DemoSoc Working Paper

Paper Number 2010--34

Is there a wage penalty for Horizontal and Vertical Mismatch?

Aleksander Kucel
E-mail: aleksander.kucel@urv.cat

and

Montserrat Vilalta-Bufí
E-mail: montsevilalta@ub.edu

March, 2010

Department of Political & Social Sciences
Universitat Pompeu Fabra
Ramon Trias Fargas, 25-27
08005 Barcelona
http://www.upf.edu/dcpis/

Abstract

This paper studies how the horizontal and vertical mismatches in the labor market affect wage. We do so by taking into account that by choosing a job, wage and mismatches are simultaneously determined. The Seemingly Unrelated Equations model also allows us to control for any omitted variable that could cause biased estimators. We use REFLEX data for Spain. Results reveal that in most cases being horizontally matched has a wage premium and being over-educated does not affect wage. Results suggest that the modeling strategy successfully accounts for some omitted variable that affects simultaneously the probability of being horizontally matched and the wage. This could explain the existence of a wage penalty for over-educated workers when the omitted variable issue is not dealt with.

Keywords

Mismatch, Over-education, wage, SURE Models.

Acknowledgements

Aleksander Kucel, Employment Observatory and Economics Department, Universitat Rovira i Virgili. Montserrat Vilalta-Bufí, Dept. Teoria Econòmica and CREB, Universitat de Barcelona. Her contribution was financially supported by the Generalitat de Catalunya through the grant 2009SGR1051 and the Ministerio de Ciencia y Educación of Spain through the grant ECO2009-06953. We would like to thank Fabian Gouret for his comments.

1. Introduction

This paper studies how the horizontal and vertical mismatches in the labor market affect wage. We refer to horizontal match when the own or a related field of study are the most appropriate for the job. When a completely different field or no particular field is appropriate for the job we assume that the individual is horizontally mismatched. Vertical mismatch instead refers to the difference between the level of education acquired and the level of education required on the job. In this paper we focus on the case of over-education, that is, when the level of education attained is higher than the level of education required in the job.

In order to properly analyze the influence of mismatches on wage we need to take into account that by choosing a job, wage and match quality are simultaneously determined. From the supply side, when an individual accepts a job offer, s/he is agreeing at once to be or not mismatched (horizontally and/or vertically) as well as accepts a particular wage. From the demand side, when the firm selects an individual for a job, it offers him/her a particular wage knowing his/her quality of vertical and horizontal match. Our modeling strategy is to use the Seemingly Unrelated Equations model, which not only allows us to have simultaneity but also to control for any omitted variable that could affect the three variables and cause biased estimators.

It is clearly established in the literature that being over-educated comes with a wage penalty. A growing body of literature is studying whether this wage penalty is caused by over-education itself or whether is the result of an omitted variable problem (Halaby, 1994, Hartog, 2000, McGuinness, 2006). At least three different strategies have been used to address this question. First of all, some authors have used panel data to control for unobserved individual heterogeneity (Bauer, 2002). Fixed effects estimations are however imprecise if most of the variation in a regressor is cross-sectional rather than over time, as it is the case with over-education. Moreover, inconsistency due to measurement error is magnified when differencing the data (Dolton and Silles, 2008). A second approach to control for unobserved individual heterogeneity has been to use quantile regressions (McGuinness and Bennett, 2007). Quantile regressions allow controlling for heterogeneity of individuals' abilities across the wage distribution. It is assumed that different quantiles in the wage distribution have different average ability. Although it gives informative results, it does not fully control for all unobserved heterogeneity. McGuinness and Sloane (2009) use instead the propensity score matching, using as control group those individuals that were over-educated in the first job and got matched in the current job. They obtain similar results to the OLS estimation, suggesting that unobserved heterogeneity is not important. Other authors introduce in the Mincerian wage equation variables measuring the otherwise omitted variables, such as some proxy for ability and mismatch of skills (Frenette, 2004, Green and McIntosh, 2007). Although there are all the more datasets with such information, these variables are likely to have large measurement error and they may not control for all the unobserved individual heterogeneity (Hartog, 2000).

