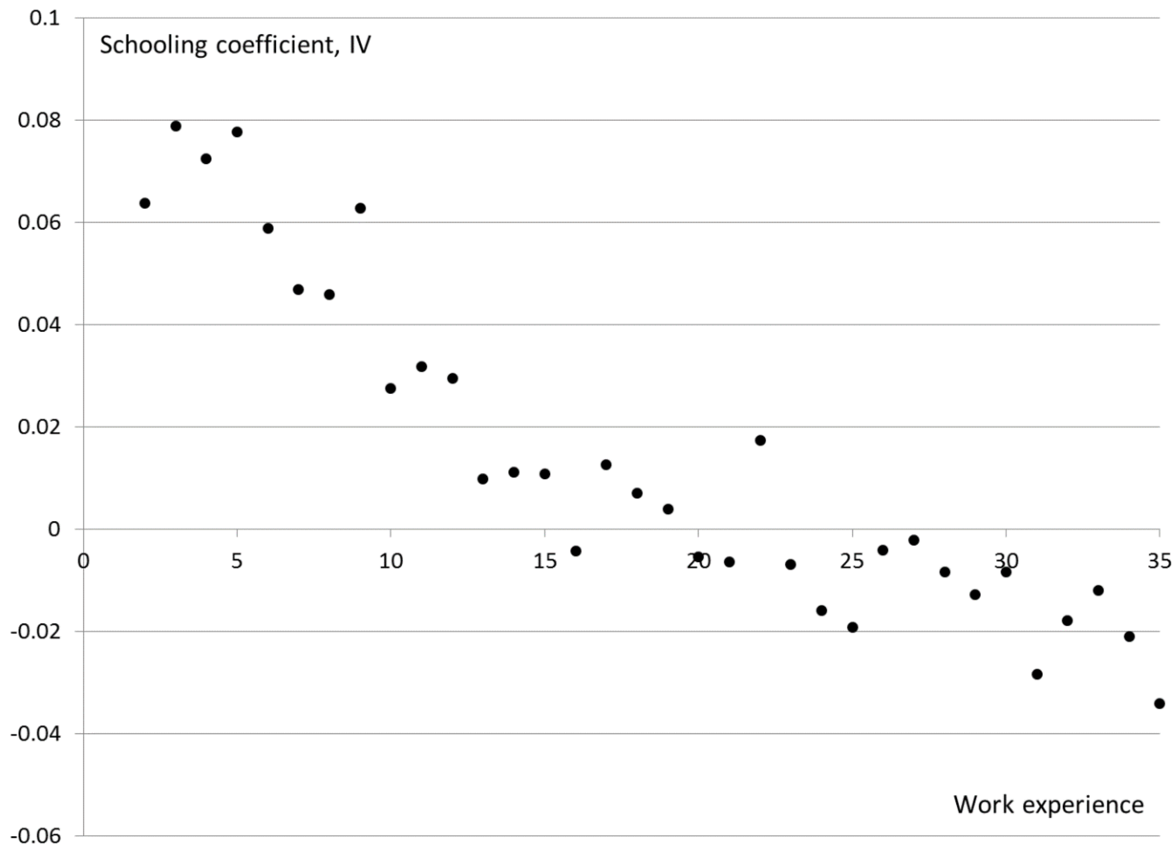


Figure 4 Estimated lagged wage coefficients for different levels of work experience, IV method

