Working papers series

WP ECON 15.15

Changes in the Return to Skills and the Variance of Unobserved Ability

Guido Matías Cortés (University of Manchester)

Manuel Hidalgo-Pérez (Universidad Pablo de Olavide)

Keywords: Return to skills; Within-group wage inequality

JEL Classification: J31; J24; E24

Department of Economics

Changes in the Return to Skills and the Variance of Unobserved Ability

Guido Matias Cortes* Manuel A. Hidalgo-Pérez†

University of Manchester and RCEA Universidad Pablo de Olavide

This Draft: December 14, 2015

Abstract

Changes in within-group wage inequality are often interpreted as reflecting changes in returns to unobservable skills. This interpretation relies on the highly restrictive assumption that the variance of skills within groups remains constant over time. We propose and implement a new identification strategy which relaxes this assumption using longitudinal data. Results based on matched CPS data for 1982-2012 show strong evidence of increases in the dispersion of unobserved skills, particularly among college graduates. Contrary to the conclusion drawn when constant within-group skill variance is assumed, our results suggest that the return to skills decreased during the 1980s and early 1990s.

JEL Codes: J31; J24; E24.

Keywords: Return to skills; Within-group wage inequality.

*Department of Economics, University of Manchester, Room 2.005 Arthur Lewis Building, Manchester, UK M13 9PL. E-mail: matias.cortes@manchester.ac.uk. We thank seminar and conference participants at ISER, Manchester, SOLE/EALE, RES and WPEG for valuable comments and suggestions.

†Department of Economics, Pablo de Olavide University, Room 3.14 Building 3, Seville, Ctra de Utrera, s/n. E-mail: mhidper@upo.es

1 Introduction

Increases in wage inequality over recent decades in many developed countries have sparked an active research agenda analyzing its sources and the driving forces behind its changes over time. An important finding in this literature is that a large proportion of the increase in wage inequality can be attributed to the dispersion of wages *within* groups of workers defined by their education and age, rather than *across* these groups (Katz and Murphy (1992), Acemoglu (2002)). This increase in within-group inequality is generally interpreted as reflecting changes in the return to unobservable skills within groups (e.g. Juhn, Murphy, and Pierce, 1993), providing support for the influential theory of Skill-Biased Technical Change (SBTC), which suggests that technology is an important driver of inequality. According to the SBTC theory, new technologies increase the demand for skilled workers (e.g. Bound and Johnson (1992), Autor, Katz, and Krueger (1998)), implying an increase in the return to observable and unobservable skills over time.

In spite of its pervasiveness, the interpretation of changes in within-group wage inequality as reflecting changes in the return to skills is not necessarily accurate. A simple specification of an individual's residual wage suggests that it is the product of the return to skill and the quantity of unobservable skills possessed by the individual. Changes in within-group wage inequality may therefore reflect not only changes in the return to skills, but also changes in the distribution of unobservable skills within the group. Implicitly or explicitly, many papers assume that the second component is constant over time, and can therefore interpret changes in within-group inequality as being driven by returns to skills only. In this paper, we relax the assumption of constant within-group skill variance and estimate the relative importance of changes in the return to skills and in the distribution of skills in driving the observed changes in within-group wage inequality. We show that relaxing this assumption has dramatic implications for estimates of the evolution of the return to skills.

Previous literature has recognized the restrictiveness of assuming that the variance of unobserved skills is constant within groups (see Chay and Lee (2000) p.16, or Lemieux (2006) footnote 11). The concerns with this assumption are twofold. From a long-term perspective, the variance in the distribution of skills within education groups may change due to differences across cohorts. Consider the expansion of tertiary education which has occurred over recent decades. The traditional assumption implies that the dispersion of unobservable skills among, for example, young university-educated workers, is the same in 2010 as it was in 1980. As the set of young college educated workers has grown, this group has become more diverse along many observable dimensions, which likely implies that they are drawn from a wider distribution of unobserved abilities and therefore the variance of unobserved skills for this group would be increasing over time. This may be due to an increase in the dispersion of intrinsic ability among college-goers or because of a wider dispersion in the quality of education (Hendricks and Schoellman (2014), Guvenen and Kuruscu (2010)).[1] Under the traditional assumption, any change in the variance of wages for this group would be attributed exclusively to changes in the return to skills, rather than to changes in the variance of skills.[2] The distribution of unobservable skills may also change across cohorts due to factors other than changes in the selection patterns into different education groups. For example, Card, Heining, and Kline (2013) find evidence of an increase over time in the variance of the worker fixed effects obtained from an Abowd, Kramarz, and Margolis (1999) decomposition for wages in Germany.

A second cause for concern with the traditional assumption of constant within-group skill variance is related to short-term changes in the distribution of skills within cohorts due to business cycle fluctuations. There is evidence that individuals who lose their job during a recession are not randomly selected within demographic group (see for example Blundell, Reed, and Stoker (2003)). Therefore, the variance of ability among *workers* from a given group in a recessionary year may be quite different from

[1]Technological change may potentially be a driver of the changes in the selection patterns into different educational groups. See Hidalgo-Pérez and Molinari (2014) and Cortes (2016) for models in which technological change affects the composition of skills within different occupations.

[2]Note that the conditional mean of unobserved skills may also be changing over time due to these changes in the composition of university graduates, or due to newer "vintages" of workers receiving different amounts of value added through the education process (Bowlus and Robinson (2012), Carneiro and Lee (2011)). This does not affect the identification strategy described below.

the variance during an expansionary year.

Using wage data to estimate the different components of residual wage inequality is challenging, as neither the quantity of skills nor their return can be directly measured. In this paper we use longitudinal data in order to solve this identification problem. We relax the traditional assumption of constant within-group skill variance, and replace it with the more nuanced assumption that the variance of skills remains constant only over short horizons and only when conditioning on people who remain employed across consecutive years. This allows the variance of unobserved skills to differ across cohorts, as well as differing within cohorts over the business cycle due to non-random attrition. We show how changes in the return to skills can be identified under our milder assumptions using data on changes in wage inequality for a common group of individuals observed at two different points in time.

A major advantage of our identification strategy – in addition to relaxing the assumption of constant within-group skill variance – is that it only requires two observations per individual. Thus, we are able to take advantage of the limited longitudinal dimension of the Current Population Survey, rather than having to rely on longer panel datasets which typically have much smaller sample sizes. We implement our new identification strategy using data from the CPS's Merged Outgoing Rotation Group (MORG) sample for the period 1982-2012 and estimate the relative role of price changes and changes in the variance of skills in accounting for changes in within-group residual wage variances.

Our results suggest that relaxing the assumption of constant within-group variance is crucial. The simple change in identification strategy that we propose completely changes the conclusions about the driving force behind the observed changes in within-group inequality. We find that the variance of unobserved skills within education-age groups experiences important changes over time which, when ignored, generate misleading results about the changes in the return to skills. Based on our identification strategy, we find that the return to skills falls during the 1980s and early 1990s, then recovers somewhat until around 2002, and remains fairly stable thereafter. This implies that the main driver of the observed increases in within-group inequality among college graduates is not an increase in the return to skills, but rather an increase in the dispersion of skills within this group. This increase is estimated to be quite substantial between the mid-1980s and mid-1990s.

We show that this key result is not driven by the fact that we use a selected sample (i.e. workers who remain employed across two consecutive years), but rather is solely due to the change in the identification strategy. Moreover, allowing for measurement error using external estimates from Lemieux (2006) amplifies the fall in the return to skills during the early period. We also allow for the possibility that returns to skills may differ across demographic groups. In this case, we find results that are consistent with the aggregate patterns except for two groups for which we find evidence of increases in the return to skills after the mid-1990s: young workers (aged 25-34) and college graduates.

Next, we extend our identification strategy to further relax the assumption that the variance of skills remains constant within a group of workers observed across two consecutive periods. Here we combine elements from another strand of the literature, which specifies an error components structure for residual wages and is primarily interested in identifying the variance of permanent and transitory shocks to earnings (e.g. Meghir and Pistaferri (2011), Moffitt and Gottschalk (2012), Blundell et al. (2014)), without particularly focusing on estimating the return to skills. In particular, we allow individuals to be subject to permanent shocks to skills, due for example to health shocks, or heterogeneities in the accumulation of human capital on the job. We also allow for temporary idiosyncratic shocks to earnings. While adopting the specification structure of permanent and transitory shocks, we maintain the idea of identifying changes in the return to skills through changes in conditional wage variances (as in Chay and Lee (2000)and Lemieux (2006)). Identification is achieved by considering changes over time in the differential between groups, when still conditioning on a consistent set of workers observed across two consecutive periods.