We contribute to this debate by considering the effect of simultaneity on the decision on wage and quality of job match. By using simultaneous equations on wage, horizontal match and over-education we control for any omitted variable that would affect the three equations at once. We use REFLEX data for Spain. Results reveal that being horizontally matched has a wage premium for males and reduces the probability of being over-educated. On the other hand, being over-educated does not affect wage when using the simultaneous estimation for the general sample and females. This result contrasts with most of the studies on over-education penalty, which find a significant lower wage for overeducated workers. Nevertheless, we obtain a wage penalty of 17% for overeducated males, in line with the wage penalty found in the literature (between 8 and 27%, with a mean of 15.3% (McGuinness, 2006)).

The rest of the paper is organized as follows. Next we briefly review the literature on wage penalty due to labor mismatches. In section 3 we describe the data and in section 4 the methodology used. Then we present the results and conclude.

2. Literature review

Most of the literature on labor mismatches has concentrated on their wage effect. [1] It is now well established that being over-educated has a wage penalty when comparing

[1] We only review here those papers directly comparable to our results, that is, those studying the effect of over-education on wage which use the Mincerian specification. There is an alternative specification, called ORU model, which consists of introducing the variables *required education*, *over-education* and *under-education* in the regression instead of acquired education and dummies for over-education and under-education. The main difference is in the interpretation of the coefficients. While in the ORU

among individuals with the same level of education (for extensive surveys see Chevalier (2003), Hartog (2000) and McGuinness (2006)). Similar results are found when studying the wage effect of being horizontally mismatched (Robst, 2007).

The main question in the last decade has been whether the wage penalty for over-education persists when other factors are controlled for. Many researchers study whether wage penalty is the result of an omitted variable problem. Several variables have been raised as possible candidates to explain this over-education wage penalty. In the first place, it is argued that over-educated individuals may be missing skills, which would explain why they earn a lower salary. Allen and van der Velden (2001) follow this argument, but they find that, even after controlling for skill-mismatch, over-education brings a strong wage penalty. Chevalier and Lindley (2009) distinguish between apparently and genuinely over-educated. They find that the genuinely over-educated miss non-academic skills (managerial and leadership), which could explain why they get a lower salary. Green and McIntosh (2007) also introduce skill mismatch in the wage analysis. They find that the impact of over-education on wage does not change, neither in size, nor in significance. Introducing measures of skill mismatch is a valid approach to unravel the effect of over-education on wage, yet other variables such as ability or personality may remain omitted. Moreover, it is often difficult to develop a good measure of these variables.

The second hypothesis is that over-educated individuals are simply less able than matched ones. Bauer (2002) controls for individual heterogeneity by using panel data. Wage differences between over-educated and matched individuals decrease or even disappear in some cases. McGuinness and Bennet (2007) use quantile regressions in the Mincer specification and find opposite results.

In conclusion, there is some evidence that the observed wage penalty for over-educated workers may at least partially exist due to missing skills or ability. In rare occasions, though, does this wage penalty completely disappear from the equation. This would mean that being over-educated in itself causes lower earnings to the individuals.

---

model over-educated individuals are compared to their matched colleagues in the same job, in the dummy model we compare them to matched individuals with the same level of education.

In this paper we present yet another strategy to try to disentangle to what extent there is a wage penalty to being over-educated and to what extent it is an artifact of the econometric modeling strategy. We consider jointly the determination of wage and vertical and horizontal mismatch, controlling for endogeneity of these variables.

3. Data

We use the REFLEX data for Spain. It consists on information of individuals who graduated in 1999/2000 and were interviewed five years later in 2005. We have information on their graduate studies, personal background, first job and current job. We restrict our sample to those individuals who are currently working more than 20 hours per week and are not self-employed. We are left with 2.581 individuals.

