Does the observed return to schooling converge to the potential one?
Evidence based on a dynamic Mincer equation with individual unobserved heterogeneity

Corrado Andini

Universidade da Madeira, Departamento de Gestão e Economia, 9000-390 Funchal, Portugal
Centro de Estudos de Economia Aplicada do Atlântico (CEEAplA), 9501-801 Ponta Delgada, Portugal
Institute for the Study of Labor (IZA), D-53072 Bonn, Germany

ABSTRACT

This paper shows how to estimate the average potential and observed wage return to schooling in a dynamic Mincerian model with individual unobserved heterogeneity. It is shown that the observed return is lower than its potential level at the beginning of the working life but converges to the latter as time goes by.

Keywords: Mincer Equation, Wages, Human Capital.
JEL codes: I21, J31.

Corresponding address:
Prof. Corrado Andini, Universidade da Madeira, Campus da Penteada, 9000-390 Funchal, Portugal.
e-mail: andini@uma.pt

1. Introduction

In 1974, Jacob Mincer published a seminal book that has been the starting point of a large body of literature dealing with the estimation of a linear model where the logarithm of the hourly wage of an individual is explained by his schooling years, labor-market experience, experience squared, and the monetary value of the individual ability at birth, which is not observable.

In spite of its wide acceptance within the profession, the spread of the framework developed by Mincer (1974) over the last forty years has not been uncontroversial. Some authors criticized it by arguing that the Mincer's model is not able to provide a good fit of empirical data. Some stressed that the average effect of schooling on earnings is likely to be non-linear in schooling. Some suggested that education levels should replace schooling years in the wage equation. Other authors proposed different arguments questioning the original Mincer's model. As a matter of example, Murphy and Welch (1990) maintained that the standard Mincer equation provides a very poor approximation of the true empirical relationship between earnings and experience, Trostel (2005) argued that the average impact of an additional year of schooling on earnings varies with the number of completed years of schooling, while Belzil (2007) argued that schooling and experience are not separable in a wage equation.

Looking at the big picture, however, besides some critical voices, the history of human-capital regressions has been characterized by a generalized attempt of consistently estimating the coefficients of the Mincer equation, under an implicit acceptance of the theoretical setup of the model.

As stressed by Polachek (2007), at present, several survey articles have been written on the Mincer earnings function. Perhaps three of the most popular have been authored by Card (1999), Heckman, Lochner and Todd (2003), and Lemieux (2006). One common feature of the reviewed works is that the empirical models have a static nature. Putting it differently, as shall be seen in the next section, they implicitly assume that the observed wage of an individual is equal to the monetary value of the individual human-capital productivity at any point in time. What actually changes from one study to another is the way the monetary value of the individual human-capital productivity is modeled. This paper tackles the issue of the estimation of the Mincer model from a different, dynamic perspective. Let us focus on it.

The monetary value of the individual human-capital productivity, or potential wage, defines what an individual may potentially earn because of his observed human-capital skills and his unobserved ability. In the basic specification, it is conceived as a linear function of ability, schooling, experience and its square. In more complex specifications, it is modeled using combinations of these variables, for instance due to complementarities between schooling and experience, or additional variables[1]. Yet, regardless of how the potential wage is modeled, the standard practice in Mincerian studies so far has been to assume that the equality between the potential wage of an individual and his observed wage happens at any point in time. As will be argued in the next section, the assumption follows directly from the original Mincer's model.

In this paper, building on earlier studies by Andini (2007, 2009, 2010 and 2012), we relax the above-mentioned assumption of equality by allowing an adjustment between observed and potential wages to take place over time. This leads to a dynamic Mincer equation where past wages play the role of additional explanatory variable. This equation allows us to measure the average adjustment speed of observed wages to

---

[1] These variables typically include individual observed characteristics such as birth cohort, occupation, sector of activity, marital status, gender, health, race, residence, etc… Sometimes these variables include indicators of labor-market mismatch, over-education or other factors affecting the individual potential wage.

human-capital productivity and also to analyze the implications of this adjustment for the estimation of the average wage return to schooling.

To highlight the specific contribution of this paper, we should first note that a dynamic Mincer equation has been already estimated by Andini (2007, 2009, 2010 and 2012) using data from the United States, Spain and Portugal. In particular, from 2007 to 2010, Andini controlled for observed heterogeneity and used quantile-regression techniques to inspect the impact of schooling not only on the mean but also on the shape of the conditional wage distribution. Instead, in 2012, the author provided, to the best of our knowledge, the first attempt to estimate a dynamic Mincer equation just focusing on the mean but controlling for *both observed and unobserved* heterogeneity. Yet, Andini (2012) did not discuss the implication of his approach for the computation of the return to schooling. In this paper, we build on Andini (2012) by providing the first attempt to estimate the average wage return to schooling in a dynamic Mincerian model with both observed and unobserved heterogeneity. In addition, the dynamic model presented here is obtained under weaker theoretical assumptions than in Andini (2012).

To discuss our main points, we use a limited set of observed controls, namely the past wage of the individual and the three classical human-capital variables mentioned before: schooling, experience and experience squared. The main reason is that, as shall be seen in Section 3, the first step towards a dynamic approach is to check whether the data actually reject the above-referred Mincer's assumption of equality between potential and observed wages.

Nevertheless, in using a simple specification, we follow important contributions to the literature such as those of Buchinsky (1994) or Martins and Pereira (2004), among others. However, we extend their sets of observed controls using one lagged wage.

The rationale for a simple specification is discussed in Andini (2007) and is consistent with the main argument of Pereira and Martins (2004) in favor of the estimation of *total* returns to schooling[2]. Yet, in order to reduce skepticism about the significance of estimation results obtained using a simplified model, we also provide some robustness checks using additional control variables.

The empirical analysis presented in this paper explores data for Belgium extracted from the European Community Household Panel (ECHP). Since the main purpose of paper is to show how to compute average wage returns to schooling in a dynamic model with unobserved heterogeneity, the choice of the dataset is relatively unimportant. The main requirement is that the data to have a longitudinal individual-level structure.

Yet, we have chosen to explore data for Belgium as they allow us to test for a specific issue. In particular, Andini (2009) has shown that a dynamic Mincer equation can be seen as the result of a simple wage-bargaining model between a worker and an employer. In his model, the outside option of the worker depends on the level of the unemployment benefit which, in turn, depends on the past wage of the worker. Hence, the author predicts that a dynamic Mincer equation should fit the data well in countries where unemployment-benefit policies exist, cover a large share of the labor force, and

---

[2] Martins and Pereira (2004) argued in favor of a simple Mincer specification for estimating the total return to schooling. Since many variables that are normally used as controls, such as industry or occupational dummies, are choice variables that depend on education, controlling for these variables implies that a share of the impact of education on wages is captured by the coefficient of these education-dependent covariates. Of course, downsizing the wage equation is a risky exercise because the lower is the number of regressors, the likelier is the possibility that the coefficients are inconsistently estimated due to omitted-variable bias. Andini (2007) proposed a method for the estimation of the total return to schooling when longitudinal data are available. The introduction of past earnings as additional explanatory variable increases the explained variability of wages and reduces the risk of inconsistency *without* implying any additional difficulty for the issue of recovering the total return to education.

the benefits depend on past earnings. Since Belgium has the highest generosity index of unemployment benefit adjusted for coverage in a sample of 12 European countries (Boeri and van Ours, 2008, p. 283) and unemployment benefits are based on past wages and contributions, the use of Belgian data allows us to test the prediction of Andini (2009). As shall be seen further below, the empirical evidence presented in this paper does not reject the prediction.