With this identification strategy, we find that the results are very sensitive to the assumption made about the nature of the permanent shocks to skills. In particular, if the variance of these shocks is assumed to be common across age groups (but different between education groups), the result suggests an increase in the return to skills in the 1980s, followed by a stronger decline during the 1990s and 2000s. The results are quite different, however, if the variance of the shocks is assumed to be common across education groups (but different between age groups). In this case, the result suggests an increase in the return to skills throughout the 1990s and early 2000s. Hence, estimates of the return to skills that account for common shocks across different groups within cohorts are very sensitive to the assumptions made about the nature of these shocks.

Our paper is not the first to challenge the assumption that the distribution of skills remains constant within groups over time. Other contributions to the literature have taken different approaches to address the identification challenge. For example, Heckman, Lochner, and Taber (1998) and Bowlus and Robinson (2012) identify a "flat spot" in workers' skill profiles: a period over the life-cycle during which workers' skill levels remain flat. Changes in the return to skills are identified by analyzing changes in the wage *levels* of workers over their flat spots (rather than wage variances). Another prominent example of a paper that challenges the assumption of constant within-group skill variance is Lochner and Shin (2014), whose paper is closely related to ours. Lochner and Shin (2014) make alternative assumptions about the nature of the shocks to skills and earnings. The main advantage of our identification strategy over the one proposed by Lochner and Shin (2014) is that we only require two observations per individual, and therefore can take advantage of the large sample size in the CPS. The identification strategy in Lochner and Shin (2014) requires more observations per individual, and therefore can only be implemented with panel datasets such as the Panel Study of Income Dynamics, which tend to have much smaller sample sizes.

The rest of the paper is organized as follows. In Section 2 we discuss our identification strategy and contrast it with the traditional strategy which assumes that the variance of unobserved skills remains constant over time within groups. Section 3 describes the dataset and discusses the details of the empirical implementation. Section 4 presents our main results, while Section 5 extends our identification strategy to account for permanent shocks to skills within cohorts over the life cycle, as well as idiosyncratic transitory shocks to earnings. Section 6 concludes.

2 Estimating the Return to Unobserved Skills

Suppose that log wages are determined by an error components model such as in Chay and Lee (2000) or Lemieux (2006):

$$w_{it} = x_{it}b_t + u_{it} \tag{1}$$

$$u_{it} = p_t e_{it} + \nu_{it} \tag{2}$$

w_{it} is the natural logarithm of the hourly wage rate for individual i at time t, x_{it} is a vector of observed skills (such as education and labor market experience) and b_t is the return (or price) of observed skills. u_{it} represents residual wages, which are composed of unobserved skills e_{it}, the return to those skills which is given by p_t, and measurement error ν_{it}, which may also be interpreted as an idiosyncratic shock to wages. We assume that the distribution of ν_{it} is independent from e_{it}.

Suppose that observable skills x_{it} can be fully characterized by individuals' education and age. For individuals with education level c and age a, given the independence assumption for ν_{it}, the within-group variance of wages is given by:

$$V_{a,c,t} = p_t^2 \sigma_{a,c,t}^2 + \sigma_{\nu,a,c,t}^2 \tag{3}$$

where $\sigma_{a,c,t}^2 \equiv Var(e_{it}|a,c,t)$ is the conditional variance of unobserved skills (conditional on education and age) and $\sigma_{\nu,a,c,t}^2 \equiv Var(\nu_{it}|a,c,t)$ is the conditional variance of the measurement error.

From equation (3) it is clear that changes over time in within-group wage inequality may be driven by three different factors:

i. Changes in p_t, the price paid to unobserved ability, due for example to changes in technology which change the demand for skills.

ii. Changes in $\sigma_{a,c,t}^2$, the dispersion of unobserved ability within groups, due for example to changes in the skills of workers selecting into particular education groups, changes in the dispersion of schooling quality, or changes in the relevance of on-the-job training.

iii. Changes in $\sigma^2_{\nu,a,c,t}$, the variance of measurement error, due for example to methodological changes in the survey, or changes in the incidence of temporary idiosyncratic shocks.

We are interested in decomposing the relative importance of each of these components. Identification is difficult due to the fact that none of these components are directly observable. For now we ignore measurement error and focus on the identification of the return to skills p_t and the within-group variance of unobserved skills $\sigma^2_{a,c,t}$. We return to the issue of measurement error in Section 5.

One common way in which previous literature has achieved identification is by assuming that the within-group variance of unobserved ability among workers in a given education-age group is constant over time, i.e. $\sigma^2_{a,c,t} = \sigma^2_{a,c}, \forall t$. Under this assumption, we have that:

$$\frac{V_{a,c,t}}{V_{a,c,t-1}} = \frac{p_t^2}{p_{t-1}^2}$$

$$\Rightarrow \ln p_t - \ln p_{t-1} = \left(\frac{1}{2}\right) \ln \left(\frac{V_{a,c,t}}{V_{a,c,t-1}}\right) \tag{4}$$

Making the normalization $p_0 = 1$, we can back out the return to unobserved skills at time t by computing:

$$\ln p_t = \left(\frac{1}{2}\right) \sum_{\tau=1}^{t} \ln \left(\frac{V_{a,c,\tau}}{V_{a,c,\tau-1}}\right) \tag{5}$$

$$= \left(\frac{1}{2}\right) \ln \left(\frac{V_{a,c,t}}{V_{a,c,0}}\right) \qquad \forall t > 0$$

This is the price series implied by the Chay and Lee (2000) and Lemieux (2006) frameworks when abstracting from measurement error.[3] Changes in

[3]More specifically, this is the price series implied by the Chay and Lee (2000) framework in what they call their "between-cohort" analysis. They also consider a "within-cohort" analysis where the identifying assumption is that the variance of unobserved skills remains constant for workers from a given cohort as they gain additional experience. This "within-cohort" assumption is closer, although not equivalent, to our identification assumption discussed in detail below.

the return to unobserved skills are identified from changes over time in within-group wage inequality.

The identification of the returns to unobserved ability based on equation (5) relies on the crucial assumption that the within-group variance of unobserved skills is constant over time. As discussed in the introduction, there are two main reasons why this very strong assumption may fail. First, from a longer term perspective, changes in the selection patterns into higher education likely imply changes over time in the distribution of unobserved skills within education-age groups. Second, selection in terms of which workers lose their jobs during recessions also implies changes over the business cycle in the distribution of unobserved skills among individuals from a particular demographic group who remain in work. Therefore, in what follows we relax this assumption in order to allow the variance of unobserved skills to change over time. In particular, we make the alternative assumption that the variance of unobserved skills remains constant only over short time frames and only when conditioning on the exact same subset of individuals across two periods.

Suppose that we have access to longitudinal data. Denote the subset of individuals whose wages are observed over two consecutive years $t-1$ and t as s^t ("stayers"). Individuals in group s^t must be employed in both periods. We make the assumption that, for this subset of individuals only, the variance of ability remains constant across periods $t-1$ and t. That is:

$$\sigma^2_{a,c,t,s^t} = \sigma^2_{a-1,c,t-1,s^t} \tag{6}$$

The overall within-group variance of unobserved skills for *all* individuals of age a and education level c, $\sigma^2_{a,c,t}$, is allowed to change freely between periods $t-1$ and t due to the exit of "non-stayers" between these two periods, and the addition of new workers to each demographic group in period t. Our assumption therefore allows for changes in the composition of skills across cohorts (for example due to changing selection patterns into higher education) by assuming that the variance of skills remains constant only within cohorts. Moreover, by restricting the assumption to stayers within a cohort only, our assumption also allows for changes in the composition of skills within a particular demographic group over the business cycle due to non-random selection out of employment. Note that this identification assumption implies that the distribution of skills for a given set of individuals does not change over the life cycle. In Section 5 we relax this assumption and allow for a deterministic life-cycle pattern in the variance of ability.

Assuming that the return to unobserved skills p_t is common for all individuals, identification is achieved by considering the changes over time in the within-group variance of wages for the subset of individuals in s^t. That is:

$$\frac{V_{a,c,t,s^t}}{V_{a-1,c,t-1,s^t}} = \frac{p_t^2}{p_{t-1}^2} \tag{7}$$

$$\Rightarrow \ln p_t - \ln p_{t-1} = \left(\frac{1}{2}\right) \ln \left(\frac{V_{a,c,t,s^t}}{V_{a-1,c,t-1,s^t}}\right) \tag{8}$$

If we make the normalization $p_0 = 1$, we can obtain an implied price series given by:

$$\ln p_t = \left(\frac{1}{2}\right) \sum_{\tau=1}^{t} \ln \left(\frac{V_{a,c,\tau,s^\tau}}{V_{a,c,\tau-1,s^\tau}}\right) \qquad \forall t > 0 \tag{9}$$

Under our identification assumption, changes in the return to unobserved skills are identified from longitudinal changes in within-group wage inequality among “stayers”. If changes in the variance of ability over time within education-age groups were unimportant, then the implied price series from equations (5) and (9) should be similar to each other.