Hmatch stands for horizontal match. It is a dummy variable that takes value 1 when the individual reports that exclusively own field or own or a related field is most appropriate for his/her work, 0 otherwise. *Overed* stands for Over-education and takes value 1 when the required level of education reported by the individual is lower than the level of education attained and 0 otherwise. Both of these variables are therefore subjective. The third dependent variable is *lnghWage*, which stands for the log of gross hourly wage. We also define the variable *Undered* which takes value 1 if the individual has less education than s/he feels required by the job, 0 otherwise.

From this sample, 30.4% of the individuals consider themselves over-educated, while only 7% think they are under-educated. In relation to horizontal mismatch, 82% of the individuals consider they are working in their own or related field, while 8% declare working in a job that requires a completely different field or no particular field. The correlation between over-education and horizontal match is large and significant($\rho = -0.49$). The 66% of the sample are vertically and horizontally matched, while 14% of the individuals are vertically and horizontally mismatched. Close to 4% are only horizontally mismatched, and above 16% are simply overeducated. The gross hourly wage ranges from 2.3 to 24 Euros, being 8.50 Euros the average value.

Significant majority of the sample is female (62.6%) and around 45% are living with a partner (variable *Married*). Immigrants represent close to 2% of the sample. Average age is close to 30 with the youngest individual being 26 and the oldest 86. Above one third of the individuals (35%) studied the field Social Sciences, Business and Law.

Science, Mathematics and Computer Science were the choice of 15% of the sample, as much as for Engineering, Manufacturing and Construction. The field Health and Welfare is reported by 12% of individuals, followed closely by Education (11%). The fields less represented in the sample are Humanities and Arts (6.5%) and Agriculture and Veterinary (below 4%).

Table 1 shows the incidence of horizontal match, over-education and the mean of wages per field. Fields such as Engineering and Health present the smallest incidence of over-education and the largest value for horizontal match. Wages are the largest in the field of Engineering as well.

Table 1. Distribution of educational mismatch and *lnghwage* by field of study.

Fields of Study		Horizontal Match	Over-education	*lnghWage*
Education	Mean	**0.76**	**0.38**	**20.1**
	N	311	313	296
Humanities and arts	Mean	**0.63**	**0.42**	**20.1**
	N	180	180	170
Social sciences, business and law	Mean	**0.81**	**0.39**	**20.1**
	N	945	942	907
Science, mathematics and computing	Mean	**0.79**	**0.28**	**20.1**
	N	400	395	373
Engineering, manufacturing and construction	Mean	**0.92**	**0.18**	**20.3**
	N	425	426	380
Agriculture and veterinary	Mean	**0.83**	**0.29**	**20.1**
	N	104	104	102
Health and welfare	Mean	**0.92**	**0.14**	**20.1**
	N	343	342	326
Total	Mean	**0.82**	**0.31**	**20.1**
	N	2708	2702	2554

Table 2 gives the summary statistics of the main variables of interest. *Experience* tells us how many months has the respondent worked since graduation and *tenure* how many months has the respondent been working for the current firm. Since Reflex data is collected on university graduates, education dummies correspond to different university degree levels in Spain, that is, *diplomatura* (3 years), *licenciatura* (4-5 years) and master or doctorate (*doctorate*). *Gradsec* is the average grade in secondary education. We use this variable as a measure of ability. *Prestige*, *vocational*, *demanding*, *broad*, *freedom* and *empfml* are dummies defining study program characteristics. Respondents were asked to indicate to what extent their study program was academically prestigious (*Prestige*), vocationally oriented (*vocational*), regarded as demanding (*demanding*), had a broad focus (*broad*), gave freedom to compose own program (*freedom*) and employers were familiar with it (*empfml*). We define each of these variables to be 1 if the answer was 4 or 5 from a 1 to 5 scale, 0 otherwise.