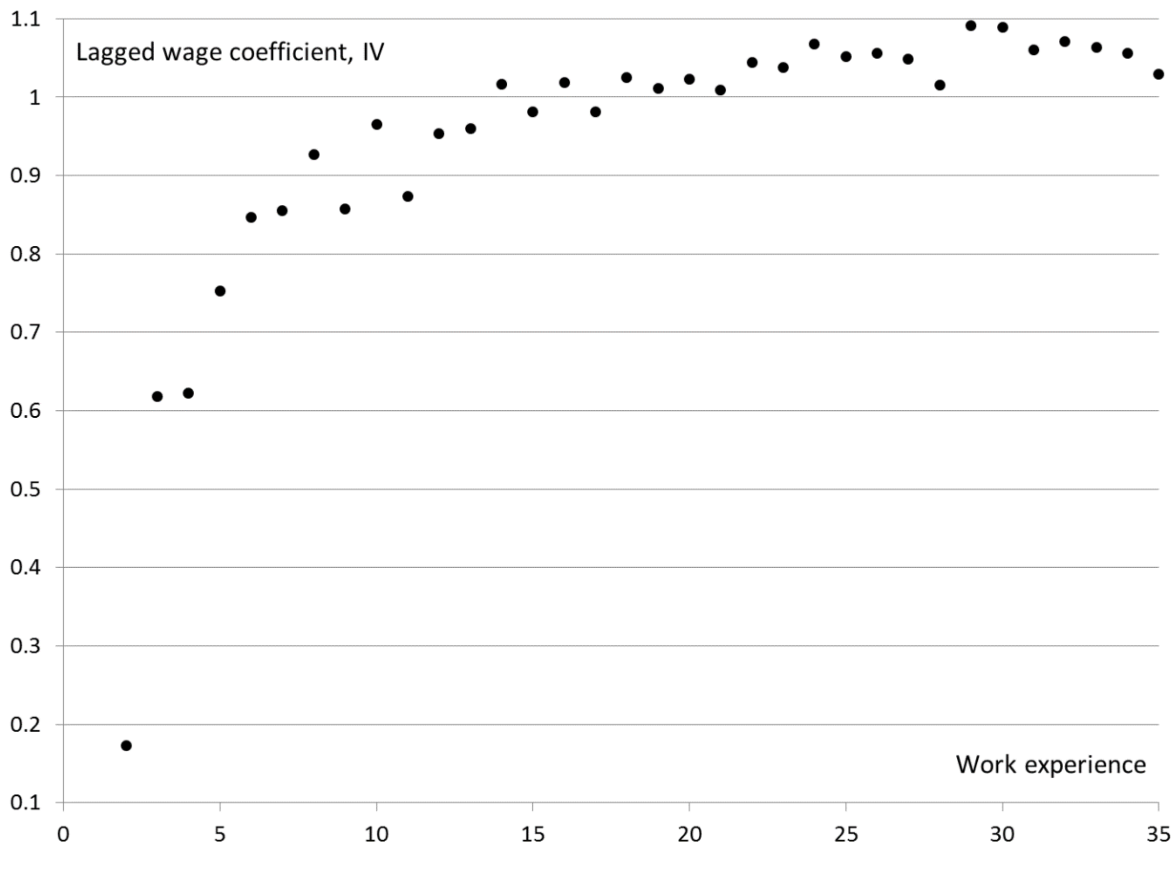


Figure 5 Standard deviation of the estimated error term for different levels of work experience, OLS method