Note that the discussion above relies on the condition that workers can be divided into education-age cells which fully characterize observable skill groups. In practice, when taking this approach to the data, the number of cells that a sample can be divided into while maintaining a large enough sample size in each cell is limited. This is particularly true in the case of our identification strategy which relies on longitudinal data. When individuals are grouped into broader skill groups, a non-negligible amount of heterogeneity in terms of observable characteristics will remain within each cell.[4] Therefore, in order to avoid confounding the effects of changes

[4] As we discuss further below, in practice we group individuals into 8 cells based on 2 education group and 4 age groups each comprising 10 years of age.

in observable within-group heterogeneity, we first estimate a log wage regression to obtain *residual* wages and implement the identification strategy described above using the within-group variances of residual (rather than real) wages.[5] In practice, this implies imposing the additional assumption that the partial correlation of age with unobserved ability (conditional on education) is constant over time; see the proof in Appendix A.

3 Empirical Implementation

We use information from the Merged Outgoing Rotation Group (MORG) sample from the monthly Current Population Survey (CPS) for the period from January 1982 until December 2012. The CPS is the main source of labor market statistics in the United States. We take advantage of the fact that the CPS is a rotating sample: households included in the survey are sampled for four consecutive months, then leave the sample for eight months before returning for another four months. Earnings information is collected when households are in their fourth and eight months in the sample (i.e. when they are in the so-called Outgoing Rotation Groups), so there is information on earnings for the same household across the same calendar month in two consecutive years. Details regarding the algorithm used to match individuals over time can be found in Nekarda (2009).

We restrict the sample to males aged 25 to 64. To obtain residual wages, we regress log real hourly earnings on a set of calendar month dummies, education dummies, age bin dummies and interactions of education dummies and a quartic in age. This regression is run separately for each calendar year.

We then categorize our observations into education-age groups, using two education categories (high school or less including those with some college education, and college graduates), and four age categories. The set of "stayers" s^t is defined as the set of individuals who are in month-in-sample 4 in period $t-1$ and month-in-sample 8 in period t, and have valid earnings data in both periods. We exclude outliers from this group, defined as individuals with residual wage changes greater than 60 log points.

[5]Details on the regression that we estimate are provided in Section 3.

Figure 1 illustrates the structure of the data for the set of workers observed in year t. Approximately half of these workers will be interviewed for the first time in year t (and re-interviewed in year $t+1$) and the other half will be in their second interview in year t. This latter group will not be re-interviewed in the following period. Some of them would have also been observed working in the previous year; these constitute the group s^t. Among those being interviewed for the first time in year t, a small number will not be re-interviewed in the following period due to attrition, and some will be re-interviewed but will not be working in year $t+1$. The remainder represent our group of stayers s^{t+1}, for whom we also have earnings data in the following year.

As mentioned above, our data is at a monthly frequency, but we observe the same individuals in the same calendar month across two consecutive years (rather than across consecutive months). Therefore, for each year t in our dataset, we pool all of the monthly observations and, as illustrated in the figure, we calculate the following within-group residual wage variances: (i) $V_{a,c,t}$, using all workers in year t, (ii) V_{a,c,t,s^t} and $V_{a-1,c,t-1,s^t}$ using workers in group s^t, and (iii) $V_{a,c,t,s^{t+1}}$ and $V_{a+1,c,t+1,s^{t+1}}$ using workers in group s^{t+1}.

One limitation of the longitudinal dimension of the matched monthly CPS data is that there are certain months when there are breaks in the CPS identifiers which make it impossible to generate matches.[6] Therefore, although for most years in our sample we have data for stayers from all 12 months within the year, there are some years for which this is not the case. Moreover, due to the change in the CPS identifiers in January 1994, we are not able to generate matches for any month in 1993. In our analysis, we assume that the price change for that particular year only is given by the change estimated for the full sample rather than the sample of stayers i.e., using equation (4) rather than equation (8).[7]

Figure 2 plots the evolution of the within-group variance of residual wages $V_{a,c,t}$ for each of our demographic groups. Within-group variances tend to increase with age and education. Over time, within-group vari-

[6]This occurs in July 1985, October 1985, January 1994, June 1995 and September 1995.

[7]As an alternative, we consider the assumption that the price change between 1993 and 1994 is equal to zero. Our results do not change in any major way under this alternative assumption.

ances are generally increasing for college graduates, while they are stable or decreasing slightly for those without a college degree.

Recall that under the traditional assumption of constant within-group variance of skills, all of the changes over time in the series plotted in Figure 2 are interpreted as reflecting changes in the return to skills. In the next section of the paper, we estimate the changes over time in the return to skills under that assumption, as well as under our more nuanced identification assumption in order to understand the extent to which the changes in within-group variance in Figure 2 can in fact be attributed to changes in the distribution of unobserved ability within each group, rather than to price changes.

4 Results

4.1 Estimated returns to skills

We use the computed variances of residual wages to estimate the return to skills based on the traditional identification strategy using equation using all workers (equation (5)), and under our proposed identification strategy which relaxes the assumption of constant within-group skill variance and focuses on stayers (equation (9)). The ratios of within-group residual wage variances are computed for each education-age group and then a weighted average across the groups is computed for each period in order to obtain a single price series.[8] The results are presented in Figure 3. The series represent price changes relative to the year 1982, when log-prices are normalized to zero.

The figure shows a clear difference between the two identification methods. By relaxing the assumption of constant within-group variance we obtain a price series that is decreasing in the 1980s and then only moderately increasing thereafter. This contrasts sharply with the consistently increasing series obtained when the assumption of constant within-group variance is imposed. Our results imply that the observed increases in within-group variance are driven by increases in the variance of unobserved ability within

[8]The weights are equal to the number of observations used to calculate the variance for each group.

groups, rather than increases in the return to skills. Figure 4 depicts the implied series for the variance of unobserved ability within each group, given our estimated returns to skills. The figure shows that the within-group variance of skills increased particularly strongly between the mid-1980s and the mid-1990s among college graduates.

Note that our identification strategy uses a deliberately selected sample (i.e. individuals who are observed twice and are working in both periods) for whom we can more reasonably assume that the variance of unobserved skills remains constant across consecutive periods. One might be concerned that the patterns of within-group wage inequality for this group may be systematically different from those of the full sample, and that this may be the source of the differences in the identified return to skills. To show that our results are not driven by our sample selection, Figure 5 plots the estimated return to skills when we apply the traditional identification strategy using our selected sample only. The results clearly show that the driving force behind our findings is not the sample selection, but rather the change in the identification strategy which allows for changes over time in the variance of unobserved skills.

4.2 Decomposition of the changes in residual wage variance

The results on the changes in skill prices and in the variance of unobserved skills can be summarized by performing a decomposition of the change in the within-group variance for each group into a price effect and a distribution effect. Specifically, the change in the within-group variance across two periods t and t' can be decomposed as follows:

$$\begin{aligned} V_{a,c,t} - V_{a,c,t'} &= p_t^2\sigma^2_{a,c,t} - p_{t'}^2\sigma^2_{a,c,t'} \\ &= \underbrace{(p_t^2 - p_{t'}^2)\sigma^2_{a,c,t}}_{\text{Price Effect}} + \underbrace{p_{t'}^2(\sigma^2_{a,c,t} - \sigma^2_{a,c,t'})}_{\text{Distribution Effect}} \end{aligned} \tag{10}$$

The price effect in equation (10) is the change in the within-group variance that would have occurred due to the change in the return to skills if the within-group distribution of skills had remained constant as in period t. The distribution effect is the portion attributable to changes in the within-group skill variance, holding prices constant at their level for period t'.

Table 1 shows the results from the decomposition for the changes in residual wage variance between the periods 1982-1984 and 2010-2012. Recall that, under the traditional assumption that the variance of skills within groups is constant, essentially all of the changes in the residual wage variances are attributed to changes in the return to skills. The decomposition results make clear how this changes dramatically when implementing our identification strategy. Our findings show that the increases in within-group inequality observed across demographic groups mask a reduction in the return to skills coupled with an increase in the dispersion of skills within groups.