The structure of the paper is as follows. Section 2 reviews the static Mincer's theory. Section 3 presents an adjustment model between observed and potential wages. Section 4 discusses issues and problems related to the estimation of an adjustment model when controlling for both observed and unobserved individual heterogeneity. Section 5 describes the dataset and the variables used in the empirical analysis. Estimation results are also presented. Section 6 provides a numerical example of how an adjustment model should be used to compute returns to schooling. Sensitivity analysis is also performed. Section 7 discusses the findings of the paper in the light of the existing literature on earnings dynamics. Section 8 concludes.

2. Static equality model

This section presents the theoretical foundations of the standard Mincer equation as reported by Heckman et al. (2003). Therefore, we make no claim of originality at this stage and mainly aim at helping the reader with notations and terminology adopted in the next sections.

Mincer argues that potential earnings today depend on investments in human capital made yesterday. Denoting potential earnings at time t as E_t, Mincer assumes that an individual invests in human capital a share k_t of his potential earnings with a return of r_t in each period t. Therefore we have:

(1) $E_{t+1} = E_t(1 + r_t k_t)$

which, after repeated substitution, becomes:

(2) $E_t = \prod_{j=0}^{t-1}(1 + r_j k_j)E_0$

or alternatively:

(3) $\ln E_t = \ln E_0 + \sum_{j=0}^{t-1} \ln(1 + r_j k_j)$.

Under the assumptions that:

- schooling is the number of years s spent in full-time investment in human capital ($k_0 = ... = k_{s-1} = 1$),

- the return to the schooling investment in terms of potential earnings is constant over time ($r_0 = ... = r_{s-1} = \beta$),

- the return to the post-schooling investment in terms of potential earnings is constant over time ($r_s = ... = r_{t-1} = \lambda$),

3

we can write expression (3) as follows:

$$(4) \qquad \ln E_t = \ln E_0 + s\ln(1+\beta) + \sum_{j=s}^{t-1} \ln(1+\lambda k_j)$$

which yields:

$$(5) \qquad \ln E_t \approx \ln E_0 + \beta s + \lambda \sum_{j=s}^{t-1} k_j$$

for small values of β, λ and k[3].

In order to build up a link between potential earnings and labor-market experience z, Mincer assumes that the post-schooling investment linearly decreases over time, that is:

$$(6) \qquad k_{s+z} = \eta\left(1 - \frac{z}{T}\right)$$

where $z = t - s \geq 0$, T is the last year of the working life and $\eta \in (0,1)$.
Therefore, using (6), we can re-arrange expression (5) and get:

$$(7) \qquad \ln E_t \approx \ln E_0 - \eta\lambda + \beta s + \left(\eta\lambda + \frac{\eta\lambda}{2T}\right)z - \left(\frac{\eta\lambda}{2T}\right)z^2.$$

Then, by subtracting (6) from (7), we obtain an expression for net potential earnings, i.e. potential earnings net of post-schooling investment costs[4]:

$$(8) \qquad \ln E_t - \eta\left(1 - \frac{z}{T}\right) \approx \ln E_0 - \eta\lambda - \eta + \beta s + \left(\eta\lambda + \frac{\eta\lambda}{2T} + \frac{\eta}{T}\right)z - \left(\frac{\eta\lambda}{2T}\right)z^2$$

which can also be written as:

$$(9) \qquad \ln npe_t \approx \alpha + \beta s + \delta z + \phi z^2 + \ln E_0$$

where $\ln npe_t = \ln E_t - \eta\left(1 - \frac{z}{T}\right)$, $\alpha = -\eta\lambda - \eta$, $\delta = \eta\lambda + \frac{\eta\lambda}{2T} + \frac{\eta}{T}$ and $\phi = -\frac{\eta\lambda}{2T}$.

Assuming that observed earnings are equal to net potential earnings at any time $t \geq s$ (a key-assumption, as shall be seen in the next section):

$$(10) \qquad \ln w_t = \ln npe_t$$

---

[3] Note that the symbol of equality (=) in expression (4) becomes a symbol of rough equality ($\approx$) in expression (5). It happens because, if a number x is close to zero, then $\ln(1+x) \approx x$.

[4] Note the post-schooling investment costs are given by $k_t E_t$ with $t \geq s$. Therefore, net potential earnings in levels are given by $E_t - k_t E_t$, or $E_t(1 - k_t)$ which, after taking logarithms, if k is small, is equal to $\ln E_t - k_t$, i.e. the left-hand side of expression (8).

4

and, using expression (9), we get:

(11) $\ln w_t \approx \alpha + \beta s + \delta z + \phi z^2 + \ln E_0$

By adding subscripts where necessary, we get:

(12) $\ln w_{it} \approx \alpha + \beta s_i + \delta z_{it} + \phi z_{it}^2 + \ln E_{0i}$

By making the model stochastic, we obtain:

(13) $\ln w_{it} = \alpha + \beta s_i + \delta z_{it} + \phi z_{it}^2 + \ln E_{0i} + e_{it}$

     Normally, the error e_{it} is assumed to be a pure well-behaved individual wage shock, uncorrelated with the explanatory variables. Instead, as $\ln E_{0i}$ represents the value of the individual potential earnings at birth, it is usually interpreted as the value of the individual unobserved ability and is therefore assumed be correlated with s_i and z_{it}. Hence, the estimation of model (13) is non-trivial.

     To conclude this section, it is important to stress that the total return to schooling in the static model (13) is given by the following expression:

(14) $\dfrac{\partial \ln w_{it}}{\partial s_i} = \beta$

and is constant over the working life, meaning *independent* of labor-market experience z. Further, because of assumption (10), the return to schooling in terms of observed earnings and the one in terms of net potential earnings coincide.

     We label β as 'the static return to schooling in terms of net potential earnings' and show, in Section 6, that our interpretation of β in terms of *net potential* rather than *observed* earnings is appropriate.

3. Dynamic adjustment model

If one takes as a starting point the presentation of the Mincer's model made in the previous section, it is possible to argue that the model is characterized by two main features. First, it provides an explanation why the logarithm of the net potential earnings of an individual at time $t = s + z$ can be approximately represented as a function of s, z and $\ln E_0$, i.e. expression (9). This expression can be seen as *the core* of the Mincer's model. Second, it is based on the assumption that, at any time $t \geq s$, the logarithm of the observed wage of an individual is equal to the monetary value of his net human-capital productivity, measured by his net potential wage, i.e. assumption (10).