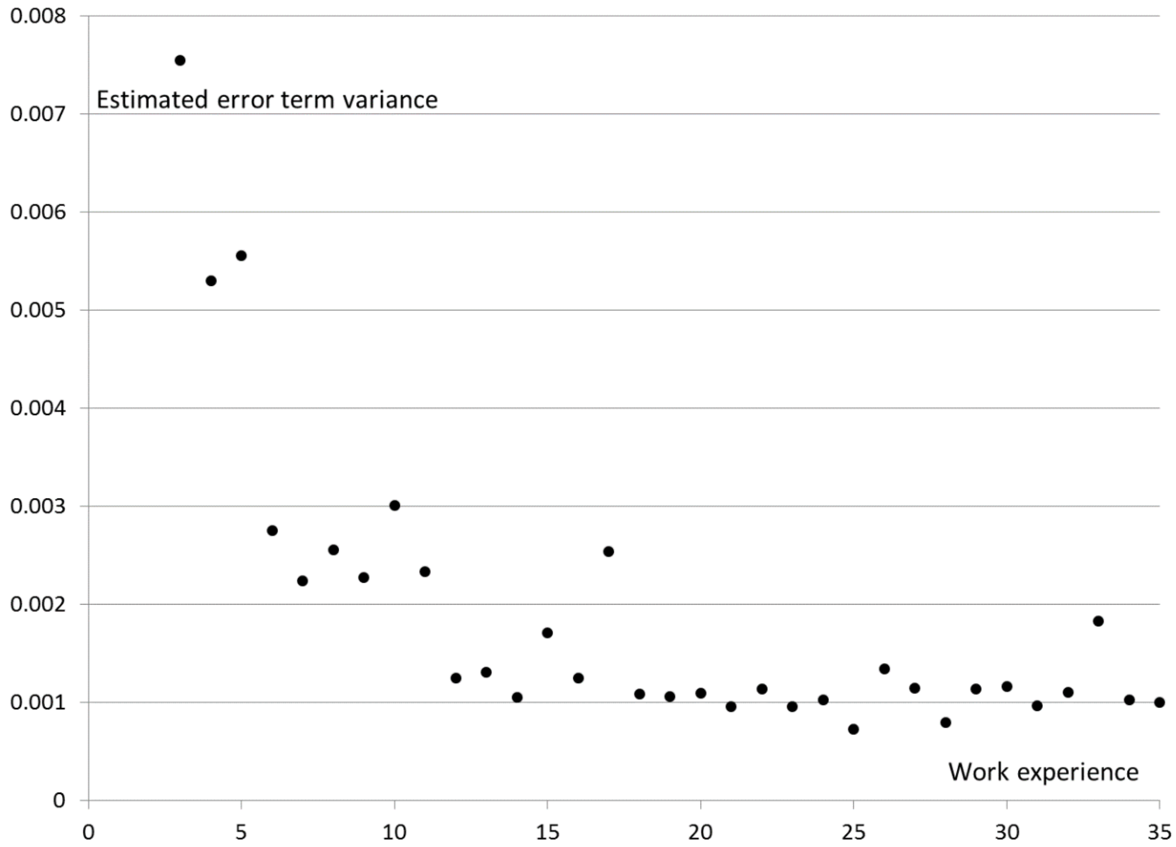
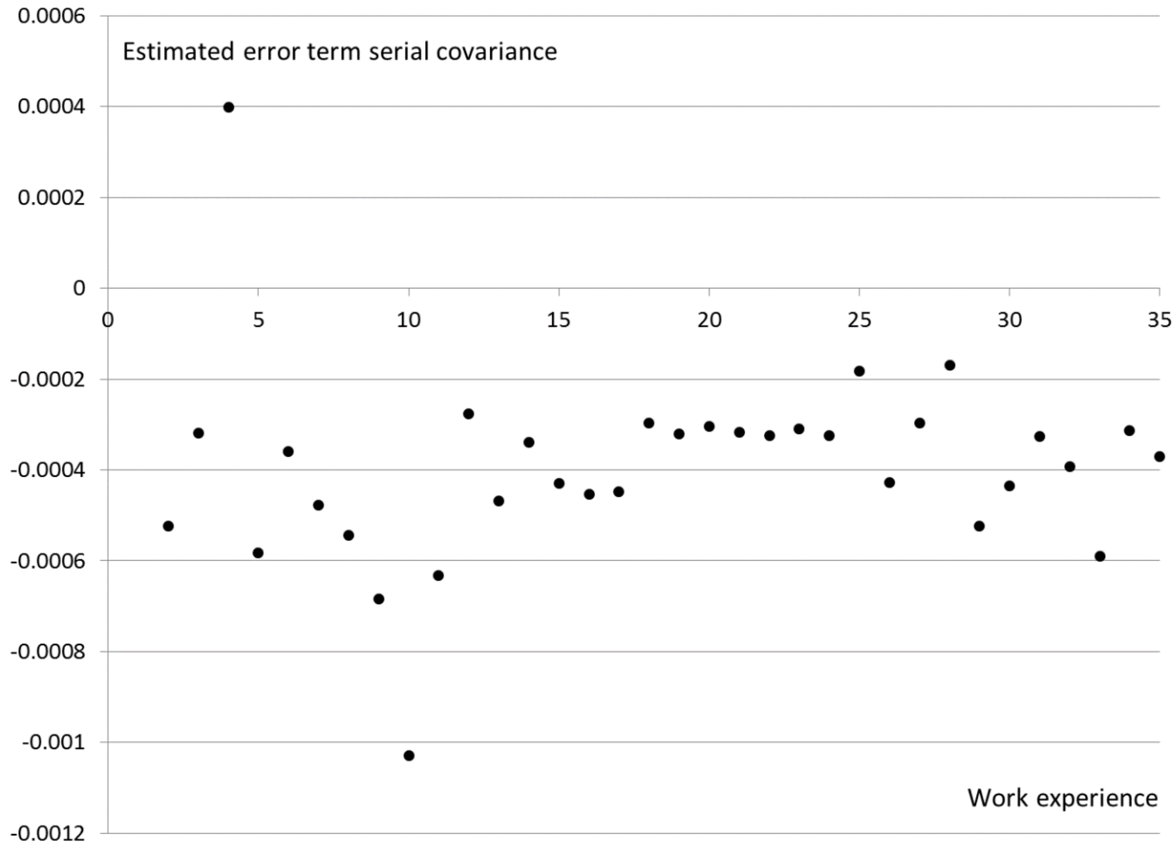


Figure 6 First-order serial covariance of the estimated error term for different levels of work experience, OLS method