4.3 Allowing for heterogeneous returns to skills across demographic groups

Our identification strategy allows us to estimate different returns to skills for different groups. Figure 6 plots the estimated return to skills separately by education group in the top panel, and by age group in the bottom panel. The aggregate patterns shown above are most closely reflected in the series for the lower education group (which is also a relatively much larger group). Among college graduates, there is evidence of increases in the return to skills over the 1990s after a stronger fall in the 1980s. This result, however, should be interpreted with caution. As Chay and Lee (2000) discuss, this group is particularly affected by changes in the top-coding procedures in the CPS that were introduced in 1989. As Figure 2 in their paper shows, the fraction of college graduates censored at the top-code increases steadily up to 1988, when usual weekly earnings were top-coded at \$999, before falling sharply in 1989, when the top-coding level was raised to \$1999 per week. For older college graduates, the top-coding rate reaches levels above 30% by 1988. The drop in the estimated return to skills before 1989 among college graduates is therefore not reliable, as it is likely driven by the increase in the fraction of top-coded workers. The data from 1989 onwards is far less affected by top-coding and therefore much more reliable, suggesting that there is a genuine increase in the return to skills among college graduates in the 1990s, in contrast to what is observed among those without a college degree.

The pattern of a decrease in the return to skills in the 1980s, followed by a period of relative stability in the 1990s and 2000s is broadly similar to what is found by Bowlus and Robinson (2012) using a different identification strategy. Their estimates also show that, after a period when the return to skills for high school graduates and college graduates move closely together, in the 1990s the return to skills among college graduates moves above that of high school graduates (see their Figure 3). This is also in line with our findings.

In terms of the differences across age groups, the bottom panel of Figure 6 shows that the estimated returns to skills are very similar among the groups of workers aged 35 and above. The patterns for the youngest group of workers is different. Here we find that after a similar decline in the returns in the 1980s, there is a fairly strong increase in the return to skills among workers aged 25-34. This could reflect heterogeneities in demand for skills among different age groups along the lines of Card and Lemieux (2001).

4.4 Measurement error: External estimate

We next consider the effects of allowing for measurement error on our estimated price series. We first use an external estimate for the variance of the measurement error. The strategy of using independent estimates of the variance of measurement error is common in the literature (see e.g. Meghir and Pistaferri (2004), Heathcote, Storesletten, and Violante (2010)). Section 5 below considers an alternative approach to account for measurement error.

Allowing for measurement error, the within-group variance of residual wages for individuals with education level c and age a is given as in Equation (3). Lemieux (2006) compares wages in the March and May/ORG CPS and finds that there is no change over time in the variance of measurement error for men in the ORG CPS over the period 1976-2003.[9] Based on

[9]Lemieux (2006) finds that the increase over time in the variance of measurement

this evidence, consider the case in which we assume that the variance of measurement error is common across groups (as in Chay and Lee (2000)) and constant over time.

Denoting the variance of measurement error as σ_ν^2, we can modify equation (7) and re-write the price ratio as:

$$\frac{V_{a,c,t,s^t} - \sigma_\nu^2}{V_{a-1,c,t-1,s^t} - \sigma_\nu^2} = \frac{p_t^2}{p_{t-1}^2} \tag{11}$$

Given an external estimate for the variance of measurement error, we can make simple adjustments to our computed within-group variances to adjust for measurement error. We consider three potential levels for the variance of the measurement error based on Lemieux (2006): a low level of 0.017 (Lemieux (2006)'s estimate for workers paid by the hour), a higher level of 0.052 (the corresponding estimate for workers not paid by the hour), and an intermediate level of 0.032 based on a weighted average of the two.[10]

The resulting estimated price series are displayed in Figure 7, along with our benchmark estimates which assume that the variance of measurement error is equal to zero. Accounting for measurement error leads to a return to skills series which falls more dramatically in the 1980s and early 1990s, and then recovers somewhat. The higher estimates for measurement error lead to larger declines in the estimated return to skills. This adjustment therefore suggests that, if anything, our benchmark price series is overestimating the return to skills and strengthens the argument in favor of the hypothesis that increases in within-group inequality are driven by factors other than an increase in the return to skills.

error for workers who are paid by the hour is offset by an increase in the fraction of men who are paid by the hour (the variance of measurement error is lower for this group relative to those who are not paid by the hour).

[10]The weigted average uses the measurement error variances in rows 1A and 1B of Table 3 in Lemieux (2006) using a weight of 0.58 for hourly workers, which is approximately equal to their average fraction in the sample (Hamermesh, 2002).

5 Extensions: Life-Cycle Patterns, Permanent and Transitory Shocks

So far our identification strategy has relied on the assumption that the variance of unobserved ability remains constant for the group of workers who remain employed over two consecutive years. In this section we relax this assumption further to allow for changes in the variance of unobserved skills among stayers due to permanent shocks to skills, which could be attributed to heterogeneities in the accumulation of human capital on the job, for example due to differences in training or on-the-job learning across occupations. We also continue to allow for measurement error, and relax the assumption that the variance of this component is constant over time.

Specifically, suppose that skills are subject to permanent idiosyncratic shocks. A similar specification is considered by Lochner and Shin (2014) and follows a long tradition of modeling the error component of earnings as a combination of permanent and transitory shocks, as reviewed by Meghir and Pistaferri (2011). In this case we would have that:

$$e_{it} = e_{it-1} + \mu_{it}$$

where e_{it} are individual i's unobserved skills at time t, as introduced in Equation (2). Assuming that μ_{it} is mean zero and allowing its variance to potentially vary over time and across demographic groups, we can re-write the within-group variance of unobserved ability among stayers as:

$$Var(e_{it}|a, c, t, s^t) \equiv \sigma^2_{a,c,t,s^t} = \sigma^2_{a-1,c,t-1,s^t} + \sigma^2_{\mu,a,c,t,s^t}$$

Allowing for measurement error – which can also be interpreted as an idiosyncratic temporary and non-persistent shocks to earnings – we have that the within-group residual wage variance for stayers would be given by:

$$\begin{aligned} V_{a,c,t,s^t} &= p_t^2 \sigma^2_{a,c,t,s^t} + \sigma^2_{\nu,a,c,t} \\ &= p_t^2 \left(\sigma^2_{a-1,c,t-1,s^t} + \sigma^2_{\mu,a,c,t,s^t}\right) + \sigma^2_{\nu,a,c,t} \end{aligned}$$

Identification will depend on the assumption that is made about the heterogeneities in the variances of the shocks across groups.

Age-specific shock variances

Assume that the variance of the permanent shock to earnings and the variance of measurement error depend on age and vary over time, but do not depend on education:

$$\sigma^2_{\mu,a,c,t,s^t} = \sigma^2_{\mu,a,t,s^t} \qquad \forall c$$
$$\sigma^2_{\nu,a,c,t} = \sigma^2_{\nu,a,t} \qquad \forall c$$

In this case, the return to skills can be identified using a Wald-type identification strategy, as suggested by Chay and Lee (2000). Specifically, price changes are identified from changes in the college-high school differential in residual wage variances over time, that is:

$$\frac{V_{a,COL,t,s^t} - V_{a,HS,t,s^t}}{V_{a-1,COL,t-1,s^t} - V_{a-1,HS,t-1,s^t}} = \frac{p_t^2\sigma^2_{a,COL,t,s^t} - p_t^2\sigma^2_{a,HS,t,s^t}}{p_{t-1}^2\sigma^2_{a-1,COL,t-1,s^t} - p_{t-1}^2\sigma^2_{a-1,HS,t-1,s^t}} = \frac{p_t^2}{p_{t-1}^2}$$

Due to the issues of top-coding for college graduates discussed earlier, we consider this identification strategy only for the years from 1989 onwards.

Education-specific shock variances

Alternatively, we could assume that the variance of the permanent shock to earnings and the variance of measurement error depend on education and vary over time, but do not depend on age:

$$\sigma^2_{\mu,a,c,t,s^t} = \sigma^2_{\mu,c,t,s^t} \qquad \forall a$$
$$\sigma^2_{\nu,a,c,t} = \sigma^2_{\nu,c,t} \qquad \forall a$$

In this case, the return to skills can be identified using a Wald-type identification strategy, as suggested by Chay and Lee (2000). Specifically, price changes are identified from changes in the college-high school differential in residual wage variances over time, that is:

$$\frac{V_{a,c,t,s^t} - V_{a',c,t,s^t}}{V_{a-1,c,t-1,s^t} - V_{a'-1,c,t-1,s^t}} = \frac{p_t^2 \sigma^2_{a,c,t,s^t} - p_t^2 \sigma^2_{a',c,t,s^t}}{p_{t-1}^2 \sigma^2_{a-1,c,t-1,s^t} - p_{t-1}^2 \sigma^2_{a'-1,c,t-1,s^t}}$$
$$= \frac{p_t^2}{p_{t-1}^2}$$

The results are presented for the estimated return to skills under the two different assumptions are presented in Figure 8. The source of identification for each of these series is very different and the results are also quite different between the two. In Panel A, when we assume that the variance of shocks is age-specific, we attribute any changes over time that are common across education groups to the life-cycle or idiosyncratic shocks. Intuitively, changes in the return to skills are identified from changes in the variance differential between college graduates and those without a college degree. The increase in this differential over time leads to the estimate of an increase in the return to skills over the 1990s and early 2000s. In contrast, in Panel B, when we assume that the variance of shocks is age-specific, and instead attribute any changes over time that are common across age groups to the life-cycle or idiosyncratic shocks, changes in the return to skills are identified from changes in the variance differential between older and younger workers. In this case we estimate an increase in the return to skills over the 1980s, followed by a sharp decrease in the 1990s and 2000s, due to the fall in the variance differential between older and younger workers over this period.