Public is a dummy variable indicating if the respondent is working in the public sector. Firm size dummy variables indicate the following size ranges: less than 10 employees (*firmsize_1*), from 10 to 49 (*firmsize_2*), from 50 to 99 (*firmsize_3*), from 100 to 249 (*firmsize_4*), from 250 to 999 (*firmsize_5*) and above 1000 employees (*firmsize_6*). *Firmtype* is yet another dummy variable which takes value 1 if the firm operates nationally or internationally and 0 if the firm operates locally or regionally. *Numemp* indicates the number of employers that the respondent had since graduation. And, finally, *supervise* indicates whether the respondent supervises other workers in his/her job. In the regression analysis we also include dummies on occupation titles and economic sectors.

Table 2. Summary statistics.

Variable	N	Mean	Std. Dev.	Min	Max
hmatch	2738	0.821	0.383	0	1
overed	2732	0.304	0.460	0	1
undered	2732	0.070	0.255	0	1
lnghwage	2581	20.139	0.393	0.836	30.180
female	2665	0.626	0.484	0	1
immigrant	2643	0.020	0.139	0	1
married	2657	0.454	0.498	0	1
experience	2598	500.758	150.460	0	84
tenure	2611	370.761	340.168	0	401
edu_1 (diplomatura)	2746	0.310	0.463	0	1
edu_2 (licenciatura)	2746	0.662	0.473	0	1
edu_3 (doctorado)	2746	0.028	0.164	0	1
gradsec	2720	20.874	0.931	1	5
prestige	2721	0.358	0.479	0	1
vocational	2714	0.235	0.424	0	1
demanding	2736	0.575	0.494	0	1
broad	2720	0.538	0.499	0	1
freedom	2723	0.317	0.466	0	1
empfml	2663	0.411	0.492	0	1
Education	2749	0.114	0.318	0	1
Humanities	2749	0.065	0.247	0	1
Social Sciences	2749	0.345	0.475	0	1
Science, Maths	2749	0.146	0.353	0	1
Engineering	2749	0.155	0.362	0	1
Agric. & Vet	2749	0.038	0.192	0	1
Health	2749	0.125	0.331	0	1
public	2727	0.327	0.469	0	1
firmsize_1	2562	0.101	0.301	0	1
firmsize_2	2562	0.179	0.384	0	1
firmsize_3	2562	0.101	0.302	0	1
firmsize_4	2562	0.094	0.292	0	1
firmsize_5	2562	0.126	0.332	0	1
firmsize_6	2562	0.399	0.490	0	1
firmtype	2703	0.565	0.496	0	1
numemp	2618	20.969	30.489	0	98
supervise	2720	0.381	0.486	0	1

4. Methodology

We base on the Seemingly Unrelated Equations model, using the Stata module cmp developed by David Roodman (2009). We estimate a system of three equations with Horizontal Match, Over-education and LnghWage as dependent variables. They form a recursive system of simultaneous equations which can be consistently estimated using the cmp module. Horizontal Match is explained by standard individual and job characteristics (gender, experience, tenure, immigrant, married, education level, field of study, field characteristics, industry, occupation, firm size, firm type, number of employers). The regression on Over-education has the same explanatory variables as the former plus the variable Horizontal Match. The justification is simple. When a worker is horizontally mismatched then it is more likely that his/her studies will not be valued in the job, and therefore s/he is more likely to be over-educated. And vice versa, a horizontally matched worker is less likely to be over-educated for the same reasoning. Finally, the regression on LnghWage includes a group of exogenous explanatory variables plus the two endogenous variables Horizontal Match and Over-education. In the LnghWage equation the set of explanatory variables contains in addition to the ones mentioned above the variable Supervise (whether you supervise other workers in the job).

$$Pr(Hmatch = 1|X) = F(X\beta)$$
$$Pr(Overed = 1|X) = F(X\delta + \gamma Hmatch)$$
$$lnghWage = Z\lambda + \theta Hmatch + \eta Overed + \zeta$$

where: $F(\cdot)$ is the standard normal cdf.