     As anticipated in Section 1, there are at least three popular surveys on the Mincer equation: Card (1999), Heckman, Lochner and Todd (2003), and Lemieux (2006).

     Card concentrated on econometric issues regarding the identification of the causal relationship between schooling and earnings, and therefore he only marginally discussed whether the theory proposed by Mincer was able to provide a good fit of the real data. In contrast, Heckman et al. concentrated on the empirical support to the theory using past and current data (and on how to best incorporate future earnings uncertainty into the Mincer framework). Analogously, Lemieux focused on how well the most common version of the Mincer earnings function fits current data. Hence, for the

purpose of this paper, the surveys by Heckman et al. and Lemieux deserve special consideration.

On the one hand, Heckman et al. tested three implications of the Mincer model: i) log-earnings experience profiles are parallel across schooling levels (i.e. the return to schooling is independent of labor-market experience); ii) log-earnings age profiles diverge across schooling levels (i.e. the return to labor-market experience increases as age increases); iii) the variance of earnings over the life-cycle has a U-shaped pattern. Using Census data on white and black males, the authors found mixed evidence of these predictions. In general, it seems that more recent data are supporting Mincer's predictions less.

On the other hand, Lemieux found that the Mincer equation remains an accurate benchmark for estimating wage equations provided that it is adjusted by i) including a quartic function of potential experience instead of a quadratic one; ii) allowing for a quadratic term in years of schooling to capture the growing convexity in the relationship between schooling and wages; and 3) allowing for cohort effects to capture the dramatic growth in returns to schooling among cohorts born after 1950.

Summing up, these influential authors basically argued that equation (11) may have some problems to fit the most recent data and, in order to solve these problems, they suggested to modify (9) instead of relaxing (10). One of the aims of this paper is to show that *relaxing (10) is a possibility that is worth exploring more*.

Hence, unlike several studies in the existing literature, this paper does not question the core of the Mincer's theory, i.e. expression (9). Although expression (9) can be criticized, and has been criticized in the past, it has a feature that is very appreciated by the applied economist: it allows the estimation of a wage model that is linear in parameters (see model (13)). In addition, and most importantly, expression (9) is theoretically well-grounded while many departures from it are not (i.e. they are justified on empirical grounds). In this paper, we show that, assuming that (9) holds (an assumption made in hundreds of studies), one can actually obtain a better estimate of the return to schooling in terms of both observed and potential earnings by relaxing assumption (10) in a simple and flexible way.

The main argument to relax assumption (10) is as follows. As we have seen, Mincer suggested that, by investing in human capital, an individual can increase the monetary value of his productivity and achieve a certain level of net potential earnings. If the labor market were characterized by perfect competition at any point in time, the net potential earnings of an individual and his observed earnings would coincide at any point in time, as in assumption (10). That is, an individual would always earn the net monetary value of his human-capital productivity. However, without departing from the perfect-competition hypothesis in the long run, there may be frictions in the labor market in the short run that may cause the observed wages to adjust to the potential wages *with some lag*. In this case, the return to the individual human-capital investment measured in terms of observed earnings - say the observed return - may be different, at some point in time, from the return to the same investment measured in terms of net potential earnings - say the potential return.

This paper investigates the above hypothesis and shows that the observed return to schooling is substantially lower than its potential level at the beginning of the working life. Andini (2009) discussed one possible source of the above-referred frictions in the labor market, namely the existence of wage bargaining at worker-employer level in a world where unemployment benefits depend on past wages, but his empirical model did not control for individual unobserved heterogeneity. Other explanations are, of course, possible. Yet, in this paper, we do not wish to enter into a theoretical discussion about the nature of the frictions. What we find important is to document that observed wages adjust to potential wages with some lag even after controlling for individual unobserved heterogeneity. In addition, and more importantly,

we aim to discuss the consequences of this adjustment for the calculation of the return to schooling.

On the lines of Flannery and Rangan (2006) among others, we argue that assumption (10) can be replaced by a more flexible assumption. Particularly, observed earnings can be seen as dynamically adjusting to net potential earnings, according to the following simple adjustment model:

$$(15) \quad \ln w_t - \ln w_{t-1} = \rho(\ln npe_t - \ln w_{t-1})$$

where $\rho \in [0,1]$ measures the speed of adjustment.

If $\rho = 1$, then assumption (10) holds, observed earnings are equal (adjust) to net potential earnings at time t (within period t), and the standard Mincerian model (11) holds. If instead $\rho = 0$, then observed earnings are constant over time, always equal to the labor-market entry earnings $\ln w_s$, and do not adjust at all to variations of net potential earnings. In general, when the speed of adjustment is neither zero nor one, by replacing expression (9) into (15), we get:

$$(16) \quad \ln w_t \approx (1-\rho)\ln w_{t-1} + \rho(\alpha + \beta s + \delta z + \phi z^2 + \ln E_0)$$

or alternatively:

$$(17) \quad \ln w_t \approx \upsilon_0 + \upsilon_1 \ln w_{t-1} + \upsilon_2 s + \upsilon_3 z + \upsilon_4 z^2 + \rho \ln E_0$$

where $\upsilon_0 = \rho\alpha$, $\upsilon_1 = 1-\rho$, $\upsilon_2 = \rho\beta$, $\upsilon_3 = \rho\delta$ and $\upsilon_4 = \rho\phi$.
By adding subscripts where necessary, we get:

$$(18) \quad \ln w_{it} \approx \upsilon_0 + \upsilon_1 \ln w_{it-1} + \upsilon_2 s_i + \upsilon_3 z_{it} + \upsilon_4 z_{it}^2 + \upsilon_i$$

where $\upsilon_i = \rho \ln E_{0i}$.
By making the model stochastic, we get:

$$(19) \quad \ln w_{it} = \upsilon_0 + \upsilon_1 \ln w_{it-1} + \upsilon_2 s_i + \upsilon_3 z_{it} + \upsilon_4 z_{it}^2 + \upsilon_i + e_{it}$$

Expression (19) is a dynamic version of the Mincer equation, which we label as the 'adjustment model'. When individual-level longitudinal data are available, the complement to one of the speed of adjustment $(1-\rho)$ can be estimated and the theory underlying (19) can be tested.

The minimum requirement for the adjustment theory to be consistent with the data is to find that the coefficient υ_1 is significantly different from zero. If it were zero, then the original Mincer's model holds, and the investment in education is fully rewarded at any point in time during the working life. If it were one, then individual wages evolve as pure random walks, and the investment in education is worthless. In the intermediate case, the investment in education is worth in the long run. The lower is υ_1 (i.e. the higher is the adjustment speed), the sooner the investment is rewarded.

To conclude this section, it is important to stress that this paper is not aimed at criticizing the original Mincer's model. And the reason is simple. As individual-level longitudinal data were not available when Mincer put forward his theory, there was no reason (and possibility) to allow for adjustment dynamics at that time. More generally,

we do not aim at criticizing the existing literature based on a static approach. Rather, we aim at complementing it.