Suggestive estimates of the parameters of model of dynamic returns to schooling

Figure 7 and Figure 8 show the parameters of Equation (10) obtained through a non-linear least squares estimation of the system of Equations (12) and (13). The estimated value for the parameter is reported on the vertical axis, while on the horizontal axis lies an interval of values for the assumed correlation between schooling and ability, $0 \leq \sigma_{as} \leq 0.8$. In other words, I change exogenously the value of the correlation between underlying ability and schooling between 0 (unrelated) to 0.8 (highly, positively correlated) and report in Figure 7 the respective values of the parameters ρ and γ, and in Figure 8 the respective values for θ and β. Together with the estimated parameters (black lines), Figure 7 also reports the lines lying two standard deviations above and below the estimated coefficients (grey lines). For the estimation, I use only levels of experience equal or larger than 10, as figures 5 and 6 shows that the error term variance and covariance seems to stabilise around that level of experience.

The results suggest that that the assumption that ρ is small relative to θ is likely to hold, as the ratio between the two parameters is constant over σ_{as} and equal to 0.049 (although θ is not precisely estimated). Both estimated parameters are positive. θ increases from about 0.2 (σ_{as}=0) to about 0.3 (σ_{as}=0.8), but it is never more than two standard deviations larger than 0. ρ increases from about 0.009 to about 0.015 across the same range, and it is always more than two standard deviations larger than 0. The estimated value for β is larger than could be expected, as it increases with the correlation between schooling and ability from about 0.4 to about 0.65. This would imply an un-standardised value of 0.15 to 0.25, which is much larger than the typical wage premium associated to one year of education, which is usually about 6-10% (Psacharopoulos and Patrinos, 2004). The estimated value for β also seems very large (about twice as large) compared to θ. Conversely, the coefficient for the dynamic returns to schooling, γ, is very close to 0 and varies from about 0 to about 0.09. γ is very small also in relation to its estimated standard error, which ranges approximately from 0.03 to 0.05.

Figure 7 Suggestive estimates for ρ and γ (black lines), with related 95% confidence intervals (grey lines)