We estimate the three equations simultaneously. Any omitted variable that affects more than one equation at once will be accounted for in the correlation of the error terms. It is sensible to think that most omitted variables affecting wage would also have an impact on mismatch, and vice versa. Take ability for instance. Low ability individuals are thought to be more likely to get over-educated and get a lower wage. Similarly, we could argue that lack of some skills would hamper the individual in both terms too. Therefore, we expect to be correcting for the main omitted variables pointed out in the literature as possibly affecting the over-education wage penalty. Moreover, the specification of the model takes into account the simultaneity decision on horizontal match, over-education and wage.

Since the system is recursive, it is identified without need for any restrictions. Notice however that with this specification we could have multicollinearity problems. Since *Hmatch* depends on exactly the same exogenous variables as *Overed*, when introducing *Hmatch* on the *Overed* equation multicollinearity could occur. We are saved, though, by the functional form (Cameron and Trivedi, 2005). In this paper, the variation in $X\hat{\beta}$ across observations is large for all the configurations, therefore the estimated *Hmatch* is not linear and this potential problem is less severe.

5. Results

Table 3 presents the main results for the whole sample (see the Appendix for complete results). The three columns in Pane A correspond to probit regressions of horizontal match and over-education (*Hmatch* and *Overed*), and the OLS regression for wages (*lnghWage*), respectively. Pane B reports the results when the three equations are regressed in a simultaneous equation system using the cmp Stata module. Both strategies give similar results in all variables but in *Overed* and *Hmatch*.

Being horizontally matched decreases the likelihood of being over-educated as expected. We find also a wage premium for being horizontally matched, which gets nearly six times larger in the simultaneous equation setting. Yet the main change occurs in the over-education variable. Although from pane A we observe the existence of a wage penalty for over-education, once we estimate the simultaneous equation model *Overed* loses significance in the wage equation. This result tells us that, once we take into account the simultaneity issue and control for omitted variables, being over-educated does not give any wage penalty.

Table 3. Main results for the whole sample.

	A			B		
	Hmatch (probit)	Overed (probit)	Lnghwage (OLS)	Hmatch	overed	lnghwage
				Simultaneous equations (cmp)		
Mismatch Dimensions						
Hmatch		-1.256***	0.0913***		-1.290***	0.525***
		(-11.83)	(3.86)		(-4.61)	(10.35)
Overed			-0.117***			-0.0972
			(-5.77)			(-1.20)
Undered			-0.0376			-0.0301
			(-1.22)			(-0.97)
Individual characteristics, Level of education (ref. Diplomatura), Educational Program Attributes, Fields of study (ref. Humanities), Job characteristics, Occupation and Industry included.						
N	2027	2008	1943	2050		
chi2/R2	561.5	900.0	0.373	1749.6		
Sig_3	0.3324***					
Rho_12	0.0188					
Rho_13	-0.721***					
Rho_23	-0.0351					

t statistics in parentheses, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Most of the other variables have the expected effect on the dependent variables. Results are shown in the Appendix. Women are more likely to be over-educated and earn less than men. Experience decreases the likelihood of being over-educated and has a positive effect on wage. Having more education is positively correlated with over-education and wage, although PhD studies give similar salary as *Diplomatura*. The measure of ability (*Gradsec*) has a negative sign for over-education and positive for wage at a 0.05 significance level. Immigrant has a positive effect on wage possibly explained by the fact that immigrants who are university graduates are not representative of the total immigrant population in Spain. Individuals living with a partner (*Married*) are paid on average a higher wage.