We conclude from this exercise that estimates of the return to skills are very sensitive to the assumptions made about the nature of the permanent life-cycle shocks to skills and the temporary idiosyncratic shocks to earnings. Whether we assume that these shocks are common across education groups or common across age groups will lead to very different answers in terms of the identified return to skills.

6 Conclusions

In this paper we develop and implement an identification strategy to estimate the return to skills which relaxes the assumption that the variance of unobserved skills is constant over time within groups of workers defined by their levels of education and age. This allows the dispersion of skills within groups to vary across cohorts due, for example, to changes in the characteristics of workers selecting into different education levels over time or changes in the dispersion of the quality of education. It also allows for changes in the dispersion of skills within cohorts over time due, for example, to changes in business cycle conditions, or heterogeneities in the prevalence of on-the-job training leading to heterogeneous permanent shocks to skills across individuals from a given cohort. Our identification strategy relies on longitudinal data, but requires only two observations per individual and thus can be implemented using a large scale dataset such as the CPS, due to its rotating sampling structure. Identification is achieved by considering changes in the within-group residual wage variance over time for workers who are observed over two consecutive years.

Our results suggest that relaxing the assumption of constant within-group variance is crucial. There have been important increases in the variance of residual wages, particularly among college-educated workers since the 1980s. Under the traditional assumption of constant skill variance within groups, these are interpreted as increases in the return to skills. By allowing for changes in the variance of skills we instead find that the return to skills in fact fell during the 1980s and early 1990s, and only partially recovered thereafter. Intuitively, this result reflects the fact that we do not observe widespread increases in the variance of wages when conditioning on the same group of workers across consecutive periods. This implies that the increase in the variance of residual wages is attributable to an increase in the dispersion of unobserved skills over time, particularly among college graduates, rather than an increase in the return to skills.

When we further relax our baseline assumption in order to allow for changes in the variance of unobserved ability even when conditioning on a common set of workers across two consecutive years, we find that the results are very sensitive to the assumption that is made about the nature of these changes. Assuming that shocks to skills differ across education groups but are common across age groups leads to dramatically different conclusions about the evolution of the return to skills as compared to the results when one assumes that the shocks differ across age groups but are common across education groups. Hence, assumptions about these shocks are not innocuous and must be carefully considered.

Overall, our results emphasize the inappropriateness of interpreting changes in within-group inequality as being indicative of changes in the return to skills. The empirical evidence does not provide robust support for the hypothesis that the return to skills has been increasing over time. Rather, the evidence suggests that increases in the distribution of unobservable skills among college graduates are an important driver of the increase in wage dispersion among this group.

References

Abowd, John M., Francis Kramarz, and David N. Margolis (1999), "High wage workers and high wage firms." *Econometrica*, 67, 251–333.

Acemoglu, Daron (2002), "Technical change, inequality, and the labor market." *Journal of Economic Literature*, 40, 7–72.

Autor, David H, Lawrence F Katz, and Alan B Krueger (1998), "Computing inequality: Have computers changed the labor market?" *The Quarterly Journal of Economics*, 113, 1169–1214.

Blundell, Richard, Michael Graber, and Magne Mogstad (2014), "Labor income dynamics and the insurance from taxes, transfers, and the family." *Journal of Public Economics*.

Blundell, Richard, Howard Reed, and Thomas M Stoker (2003), "Interpreting aggregate wage growth: The role of labor market participation." *The American Economic Review*, 93, 1114–1131.

Bound, John and George Johnson (1992), "Changes in the structure of wages in the 1980's: An evaluation of alternative explanations." *The American Economic Review*, 82, 371–392.

Bowlus, Audra J and Chris Robinson (2012), "Human capital prices, productivity, and growth." *The American Economic Review*, 102, 3483–3515.

Card, David, Jörg Heining, and Patrick Kline (2013), "Workplace heterogeneity and the rise of west german wage inequality." *The Quarterly Journal of Economics*, 128, 967–1015.

Card, David and Thomas Lemieux (2001), "Can Falling Supply Explain The Rising Return To College For Younger Men? A Cohort-Based Analysis." *The Quarterly Journal of Economics*, 116, 705–746.

Carneiro, Pedro and Sokbae Lee (2011), "Trends in quality-adjusted skill premia in the United States, 1960–2000." *The American Economic Review*, 101, 2309–2349.

Chay, Kenneth Y and David S Lee (2000), "Changes in relative wages in the 1980s: Returns to observed and unobserved skills and black–white wage differentials." *Journal of Econometrics*, 99, 1–38.

Cortes, G.M. (2016), "Where Have the Middle-Wage Workers Gone? A Study of Polarization Using Panel Data." *Forthcoming, Journal of Labor Economics.*

Guvenen, Fatih and Burhanettin Kuruscu (2010), "A quantitative analysis of the evolution of the US wage distribution, 1970-2000." In *NBER Macroeconomics Annual 2009, Volume 24*, 227–276, University of Chicago Press.

Hamermesh, Daniel S. (2002), "12 million salaried workers are missing." *Industrial and Labor Relations Review*, 55, 649–666.

Heathcote, Jonathan, Kjetil Storesletten, and Giovanni L. Violante (2010), "The Macroeconomic Implications of Rising Wage Inequality in the United States." *Journal of Political Economy*, 118, 681–722.

Heckman, James J, Lance Lochner, and Christopher Taber (1998), "Explaining rising wage inequality: Explorations with a dynamic general equilibrium model of labor earnings with heterogeneous agents." *Review of Economic Dynamics*, 1, 1–58.

Hendricks, Lutz and Todd Schoellman (2014), "Student abilities during the expansion of US education." *Journal of Monetary Economics*, 63, 19–36.

Hidalgo-Pérez, Manuel and Benedetto Molinari (2014), "Learning the new technology: the polarization of the wage distribution." *Working Paper.*

Juhn, Chinhui, Kevin M. Murphy, and Brooks Pierce (1993), "Wage inequality and the rise in returns to skill." *Journal of Political Economy*, 410–442.

Katz, Lawrence F. and Kevin M. Murphy (1992), "Changes in relative wages, 1963–1987: Supply and demand factors." *The Quarterly Journal of Economics*, 107, 35.

Lemieux, Thomas (2006), "Increasing residual wage inequality: Composition effects, noisy data, or rising demand for skill?" *The American Economic Review*, 461–498.

Lochner, Lance and Youngki Shin (2014), "Understanding earnings dynamics: Identifying and estimating the changing roles of unobserved ability, permanent and transitory shocks." *Working Paper, National Bureau of Economic Research*.

Meghir, Costas and Luigi Pistaferri (2004), "Income variance dynamics and heterogeneity." *Econometrica*, 72, 1–32.

Meghir, Costas and Luigi Pistaferri (2011), "Earnings, consumption and life cycle choices." *Handbook of Labor Economics*, 4, 773–854.

Moffitt, Robert A and Peter Gottschalk (2012), "Trends in the transitory variance of male earnings: Methods and evidence." *Journal of Human Resources*, 47, 204–236.

Nekarda, Christopher J (2009), "A Longitudinal Analysis of the Current Population Survey: Assessing the Cyclical Bias of Geographic Mobility." *Federal Reserve Board of Governors*.

Wooldridge, J.M. (2002), *Econometric analysis of cross section and panel data*. The MIT press.