4. Methods

To explore wage dynamics as those described in model (19), due to the presence of initial potential earnings $\ln E_{0i}$, we need to estimate a dynamic panel-data model with individual unobserved heterogeneity of the following type:

$$(20) \qquad Y_{it} = \upsilon_i + \upsilon_1 Y_{it-1} + \upsilon_2 X_{it} + \upsilon_3 S_i + e_{it}$$

where υ_i is correlated with S_i and X_{it} by assumption, while e_{it} is a pure wage shock (white noise), orthogonal to the explanatory variables.

Since $Y_{it-1} = \upsilon_i + \upsilon_1 Y_{it-2} + \upsilon_2 X_{it-1} + \upsilon_3 S_i + e_{it-1}$, then υ_i is also correlated with Y_{it-1}. Therefore, as the OLS estimator assumes the orthogonality of the composite error term $\upsilon_i + e_{it}$ with the explanatory variables, and this condition is violated, the OLS estimates of model (20) are inconsistent.

A transformation that eliminates υ_i is the first-difference transformation:

$$(21) \qquad Y_{it} - Y_{it-1} = \upsilon_1 (Y_{it-1} - Y_{it-2}) + \upsilon_2 (X_{it} - X_{it-1}) + (e_{it} - e_{it-1})$$

Based on model (21), Anderson and Hsiao (1981) propose to use Y_{it-2} as instrument for $Y_{it-1} - Y_{it-2}$. This instrument is mathematically linked to (hence correlated with) $Y_{it-1} - Y_{it-2}$ and uncorrelated with $e_{it} - e_{it-1}$, as long as e_{it} is not serially correlated.

Arellano and Bond (1991) provide a useful test for autocorrelation in the errors. The test has a null hypothesis of 'no autocorrelation' and is applied to the differenced residuals $\Delta e_{it} = \vartheta_1 \Delta e_{it-1} + \vartheta_2 \Delta e_{it-2} + \omega_{it}$. The test of $\vartheta_1 = 0$ (ABAR1) should reject the null hypothesis as Δe_{it-1} is mathematically linked to Δe_{it} through e_{it-1}. Instead, the test of $\vartheta_2 = 0$ (ABAR2) should not reject the null. That is, we should have $\vartheta_2 = 0$ otherwise the residuals in levels would be serially correlated of order one. This would make Y_{it-2} an invalid instrument since it would be correlated with Δe_{it}. In this case, one may test Y_{it-3} and so on.

The procedure suggested by Anderson and Hsiao (1981) provides consistent but not efficient estimates because it does not exploit all the available moment conditions. Arellano and Bond (1991) provide a more efficient GMM procedure that uses *all* the orthogonality conditions between the lagged values of both Y_{it} and X_{it} and the first differences of e_{it}. Their estimator is usually called the Difference GMM estimator (GMM-DIF).

A problem with the estimator of both Arellano and Bond (1991) and Anderson and Hsiao (1981) is that time-invariant variables are eliminated by the first-difference transformation. Since the major issue in this paper is to estimate the wage return to schooling, and schooling is time-invariant, we use the estimator proposed by Blundell and Bond (1998) who (building on Arellano and Bover, 1995) suggest to instrument the variables in levels in model (20) with their lagged first differences. The estimator of Blundell and Bond is usually called the System GMM estimator (GMM-SYS) because both (20) and (21) are used as a system.

The validity of the GMM-SYS additional moment conditions depends on the validity of initial-condition restrictions which, as argued by Blundell and Bond (1998, 2000) and Bond (2002), hold under (sufficient but not necessary) assumptions of mean

stationarity of the Y and X series. In our specific case, since only lagged wage differences are used as instruments for model (20), the key condition to identify the schooling parameter is the mean stationarity of the stochastic process that generates the logarithm of the hourly wage for each individual i at each time t.

We simply test this hypothesis by estimating an AR1 process (with a constant) for the wage logarithm using the OLS estimator. In presence of individual unobserved heterogeneity, the OLS estimator is biased upward and therefore provides an *upper-bound* estimate of the true autoregressive coefficient. In particular, we find that the OLS estimate of the autoregressive coefficient for Belgium is 0.824 (significant at 1% level). This value is well below the critical value of 1, thus providing evidence that the logarithm of the hourly wage is mean stationary in this country.

Summing up, in this paper, model (19) is estimated in both levels and difference using equations (20) and (21) as a system. The explanatory variables, i.e. past wages, schooling years, experience and experience squared, are all considered endogenous because they are all correlated with individual unobserved heterogeneity. The instruments for the equation in levels are the lagged differences in the logarithm of the hourly wage ($\Delta \ln w_{it-1}, \Delta \ln w_{it-2}$, …). The instruments for the difference equation are the lagged levels of all the time-varying explanatory variables ($\ln w_{it-2}$, $\ln z_{it-2}$, $\ln z_{it-2}^2$, $\ln w_{it-3}$, …).

The null hypothesis of 'the instruments as a group are exogenous' is tested using the Hansen J test. As the GMM-SYS method can generate a very high number of instruments, the evidence can suffer a problem of instruments proliferation, meaning that the endogenous variables can be over-fitted, and the power of the Hansen test to detect instruments joint-validity can be weakened. Hansen test p-values equal to 1, or very close to 1, should be seen as a warning (Roodman, 2006). Yet, as shall be seen further below, the latter does not seem to be an issue in this paper.

5. Data and estimates

As anticipated in Section 1, the empirical application proposed in this paper is based on data on male workers, aged between 18 and 65, for Belgium. The data are extracted from the European Community Household Panel (ECHP) and cover the period of 1994-2001.

Table 1 contains a description of the sample statistics. We restricted the analysis to males in order to minimize the classical sample-selection problems that would arise with females. An extension of this paper's analysis to the case of females with the introduction of a participation equation would be an interesting topic for future research. Thus, one limitation of this study is that selection is not considered.

To be consistent with the standard Mincerian model where the representative agent first stops schooling and then starts working, we selected a sample of individuals who started working after leaving school. Then, we defined the human-capital variables as usual with ECHP data.

Specifically, we associated a given number of schooling years to each completed education level[5], making sure that the schooling variable was time-invariant by

---

[5] As usual, the education system in Belgium has three levels: primary, secondary and tertiary. Regularly completing the primary level means six years of schooling. Regularly completing the secondary level means six years of schooling as well, while the tertiary level implies at least three years of schooling. In the ECHP dataset, there are six education levels: 1) higher university degree or post graduate 2) initial university degree or equivalent 3) tertiary-level education other than university degree 4) upper secondary 5) lower secondary 6) less than lower secondary. In this paper, we used the following conversion rule:

Education level	1	2 or 3	4	5	6
Schooling years	18	15	12	9	6

dropping individuals with time-varying schooling years. Then, we defined potential experience as age–schooling years–6.