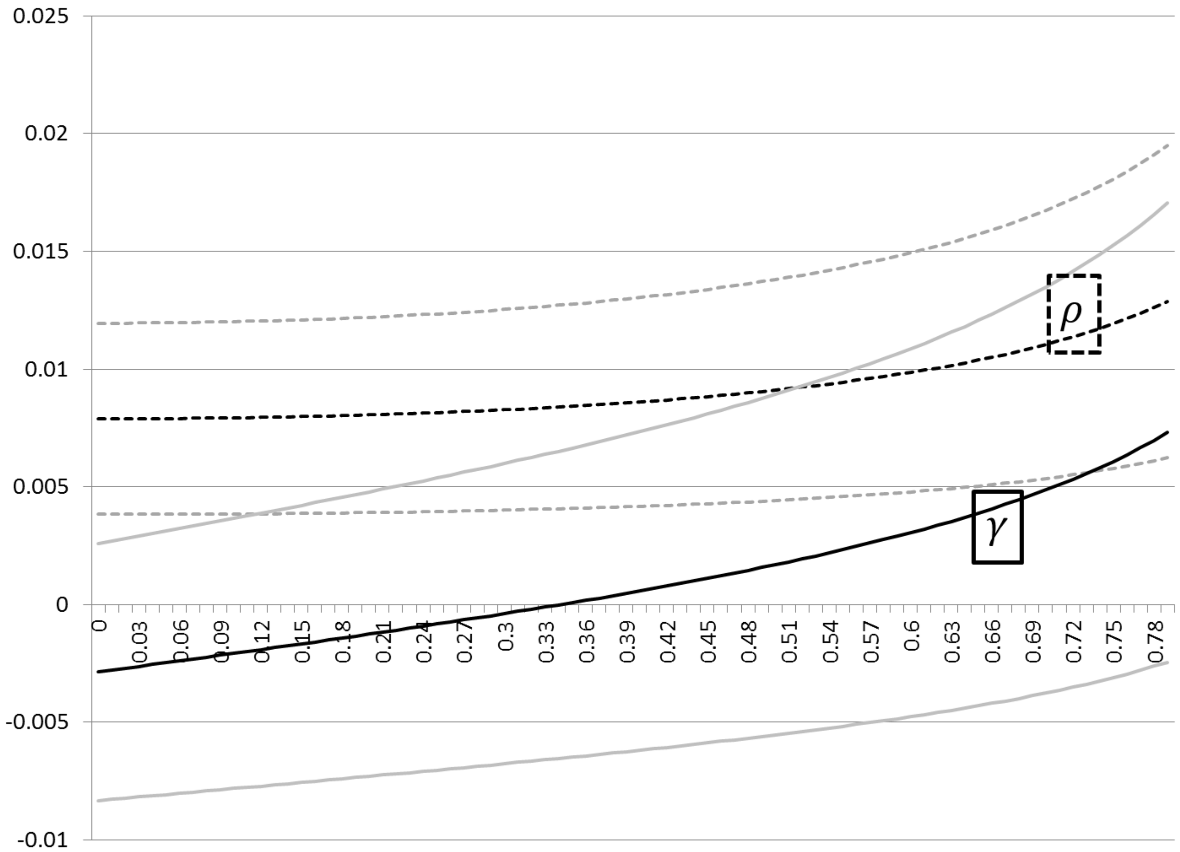
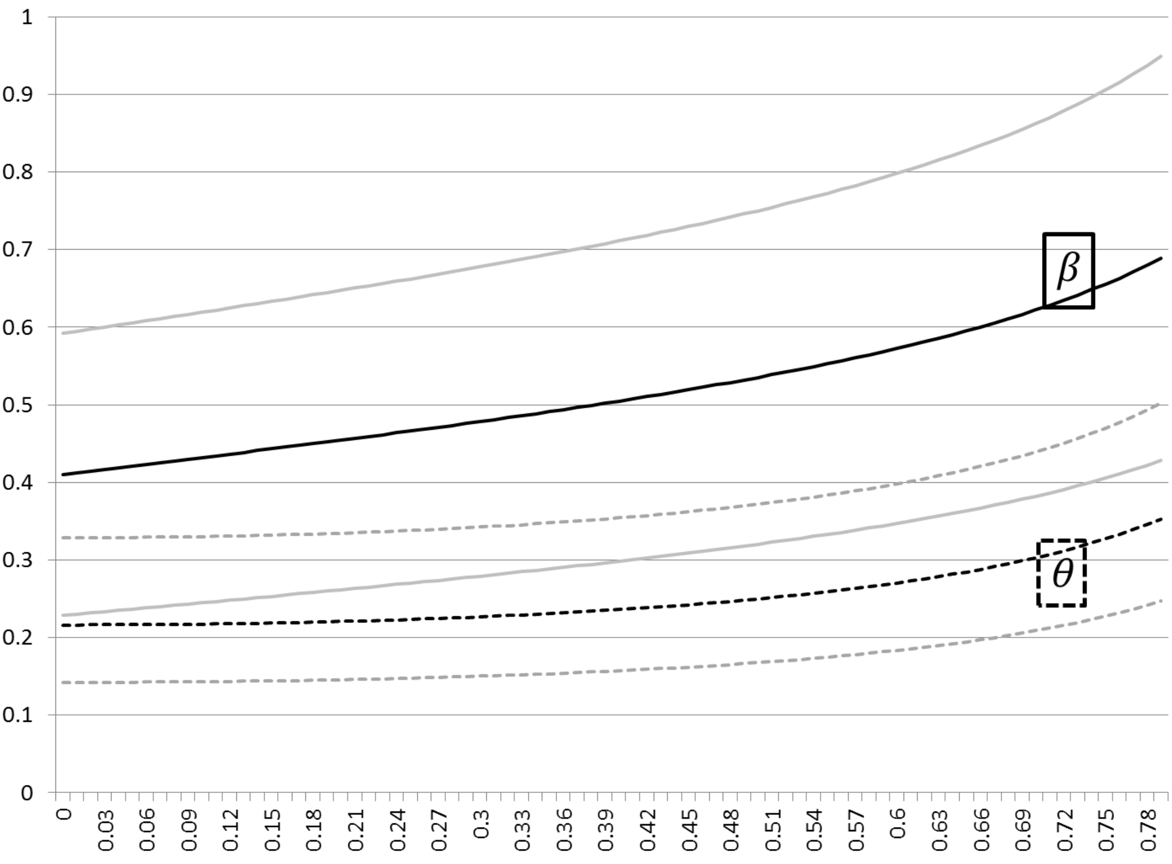


Figure 8 Suggestive estimates for θ and β (black lines), with related 95% confidence intervals (grey lines)