Figure 1: Data structure: Workers in period t

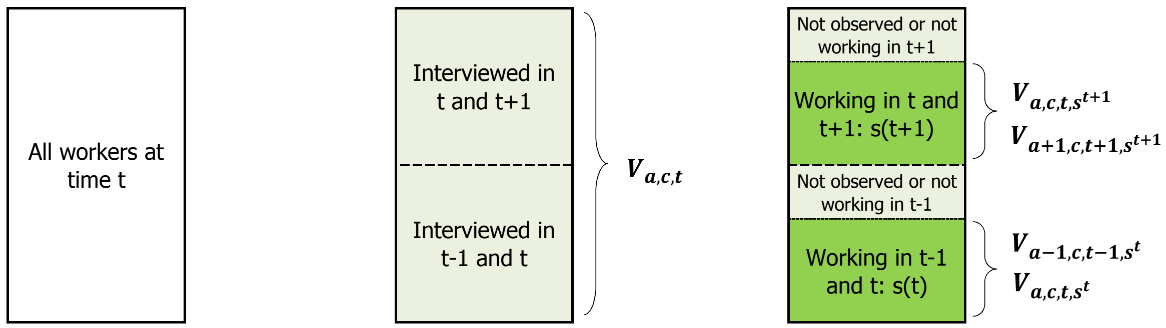


Figure 2: Within-group variance of residual wages

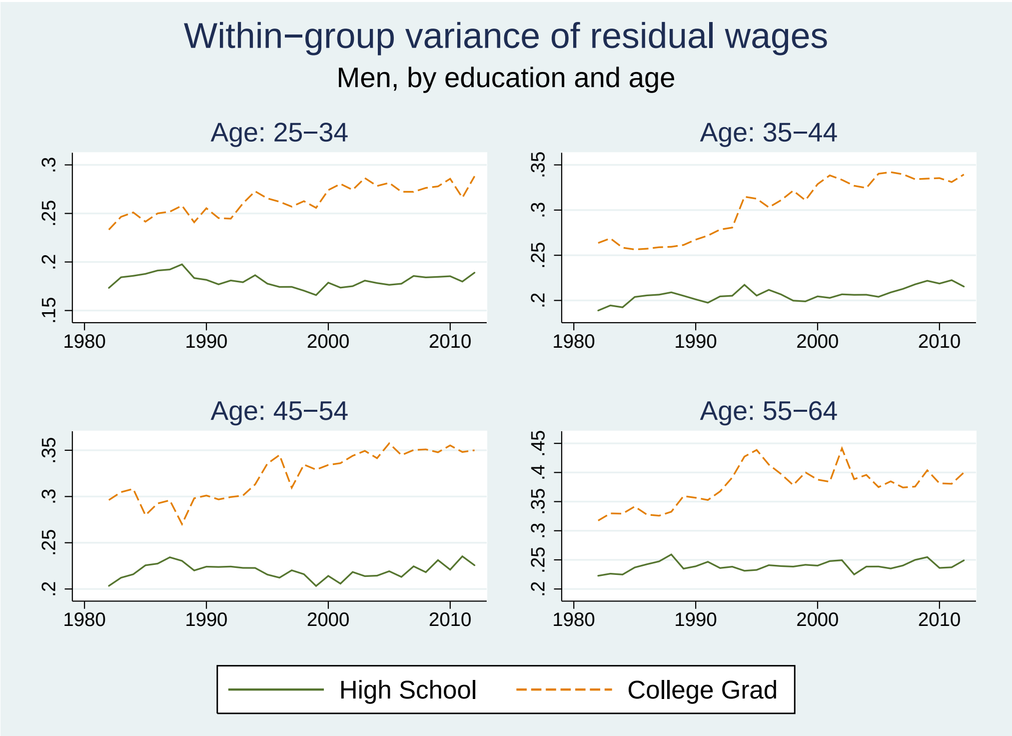


Figure 3: Estimated Return to Skills

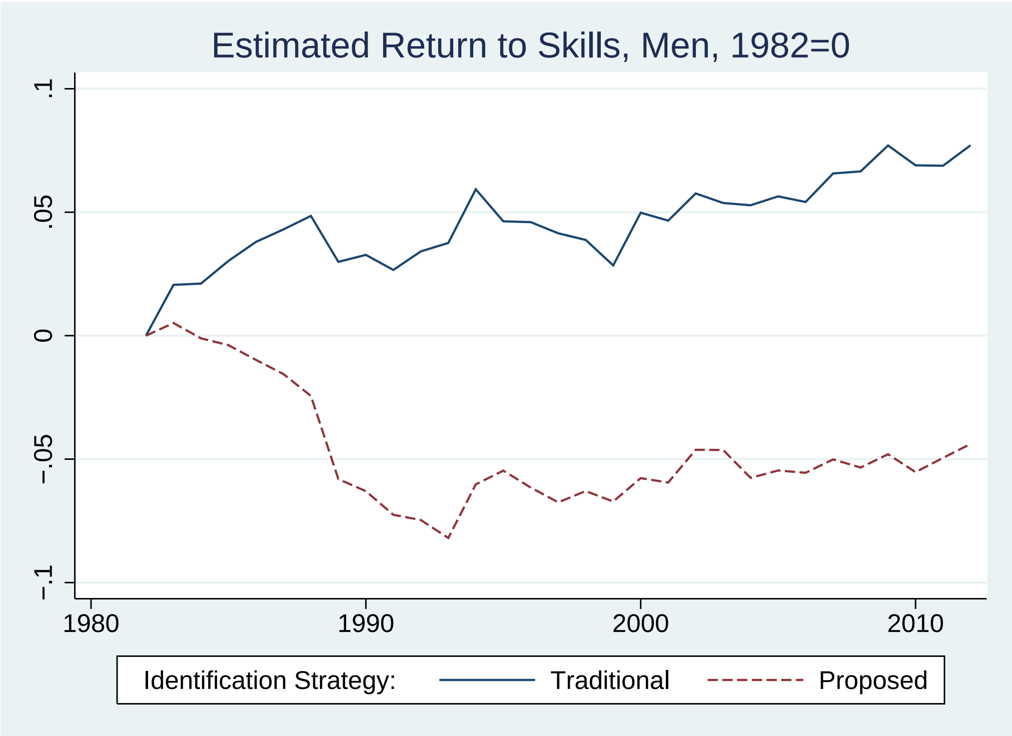


Figure 4: Estimated Changes in Variance of Unobserved Ability

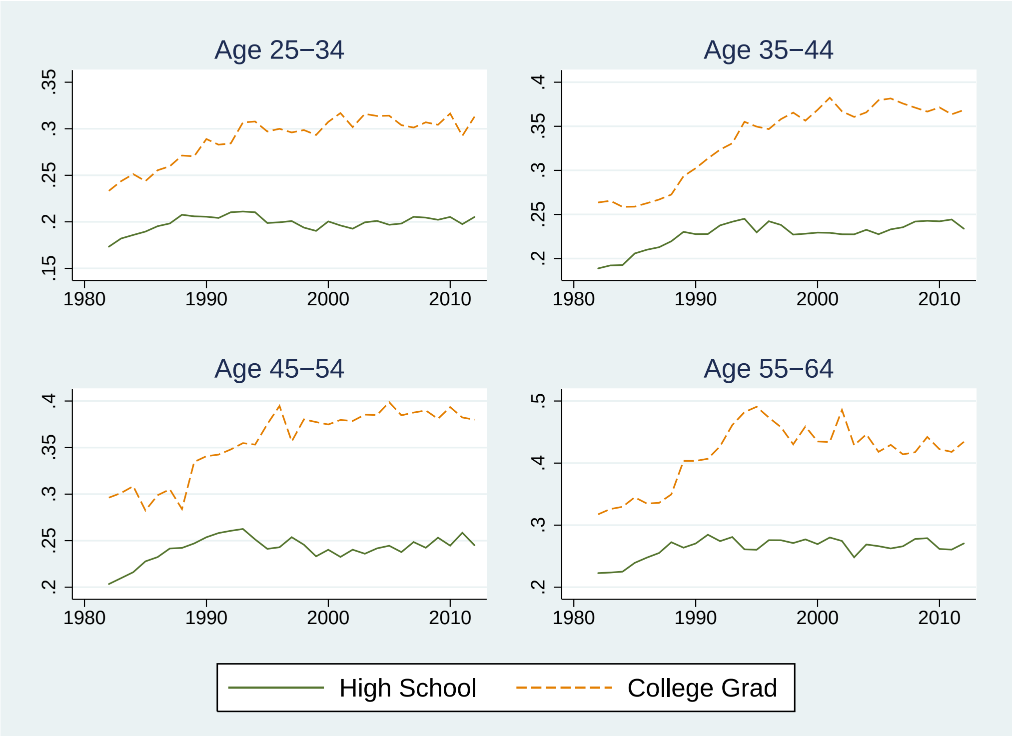


Figure 5: Sample Selection