As we control for individual unobserved heterogeneity in our empirical model, we implicitly take into account not only that individuals have different abilities but also that *the explanatory variables can be measured with error*. Hence, we believe that the way we measure the human-capital variables in the context of this paper is not problematic.

The natural logarithm of the individual gross hourly wage is also measured as usual. From the gross monthly wage, we obtained the weekly wage. Dividing the latter by the number of weekly hours of work, we obtained the hourly wage.

Table 2 presents estimates of model (19) based on both OLS and GMM techniques. Our preferred estimates are the GMM-SYS estimates, accounting for endogeneity, individual heterogeneity and time effects. Specifically, as referred in Section 4, these estimates are obtained using the estimator of Blundell and Bond (1998). In our preferred estimates, the coefficient $\upsilon_1 = 1 - \rho$ is statistically different from zero and estimated at 0.335. This implies that the speed of adjustment ρ is statistically different from one and estimated at 0.665. In addition, the standard Mincerian covariates, related to the individual human capital, are generally found to be significant. Note that all the standard specification tests are passed.

In sum, our preferred estimates in Table 2 provide evidence that model (19) fits the data well and that the Mincer assumption (10) or $\rho = 1$ is rejected.

As expected, the OLS estimator over-estimates the autoregressive coefficient[6] while the GMM-SYS estimates without year effects are not reliable because the model without time effects does not pass both the Hansen J test and the Arellano-Bond 2nd order autocorrelation test.

6. Computation of returns to schooling and sensitivity analysis

Using model (19), it can be easily shown that 'the return to schooling in terms of observed earnings' is given by the following expression:

$$(22) \quad \beta(z) = \frac{\partial \ln w_{it}}{\partial s_i} = \upsilon_2(1 + \upsilon_1 + \upsilon_1^2 + ... + \upsilon_1^Z) = \rho\beta\left[1 + (1-\rho) + (1-\rho)^2 + ... + (1-\rho)^Z\right]$$

and is, in general, dependent of labor-market experience z.

The return in expression (22) is, in general, lower than the return in expression (14), although the former converges to the latter as z increases. Indeed, for a value of $\rho \in (0,1)$, the following expression holds:

$$(23) \quad \beta(\infty) = \lim_{z \to \infty} \beta(z) = \frac{\upsilon_2}{1 - \upsilon_1} = \frac{\rho\beta}{1 - (1-\rho)} = \beta \,.$$

Therefore, the adjustment model (19) is able to provide a measure of β comparable with expression (14). We label $\beta(\infty) = \dfrac{\upsilon_2}{1 - \upsilon_1}$ as 'the dynamic return to schooling in terms of net potential earnings' to distinguish it from the 'the static return to schooling in terms of net potential earnings' defined in Section 2.

Expression (23) helps to show that the interpretation of β in terms of *net potential* rather than *observed* earnings, made in Section 2, is appropriate because nobody can live and work forever. To the extent of T being a finite number, the return

---

[6] Although not reported in Table 2, as one would expect, the FE estimator under-estimates the autoregressive coefficient.

to schooling in terms of observed earnings $\beta(z)$ can never be equal to β, but in the very special case of $\rho = 1$ (which is rejected in our empirical application).

As a matter of example, we use the adjustment model (19) to compute returns to schooling in terms of both net potential and observed earnings, using our preferred estimates in Table 2 (GMM-SYS, controlling for year effects).

Using expression (23), one can easily calculate that the return to schooling in terms of potential earnings $\beta(\infty)$, i.e. the equivalent of the static β return in the standard Mincer model[7], is equal to 0.093. For comparison, Figure 1 also reports the standard schooling coefficient of the static Mincer equation, i.e. expression (14), estimated at 0.110 in column 6 of Table 3, our preferred estimate.

In addition, we can use expression (22) to calculate the return to schooling in terms of observed earnings over the working life $\beta(z)$. As shown in Figure 1 (the horizontal axis measures potential labor-market experience z), the standard static Mincerian model would not capture the fact that the return to schooling is increasing over time at the beginning of the working life and that the observed return to schooling at labor-market entry $\beta(0)$, estimated at 0.062, is well below the potential one $\beta(\infty)$, estimated at 0.093.

The reminder of this section discusses the potential weaknesses of the analysis presented so far. In particular, we focus on the use of a simplified model which, we believe, is the major issue here. In Section 1, we discussed the rationale behind the use of a simple specification. In this section, we present some estimates supporting the arguments proposed in Section 1. In particular, the argument here is that using an extended model *does not allow* to recover the total return to schooling.

Specifically, Table 4 presents estimates of model (19) using the GMM-SYS estimator, controlling not only for individual unobserved heterogeneity, year effects, past wage and human-capital variables but also for other observed individual characteristics. This implies assuming that expression (9) does not hold, which may be reasonable but is not consistent with the original aim of this paper.

For identification purposes, we limit our control set to variables that exhibit time variation over the sample period. This allows us to treat these variables as endogenous and to instrument them using their lagged levels or lagged differences. This point is important because all the additional explanatory variables considered in this section are actually endogenous as they are choice variables that depend on individual unobserved characteristics.

A more elegant treatment of this type of endogeneity in a dynamic panel-data framework would require the specification and the estimation of several first-stage discrete-choice models (for instance, modeling the decision to work in the private rather than in the public sector) to produce fitted probabilities which can be used as instruments (see Wooldridge 2002, p. 625) in a GMM-SYS setting (see also Table 2 in Andini, 2012). However, this method would complicate a lot in presence of multiple-choice models, such as occupational-choice models[8], and the results would strongly depend on how the first-stage models are specified.

The extension of the control set implies a substantial loss of observations (from 4787 to 1581), after excluding not-applicable or missing values (categories -8 and -9 in the ECHP dataset).

---

[7] This does not mean that the two models, the dynamic one and the static one, must give the same estimates. The argument here is that the dynamic model allows us to obtain a better estimate of the coefficient because the adjustment process is taken into account.

[8] The choice of the occupation is not just between 0 and 1 but among 0, 1, 2 or 3, etc…

In particular, the control set includes information on occupations[9], job status (whether the individual is supervisor or not), marital status (whether the individual is married or not), health (whether the individual has chronic health problems or not), sector of production (whether the individual works in agriculture or not), migration status (whether the individual is immigrant or not), and finally sector of activity (whether the individual works in the private sector or not).

Columns from 1 to 8 gradually extend the model using a sequence of additional controls. The first model, in column 1, is model (19) estimated with the restricted sample (1581 observations) and hence with no additional controls (besides year effects). The last model, in column 8, includes the whole control set.

As one would reasonably expect, the results are consistent with the predictions of Pereira and Martins (2004). Since all the control variables are choice variables that are somehow dependent on education, the insertion of these variables into a human-capital regression model implies that a share of the impact of education on wages is captured by the coefficients of these education-dependent covariates.