6. Conclusions

In general, the results of my analysis of the results from a sample of 17841 German male workers suggest the following conclusions. i) The coefficients for schooling and lagged wage in a dynamic panel data model of the earnings curve display substantial variation by work experience. This suggests that using a pooled sample of workers at different levels of experience for a dynamic panel data regression (e.g. Andini, 2013b) can be misleading. For workers at advanced levels of work experience, the IV coefficient for lagged log wage is close to 1 and the IV coefficient for schooling is close to 0. ii) The (estimated) serial correlation and standard deviation of the earnings equation's error term is much higher and more unstable for workers in their first years on the labour market, suggesting that this phase of the workers' careers is structurally different from the later phases. This gives credit to the efforts to analyse separately the labour market experiences of young workers from an academic and policy perspective (Allen and van der Velden, 2007). iii) The Mincer model with underlying ability is not consistent with the finding of a high variability of the estimated coefficients by level of experience; the model of dynamic returns to schooling and ability, in contrast, is not inconsistent with this variability.

Among other limitations, the sample used for the estimations presented in this paper is composed only of male workers living in Germany. It would be interesting to extend the analysis to other economies and, perhaps more importantly, to female workers.

References:

Allen, J. and van der Velden, R. (2007). The Flexible Professional in the Knowledge Society: General Results of the REFLEX Project. ROA, Maastricht.

Andini, C. (2007). Returns to Education and wage equations: A dynamic approach. *Applied Economic Letters*, Vol. 14(8), pp. 577-579.

Andini, C. (2012). How well does a dynamic Mincer equation fit NLSY data? Evidence based on a simple wage-bargaining model. *Empirical Economics*, DOI: 10.1007/s00181-012-0581-5.

Andini, C. (2013a). Persistence bias and the wage-schooling model. *IZA Discussion Papers*, no. 7186.

Andini, C. (2013b). Earnings persistence and schooling returns. *Economics Letters*, Vol. 118, pp. 482-484.

Arrazola, M. and de Hevia, J. (2006). More on the estimation of the human capital depreciation rate. *Applied Economics Letters*, Vol. 11(3), pp. 145-148.

Altonij, J. and Pierret, C. (2001). Employer learning and statistical discrimination. *The Quarterly Journal of Economics*, Vol. 116(1), pp. 313-350.

Arcidiacono, P., Bayer, P. and Hizmo, A. (2008). Beyond signaling and human capital: Education and the revelation of ability. *NBER Working Paper Series*, No. 13951.

Bauer, T. and Haisken-deNew, J. (2001). Employer learning and the returns to schooling. *Labour Economics*, Vol. 8, pp. 161-180.

Brunello G. (2001). On the complementarity between education and training in Europe. *IZA Discussion Papers*, No. 309.

Brunello, G. and Comi, S. (2004). Education and earnings growth: Evidence from 11 European countries. *Economics of Education Review*, Vol. 23, pp. 75-83

Greene, W. (2008). *Econometric Analysis* (6th ed.). Upper Saddle River (NJ): Pearson Prentice Hall.

Chevalier, A., Harmon, C., Walker, I. and Zhu, Y. (2004). Does education raise productivity, or just reflect it? *The Economic Journal*, Vol. 114, pp. 499-517.

Harmon, C., Oosterbeek, H., and Walker, I. (2003). Returns to education: Microeconomics. *Journal of Economic Surveys*, pp. 115-156, Vol. 17.

Heckman, J., Lochner, L. and Todd, P. (2006). Earning functions, rates of return and treatment effects: The Mincer equation. In Hanushek and Welch (eds.), *Handbook of the economics of education*, pp. 307-458. Amsterdam: North Holland.

Mincer, J. (1974). *Schooling, experience, and earnings*. Available at: www.nber.org

Neuman, S. and Weiss, A. (1995). On the effects of schooling vintage on experience-earnings profiles: Theory and evidence. *European Economic Review*, Vol. 39, pp. 943-955.

Psacharopoulos, G. and Patrinos, H. (2004). Returns to investment in education: a further update. *Education Economics*, Vol. 12(2), pp. 111-134.

Psacharopoulos, G. and Layard, R. (1979). Human capital and earnings: British evidence and a critique. *Review of Economic Studies*, Vol. 46 (3), pp. 485-506.

Weber, S. (2009). Human capital depreciation and education level: Some evidence for Switzerland. *Université de Genève, Paper Serie Compétences des adultes et formation professionnelle*.

Zhelyazkova, N. (2015). *Work-Family Reconciliation and Use of Parental Leave in Luxembourg*. Maastricht: Boekenplan. Retrieved from http://www.merit.unu.edu/publications/uploads/1424769568.pdf