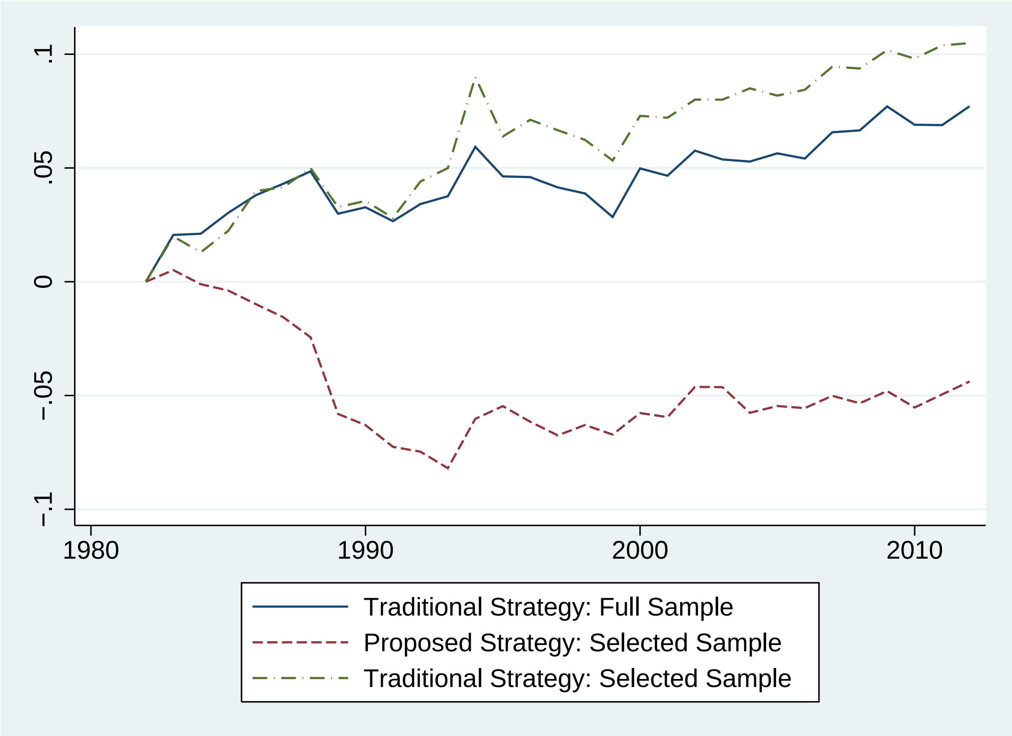


Figure 6: Group-specific returns to skills

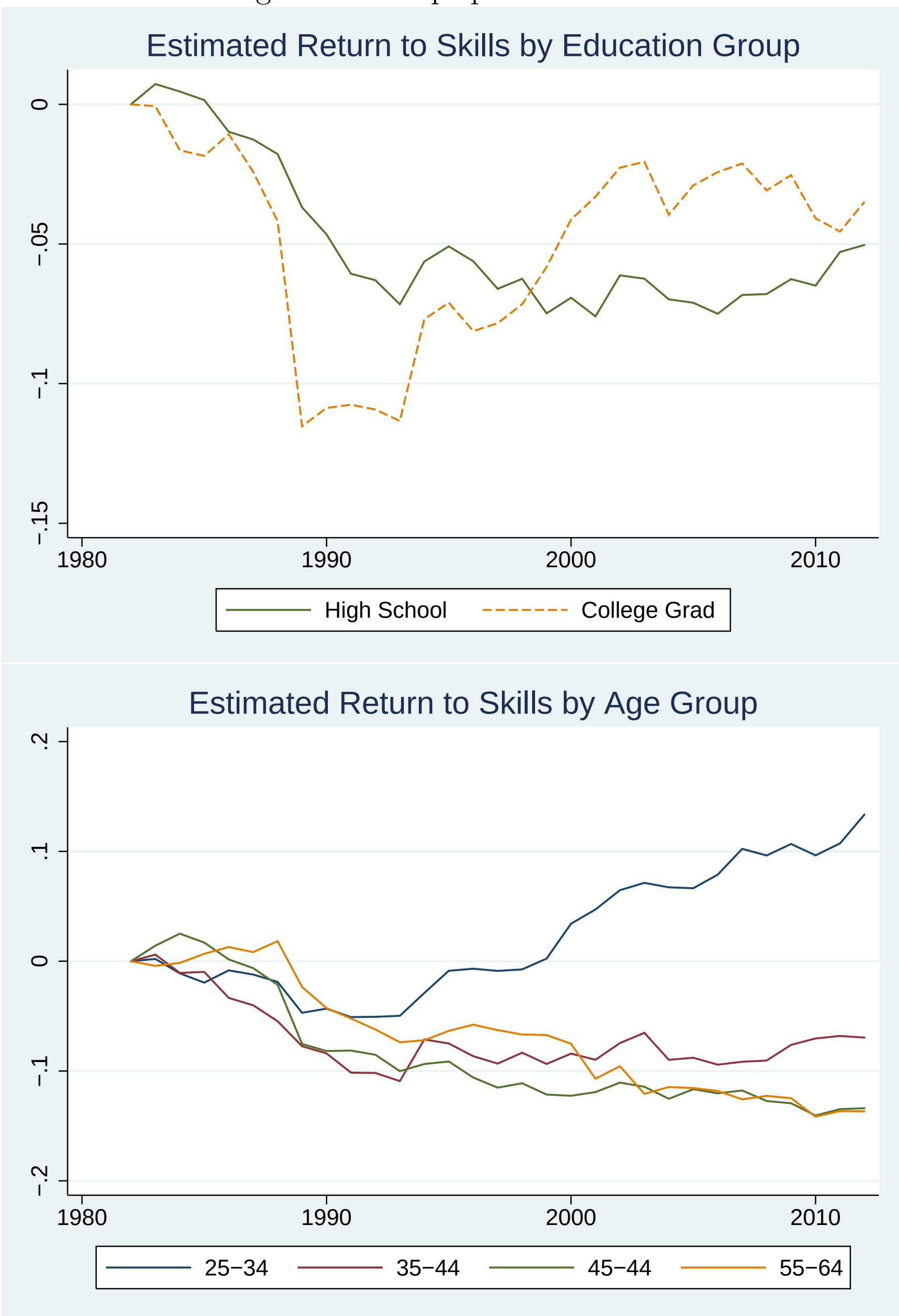
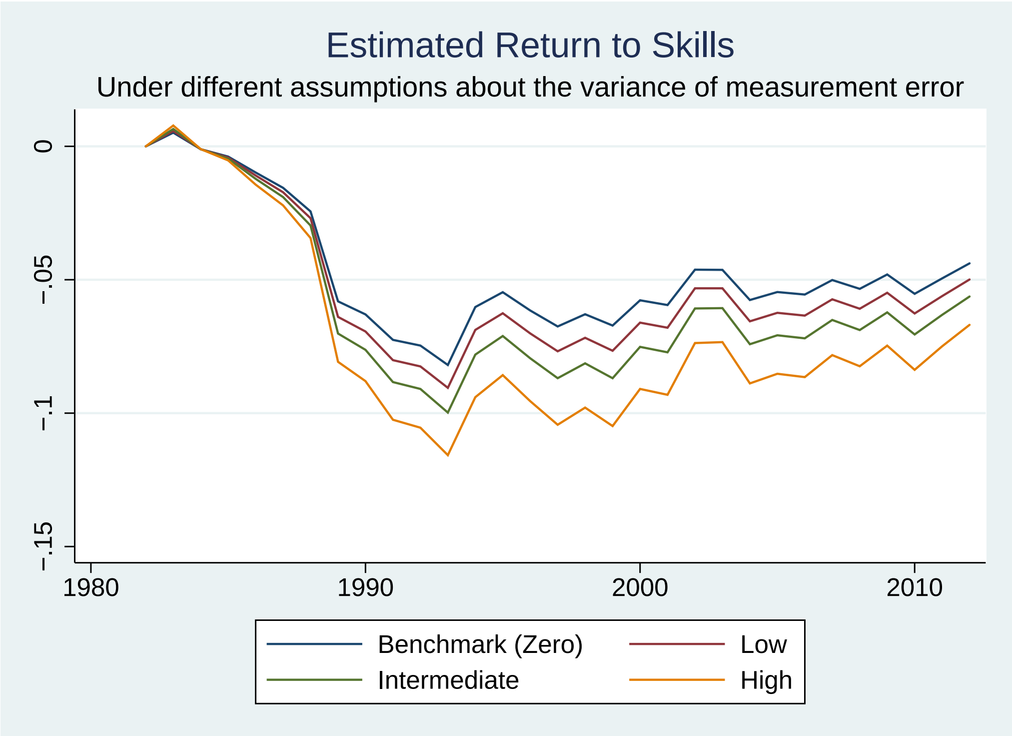


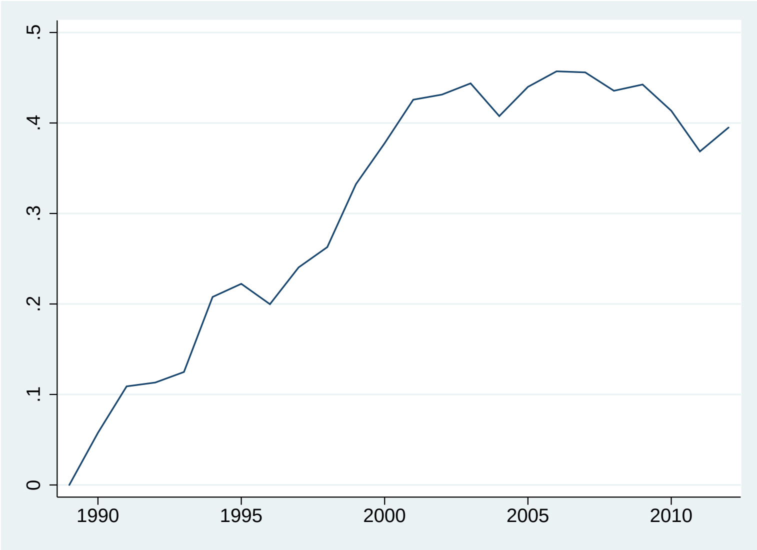
Figure 7: External estimates for measurement error



Note: The low, intermediate and high assumptions about the variance of measurement error are 0.017, 0.032, and 0.052, based on estimates from Lemieux (2006).