Typically, when the correlation between a schooling-dependent covariate and schooling is positive (negative), controlling for this schooling-dependent covariate lowers (increases) the coefficient of schooling. Yet, when there are many schooling-dependent covariates in the control set, it is difficult to predict whether the schooling coefficient increases or decreases. In our specific case, described in Table 4, it is found that the coefficient of schooling lowers. The issue is even clearer if one compares the estimate of the return to schooling in terms of potential earnings $\beta(\infty)$ based on column 1 (0.088) with the one based on column 8 (0.058).

In addition, the extension of the control set also affects the coefficients of the potential-experience variables, which become less statistically significant, and the coefficient of the past wage, which lowers. This is again consistent with the predictions of Pereira and Martins (2004) as past wage and experience are education-dependent covariates themselves.

Finally, the extension of the control set does not notably improve the explanatory power of the regression model in Belgium. In column 8, the only variables that are statistically significant are the indicator variables for the occupation as a clerk, the role of supervisor, and the private sector (the latter at 10% level), suggesting that individuals who work in the private sector, are clerks and supervisors earn on average more than their colleagues with the same observed and unobserved characteristics who work in the public sector, have a different occupation or are not supervisors.

From an empirical point of view, the latter suggests that individual wages are well explained by model (19) even if the dataset does not allow to control for a large set of covariates. We may interpret this result as another major finding in the paper. Yet, for the purpose of this paper, we prefer to focus, again, on the finding that the hypothesis of $\rho = 1$ is rejected, implying the need of using dynamic-Mincer-equation estimates for the calculation of the return to schooling in terms of both potential and observed earnings.

Since the procedure has been already described at the beginning of this section, we do not repeat it here. Nevertheless, it is important to stress that using column 8's estimates would not allow us to recover *total* returns to schooling as the estimate of υ_2 is 'biased' by the presence of schooling-dependent covariates in the control set.

---

[9] The occupation categories are nine: 1) legislators, senior officials and managers; 2) professionals; 3) technicians and associate professionals; 4) clerks; 5) service workers and shop and market sales workers; 6) skilled agricultural and fishery workers; 7) craft and related trades workers; 8) plant and machine operators and assemblers; 9) elementary occupations.

7. Discussion

To the best of our knowledge, there are two alternative approaches to modelling income dynamics in the literature (Guvenen, 2009). Both these approaches rely on the following specification of the income model:

$$(24) \quad \ln w_{it} = \varphi_i + \upsilon_2 X_{it} + u_{it}$$

$$(25) \quad u_{it} = \upsilon_1 u_{it-1} + e_{it}$$

where u_{it} is an autoregressive income shock, φ_i captures individual unobserved heterogeneity and X_{it} is a set of time-varying regressors.

These two approaches differ for how they perform the estimation of the parameters. The first, based on the hypothesis of no unobserved individual heterogeneity ($\varphi_i = \varphi \; \forall i$), finds that income shocks are highly persistent and is called the Restricted Income Profiles (RIP) hypothesis. The second, allowing for the presence of heterogeneity, finds modest persistence of income shocks and is called the Heterogeneous Income Profiles (HIP) hypothesis.

In the RIP specification, individuals are subject to extremely persistent - near random walks - shocks while facing similar life-cycle income profiles (conditional of the observed characteristics). In the HIP specification, individuals are subject to shocks with modest persistence, while facing life-cycle profiles that are individual-specific.

Guvenen (2009) found that disregarding individual unobserved heterogeneity, when in fact is present (as in the Mincer model), implies an overestimation of the persistence estimates. In addition, heterogeneity is estimated to be substantial.

It is easy to show that expressions (24) and (25) lead to the following dynamic specification, usually called 'common-factor restricted form':

$$(26) \quad \ln w_{it} = \upsilon_i + \upsilon_1 \ln w_{it-1} + \upsilon_2 X_{it} + \upsilon_3 X_{it-1} + e_{it}$$

where $\upsilon_i = (1 - \upsilon_1)\varphi_i$ and $\upsilon_3 = -\upsilon_1\upsilon_2$ is the common-factor restriction.

The main improvement of the model proposed in this paper with respect to the literature on income profiles is that education is explicitly controlled for. For example, Guvenen (2009)'s analysis just uses subsamples of individuals by education level but education (or schooling) never enters the regression model as an additional explanatory variable. Controlling for schooling, we still find evidence in favour of the HIP hypothesis. In addition, model (26) is not suitable to be applied to Mincerian equations because of the presence of both X_{it} and X_{it-1}, which implies multicollinearity among current and lagged experience variables. Finally, while model (19) is the result of combining the Mincerian theory of potential earnings with a simple and flexible (adjustment) assumption, model (26) is the result of two assumptions, (24) and (25), which can be theoretically justified but still imply a common-factor restriction uneasy to be explained in a Mincerian context.

8. Conclusions

Consistently with the original Mincer's model, the adjustment model presented in this paper suggests that the potential return to schooling and the observed return coincide in the long-run equilibrium because the latter converges to the former as time increases. However, unlike the static Mincer equation, the model presented here allows to characterize the adjustment of the observed return to schooling towards its long-run potential equilibrium level. Further, the model is able to provide a measure of the

potential return, alternative to the standard Mincerian beta. Finally, the model highlights that there may be a difference between the observed return and its potential level and that the size of the difference depends on the magnitude of the adjustment speed. In particular, the empirical data for Belgium suggest that the observed return is substantially lower than its potential level at the beginning of the working life.

In sum, under the assumption that the Mincerian theory of the individual human-capital productivity (or potential wage) holds, this paper shows that the return to schooling in terms of both observed and potential earnings can be better estimated by allowing a dynamic wage adjustment process to take place rather than imposing an equality between observed and potential earnings at any point in time. An interesting implication of this approach is its consistency with the argument, proposed by Heckman et al. (2003 and 2005), that the observed return to schooling may be not independent of labor-market experience.

Overall, the empirical evidence supports previous results by Andini (2007, 2009, 2010 and 2012) in favor of dynamic Mincerian specifications but improves 'the state of the art' by estimating the average total (potential and observed) wage return to schooling over the individual life-cycle in a dynamic Mincerian model with individual unobserved heterogeneity. It shows that disregarding individual fixed effects (using OLS) implies an over-estimation of the autoregressive coefficient and, therefore, an under-estimation of the adjustment speed. The latter, in turn, biases the estimation of schooling returns, in terms of both observed and potential earnings.
Summing up, in this paper, we make the following contributions:

a) We estimate a dynamic Mincer equation controlling for both observed and unobserved individual heterogeneity
b) We find that observed earnings do not adjust to human-capital productivity as rapidly as assumed by Mincer (1974)
c) We compute the total return to schooling in terms of both potential and observed earnings using a)'s estimates
d) We find that the observed return is lower that its potential level at the beginning of the working life but converges to the latter as time goes by

Contributions a) and b) are in line with earlier findings by Andini (2012). Contributions c) and d) represent the real added-value of this paper. To the best of our knowledge, contribution c) is novel. Instead, regarding contribution d), it is worth stressing that, while there are alternative explanations in the literature for an 'under-return' at early stages of career, the dynamic panel-data evidence that the 'under-return' tends to disappear as time goes by is novel. This result is important because, if the estimated adjustment speed would have been zero, then individual wages would have evolved as random walks, and the investment in education would have been worthless. Instead, our evidence suggests that schooling is worth in the long run.

Of course, our results focus on the average and the issue may be more complicated if one looks at the dynamics along the conditional wage distribution. Recent advances in quantile regression for dynamic panel data (Galvão, 2011) may allow for further research in this field.

References

Anderson T.W., and Hsiao C. (1981) Estimation of Dynamic Models with Error Components, Journal of the American Statistical Association, 76(375): 598-606.

Andini C. (2007) Returns to Education and Wage Equations: A Dynamic Approach, Applied Economics Letters, 14(8): 577-579.

Andini C. (2009) Wage Bargaining and the (Dynamic) Mincer Equation, Economics Bulletin, 29(3): 1846-1853.

Andini C. (2010) A Dynamic Mincer Equation with an Application to Portuguese Data, Applied Economics, 42(16): 2091-2098.

Andini C. (2012) How Well Does a Dynamic Mincer Equation Fit NLSY Data? Evidence Based on a Simple Wage-Barganing Model, Empirical Economics, doi 10.1007/s00181-012-0581-5.

Arellano M., and Bond S.R. (1991) Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations, Review of Economic Studies, 58(2): 277–297

Arellano M., and Bover O. (1995) Another Look at the Instrumental Variable Estimation of Error-Components Models, Journal of Econometrics, 68(1): 29-52.

Belzil C. (2007) Testing the Specification of the Mincer Equation, IZA Discussion Papers, nº 2650, Institute for the Study of Labor.

Boeri T., and van Ours J. (2008) The Economics of Imperfect Markets, Princeton NJ: Princeton University Press.

Bond S.R. (2002) Dynamic Panel Data Models: A Guide to Micro Data Methods and Practice, Portuguese Economic Journal, 1(2), 141-162

Blundell R.W., and Bond S.R. (1998) Initial Conditions and Moment Restrictions in Dynamic Panel Data Models, Journal of Econometrics, 87(1): 115–143

Blundell R.W., and Bond S.R. (2000) GMM Estimation with Persistent Panel Data: an Application to Production Functions, Econometric Reviews, 19(3), 321-340.

Buchinsky M. (1994) Changes in the U.S. Wage Structure 1963-1987: Application of Quantile Regression, Econometrica, 62(2): 405-458.

Card D. (1999) The Causal Effect of Education on Earnings, in: Ashenfelter O., Card D. (Eds.) Handbook of Labor Economics, New York, North-Holland.

Flannery M.J., and Rangan K.P. (2006) Partial Adjustment toward Target Capital Structures, Journal of Financial Economics, 79(3): 469-506.

Galvão A.F. (2011) Quantile Regression for Dynamic Panel Data with Fixed Effects. Journal of Econometrics, 164(1): 142-157.

Guvenen F. (2009) An Empirical Investigation of Labor Income Processes, <u>Review of Economic Dynamics</u>, 12(1): 58-79.

Heckman J.J, Lochner L.J., and Todd P.E. (2003) Fifty Years of Mincer Earnings Regressions, <u>NBER Working Papers</u>, nº 9732, National Bureau of Economic Research.

Heckman J.J., Lochner L.J., and Todd P.E. (2005) Earnings Functions, Rates of Return, and Treatment Effects: the Mincer Equation and Beyond, <u>NBER Working Papers</u>, nº 11544, National Bureau of Economic Research.

Lemieux T. (2006) <u>The 'Mincer Equation' Thirty Years after Schooling, Experience, and Earnings</u>, in: Grossbard S. (ed.), Jacob Mincer: A Pioneer of Modern Labor Economics, New York: Springer.

Martins P.S., and Pereira P.T. (2004) Does Education Reduce Wage Inequality? Quantile Regression Evidence from 16 Countries, <u>Labour Economics</u>, 11(3): 355-371.

Mincer J. (1974) <u>Schooling, Experience and Earnings</u>, Cambridge MA: National Bureau of Economic Research.

Murphy K.M., and Welch F. (1990) Empirical Age-Earnings Profiles, <u>Journal of Labor Economics</u>, 8(2): 202-229.

Pereira P.T., and Martins P.S. (2004) Returns to Education and Wage Equations, <u>Applied Economics</u>, 36(6): 525-531.

Polachek S.W. (2007) Earnings Over the Lifecycle: The Mincer Earnings Function and Its Applications, <u>IZA Discussion Papers</u>, nº 3181, Institute for the Study of Labor.

Roodman D. (2006), How to do xtabond2: An Introduction to "Difference" and "System" GMM in Stata, <u>CGD Working Papers</u>, nº 103, Centre for Global Development.

Trostel P.A. (2005) Nonlinearity in the Return to Education, <u>Journal of Applied Economics</u>, 8(1): 191-202.

Wooldridge J. M. (2002) <u>Econometric Analysis of Cross Section and Panel Data</u>, Cambridge MA: MIT Press.

Table 1. Sample statistics, Belgium 1994-2001 (ECHP)

	Obs.	Mean	Std. Dev.	Min	Max
Log. of gross hourly wage	6873	6.164	0.433	2.815	8.697
Schooling years	6873	13.858	3.240	6	18
Potential labor-market experience	6873	19.521	10.362	0	53

Table 2. Adjustment model

Dependent variable: Logarithm of gross hourly wage

OLS

Constant	1.223 (0.000)
Logarithm of gross hourly wage (-1)	0.757 (0.000)
Schooling years	0.016 (0.000)
Potential labor-market experience	0.005 (0.001)
Potential labor-market experience squared	-0.000 (0.168)

OLS, controlling for year effects

Constant	1.252 (0.000)
Logarithm of gross hourly wage (-1)	0.754 (0.000)
Schooling years	0.016 (0.000)
Potential labor-market experience	0.006 (0.000)
Potential labor-market experience squared	-0.000 (0.094)

GMM-SYS

Constant	2.102 (0.000)
Logarithm of gross hourly wage (-1)	0.443 (0.000)
Schooling years	0.073 (0.000)
Potential labor-market experience	0.022 (0.000)
Potential labor-market experience squared	-0.000 (0.116)
ABAR1 test (p-value)	(0.000)
ABAR2 test (p-value)	(0.065)
Hansen J test (p-value) – all instruments	(0.030)
Hansen J test (p-value) – instruments for eq. in levels	(0.943)
Obs.	4787

GMM-SYS, controlling for year effects

Constant	2.901 (0.000)
Logarithm of gross hourly wage (-1)	0.335 (0.000)
Schooling years	0.062 (0.000)
Potential labor-market experience	0.032 (0.000)
Potential labor-market experience squared	-0.000 (0.000)
ABAR1 test (p-value)	(0.000)
ABAR2 test (p-value)	(0.121)
Hansen J test (p-value) – all instruments	(0.256)
Hansen J test (p-value) – instruments for eq. in levels	(0.877)
Obs.	4787

P-values of estimated coefficients, in parentheses, are based on White-corrected standard errors for OLS and on Windmeijer-corrected standard errors for GMM-SYS.