Figure 8: Estimated Return to Skills Allowing for Permanent and Transitory Shocks

Panel A: Assuming variance of shocks is age-specific – identification from changes in residual wage variance between education groups (base year: 1989)



Panel B: Assuming variance of shocks is education-specific – identification from changes in residual wage variance between age groups (base year: 1982)

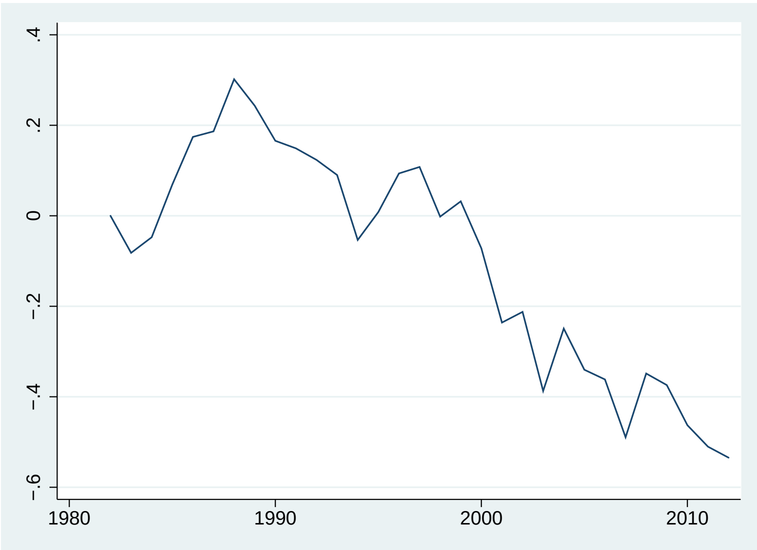


Table 1: Decomposition Results

	Residual wage variance			Decomposition	
	1982-1984	2010-2012	Change	Price Effect	Distrib Effect
High School					
25 to 34	0.181	0.185	0.004	-0.019	0.022
35 to 44	0.192	0.219	0.027	-0.022	0.049
45 to 54	0.211	0.227	0.017	-0.023	0.040
55 to 64	0.225	0.241	0.016	-0.024	0.041
College Grad					
25 to 34	0.244	0.280	0.036	-0.028	0.065
35 to 44	0.264	0.335	0.072	-0.034	0.106
45 to 54	0.303	0.351	0.048	-0.036	0.084
55 to 64	0.326	0.387	0.062	-0.039	0.101

Appendix A Assumptions about the relationship between observable and unobservable skills

Suppose that wages depend on education, age, and interactions of these variables. For simplicity, consider the case where education is a categorical variable which may adopt one of three values (high school dropout, high school graduate, college graduate), and the effect of age on wages is linear. The wage equation is given by:

$$w_{it} = \beta_{0t}+\beta_{1t}A_{it}+\beta_{2t}E_{2it}+\beta_{3t}E_{3it}+\beta_{4t}E_{2it}A_{it}+\beta_{5t}E_{3it}A_{it}+p_t e_{it}+\nu_{it} \quad \text{(A.1)}$$

Suppose that unobservable skills e_{it} are correlated with education. Following Wooldridge (2002), we can write the linear projection of e_{it} onto the observable explanatory variables as:

$$e_{it} = \pi_{0t} + \pi_{1t}A_{it} + \pi_{2t}E_{2it} + \pi_{3t}E_{3it} + \pi_{4t}E_{2it}A_{it} + \pi_{5t}E_{3it}A_{it} + \epsilon_{it} \quad \text{(A.2)}$$

Substituting equation (A.2) into (A.1) gives:

$$\begin{aligned} w_{it} = (\beta_{0t} + p_t\pi_{0t}) + (\beta_{1t} + p_t\pi_{1t})A_{it} + (\beta_{2t} + p_t\pi_{2t})E_{2it} + (\beta_{3t} + p_t\pi_{3t})E_{3it} \\ + (\beta_{4t} + p_t\pi_{4t})E_{2it}A_{it} + (\beta_{5t} + p_t\pi_{5t})E_{3it}A_{it} + p_t\epsilon_{it} + \nu_{it} \end{aligned}$$

The error term $p_t\epsilon_{it} + \nu_{it}$ has zero mean and is uncorrelated with each regressor. The plim of the OLS estimators from the wage regression are therefore $\beta_{kt} + p_t\pi_{kt}$, $k = 1, 2, \ldots 5$. In other words, the residual wages that we obtain are:

$$\begin{aligned} \widetilde{u}_{it} &= p_t\epsilon_{it} + \nu_{it} \\ &= p_t(e_{it} - \pi_{0t} - \pi_{1t}A_{it} - \pi_{2t}E_{2it} - \pi_{3t}E_{3it} - \pi_{4t}E_{2it}A_{it} - \pi_{5t}E_{3it}A_{it}) + \nu_{it} \end{aligned} \quad \text{(A.3)}$$

rather than $p_t e_{it} + \nu_{it}$.

The variance of this residual within an education-age group (where education c is one of the three categories mentioned above and age group

a comprises a 10-year age group) at a given time t would be given by:

$$\begin{aligned}
V_{a,c,t} &\equiv Var(\widetilde{u}_{it} \mid a, c, t) \\
&= Var\left[p_t(e_{it} - \pi_{0t} - \pi_{1t}A_{it} - \pi_{2t}E_{2it} - \pi_{3t}E_{3it} - \pi_{4t}E_{2it}A_{it} - \pi_{5t}E_{3it}A_{it}) + \nu_{it} \mid a, c, t\right] \\
&= Var\left[p_t(e_{it} - \widetilde{\pi}_{c,t}A_{it}) + \nu_{it} \mid a, c, t\right] \\
&= p_t^2\sigma_{a,c,t}^2 + \sigma_{\nu t}^2 - p_t^2\widetilde{\pi}_{c,t}^2 Var(A_{it} \mid a, c, t) \qquad \text{(A.4)}
\end{aligned}$$

where we have used the fact that education does not vary within a group, but there is some variation in age given that we aggregate workers into 10-year age groups. All terms that are constant across individuals within groups have been dropped as they do not affect the within-group variance at time t. $\widetilde{\pi}_{c,t}$ is equal to π_{1t} for high school dropouts, $\pi_{1t} + \pi_{4t}$ for high school graduates and $\pi_{1t} + \pi_{5t}$ for college graduates.

The key assumption for our identification strategy to be valid is $\pi_{a,c,t} = \pi_{a,c}\ \forall\ t$. This means that the partial correlation of age with unobserved ability (conditional on education) is constant over time. We also require that $Var(A_{it} \mid a, c, t)$ is constant over time. For our identification strategy, we only require this variance to remain constant over consecutive periods for stayers, which holds by construction.[1] Define $\sigma_{A,c}^2 \equiv Var(A_{it} \mid a, c, t)\ \forall\ t$. With these assumptions, and abstracting from measurement error, we have:

$$V_{a,c,t} = p_t^2\left(\sigma_{a,c,t}^2 - \pi_{a,c}^2\sigma_{A,c}^2\right) \qquad \text{(A.5)}$$

Therefore, under these assumptions, the identification approach described in Section 2 to identify changes in the return to unobservable skills can be implemented using changes over time in within-group variances of *residual* (rather than real) wages. The estimated within-group variance of ability includes the term $-\pi_{a,c}^2\sigma_{A,c}^2$, but we focus on the *changes* in the within-group variance over time, so as long as this term is (approximately) constant, this does not affect our results.

Note that the above discussion requires that each of our education-experience groups is composed of workers who are homogenous in terms of their education level. In practice, due to the small number of high school

[1]In the data, there are minor deviations for stayers due to measurement error in age.

dropouts in our sample (particularly for more recent years and when conditioning on experience level and on staying in the sample over consecutive periods), we pool high school dropouts and high school graduates. However, we have verified that our empirical findings go through if we exclude high school dropouts from our sample.