Table 3. Static returns to schooling in terms of net potential earnings

	(1) OLS	(2) OLS	(3) RE	(4) RE	(5) GMM-SYS	(6) GMM-SYS
	0.067 (0.000)	0.066 (0.000)	0.055 (0.000)	0.050 (0.000)	0.163 (0.000)	0.110 (0.000)
Control for individual effects	no	no	yes	yes	yes	yes
Control for year effects	no	yes	no	yes	no	yes
Control for endogeneity	no	no	no	no	yes	Yes

All the regressions control include constant term, experience and experience squared.
P-values of estimated coefficients, in parentheses, are based on White-corrected standard errors for OLS and on Windmeijer-corrected standard errors for GMM-SYS.

Table 4. Adjustment model with additional controls

Dependent variable: Log. of gross hourly wage	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Constant	2.660	2.879	3.280	3.229	3.222	3.287	3.239	3.422
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Log. of gross hourly wage (-1)	0.415	0.369	0.335	0.337	0.338	0.324	0.329	0.303
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Schooling years	0.052	0.046	0.040	0.040	0.040	0.040	0.043	0.041
	(0.000)	(0.009)	(0.018)	(0.020)	(0.020)	(0.022)	(0.015)	(0.027)
Potential experience	0.023	0.032	0.023	0.023	0.023	0.025	0.025	0.023
	(0.005)	(0.000)	(0.031)	(0.028)	(0.031)	(0.027)	(0.030)	(0.053)
Potential experience squared	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000
	(0.108)	(0.013)	(0.129)	(0.120)	(0.132)	(0.113)	(0.139)	(0.207)
Occupation 1		0.163	-0.066	-0.075	-0.074	-0.032	-0.054	-0.044
		(0.411)	(0.745)	(0.727)	(0.734)	(0.883)	(0.795)	(0.833)
Occupation 2		0.143	0.147	0.142	0.142	0.172	0.149	0.190
		(0.397)	(0.362)	(0.402)	(0.399)	(0.318)	(0.394)	(0.268)
Occupation 3		0.141	0.024	0.019	0.019	0.020	0.013	0.078
		(0.394)	(0.887)	(0.917)	(0.914)	(0.911)	(0.941)	(0.669)
Occupation 4		0.383	0.322	0.317	0.317	0.369	0.361	0.418
		(0.006)	(0.018)	(0.026)	(0.026)	(0.013)	(0.012)	(0.004)
Occupation 5		-0.157	-0.078	-0.082	-0.081	-0.077	-0.084	-0.038
		(0.281)	(0.568)	(0.554)	(0.569)	(0.598)	(0.558)	(0.803)
Occupation 6		-1.579	-2.521	-2.513	-2.503	-2.979	-3.002	-4.663
		(0.555)	(0.405)	(0.406)	(0.404)	(0.358)	(0.354)	(0.262)
Occupation 7		0.109	-0.028	-0.031	-0.030	-0.016	-0.024	-0.089
		(0.468)	(0.839)	(0.822)	(0.835)	(0.912)	(0.866)	(0.553)
Occupation 8		0.0851	-0.051	-0.064	-0.062	-0.045	-0.042	-0.123
		(0.547)	(0.719)	(0.712)	(0.718)	(0.797)	(0.810)	(0.519)
Job status (1 if supervisor)			0.304	0.300	0.299	0.298	0.296	0.203
			(0.002)	(0.002)	(0.003)	(0.004)	(0.004)	(0.041)
Marital status (1 if married)				0.048	0.048	0.035	0.024	-0.044
				(0.803)	(0.803)	(0.855)	(0.902)	(0.820)
Chronic health problem (1 if yes)					0.009	0.008	0.000	-0.002
					(0.963)	(0.967)	(0.997)	(0.990)
Sector of production (1 if agriculture)						0.324	0.319	0.256
						(0.401)	(0.407)	(0.525)
Migration status (1 if immigrant)							-0.042	-0.148
							(0.746)	(0.359)
Sector of activity (1 if private sector)								0.134
								(0.069)
ABAR1 test (p-value)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
ABAR2 test (p-value)	(0.235)	(0.437)	(0.702)	(0.717)	(0.717)	(0.730)	(0.702)	(0.740)
Hansen J test (p-value)								
- all instruments	(0.428)	(0.435)	(0.578)	(0.518)	(0.512)	(0.475)	(0.449)	(0.531)
- instruments for eq. in levels	(0.849)	(0.605)	(0.491)	(0.568)	(0.601)	(0.611)	(0.562)	(0.555)
Observations	1581	1581	1581	1581	1581	1581	1581	1581

All the regression models control for year effects. Occupation 9 is the excluded category.
P-values of estimated coefficients, in parentheses, are based on Windmeijer-corrected standard errors.

Figure 1. Returns to schooling in terms of observed earnings

Belgium
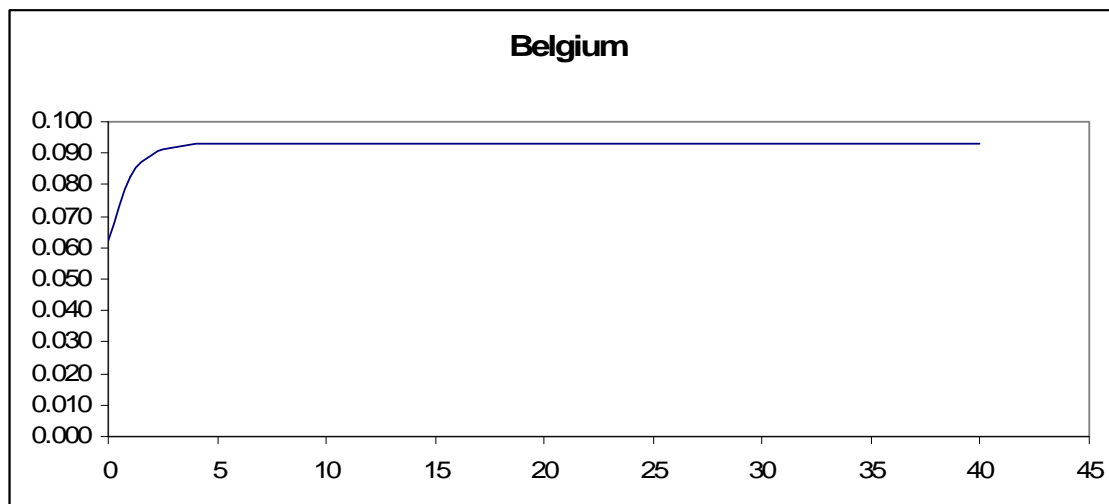
$\beta(0) = 0.062$, $\beta(\infty) = 0.093$, and $\beta = 0.110$

The horizontal axis measures potential labor-market experience z