From the group of variables describing characteristics of the study program (*Prestige, Vocational, Demanding, Broad, Freedom, Empfml*) only studying a prestigious one turns out significant and it is so for being horizontally matched. Engineering, Health and Social Sciences give the highest likelihood to being horizontally match, in this order. The likelihood of being overeducated is higher for those who studied Education, Science, Engineering and, to a lesser extent, Social Sciences. The effects of fields of study on wage change from pane A to pane B. While in pane A Engineering and Education have a positive effect on wage, when solving for the simultaneous equation system this positive effect disappears and Social Sciences and Health turn out to have a negative effect on wage.

Results reveal that working in the public sector decreases the chances of being overeducated and pays higher wages than working in the private sector. Larger firms also pay higher wages than smaller firms, as found in the literature. We find, however, that those firms operating nationally or internationally pay less than those operating regional or locally, although they decrease the likelihood of being overeducated. The number of employers since graduation has a small positive effect on over-education, while theories of job mobility clearly suggest the opposite. This is possibly explained by the fact that those individuals who change job more often are having precarious temporary employment doing substitutions and seasonal jobs. Finally, we obtain that supervising other workers is rewarded in the labor market.

We find a negative correlation between the errors of *Hmatch* and Wage equation. This indicates that we are omitting some variable that affect in opposite direction Hmatch and Wage. Our guess is that the level of vocation for the job could be this omitted variable. It would affect positively *Hmatch* and negatively *lnghWage*, since individuals with strong vocation for the job are willing to work at lower salaries in order to work on their field of interest.

Table 4 and 5 present the same analysis but for males and females, respectively.[2] While the negative effect of horizontal match on over-education holds in all the setups, the result on over-education wage penalty does not. We find that over-education has a wage penalty for males. However, while in pane A this effect is strongly significant, in the simultaneous equations estimation (pane B, Table 4) significance of over-education is only at the 0.10 level. Another interesting result is that the wage premium for being horizontally matched does not come up significant for females.

[2] Full results are reported in the appendix.

Table 4. Main results for males.

	A			B		
	Hmatch (probit)	Overed (probit)	Lnghwage (OLS)	Hmatch	overed	lnghwage
				Simultaneous equations (cmp)		
Mismatch Dimensions						
Hmatch		-1.239***	0.0678*		-0.933**	0.587***
		(-6.67)	(1.69)		(-2.30)	(9.18)
Overed			-0.116***			-0.174*
			(-3.42)			(-1.87)
Undered			-0.0888*			-0.0566
			(-1.92)			(-1.27)
Individual characteristics, Level of education (ref. Diplomatura), Educational Program Attributes, Fields of study (ref. Humanities), Job characteristics, Occupation and Industry included.						
N	789	794	744	805		
chi2/R2	146.8	245.7	0.370	676.4		
Sig_3	0.3328***					
Rho_12	-0.1676					
Rho_13	-0.8927***					
Rho_23	0.1662					

t statistics in parentheses, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

There are some other differences in the results, as can be seen in the Appendix. Let us discuss briefly only those more interesting. First of all, while working in the public sector decreases the likelihood of over-education and pays higher wages in the general sample and for females, males' results show that public sector increases the probability of being horizontally matched. Second, fields of study turn out significant when using the general and female samples, but field characteristics seem to take their role when using the male sample. Third, experience and ability (*gradsec*) decrease the likelihood of being overeducated for males (and also in the general sample), but they do not affect the likelihood of being overeducated for females. Also, living with a partner (*married*) has a positive effect on wage in all the analysis, yet the effect is double for males than for females. Finally, the correlation of the errors for horizontal match and wage is always significant, yet it is negative for males and positive for females. This suggests that different variables are omitted in each analysis, being a possibility "vocation" for the male sample and "ambition" in the female sample.

Table 5. Main results for females.

	A			B		
	Hmatch (probit)	Overed (probit)	Lnghwage (OLS)	Hmatch	overed	lnghwage
				Simultaneous equations (cmp)		
Mismatch Dimensions						
Hmatch		-1.342^{***}	0.0985^{***}		-1.670^{***}	-0.0857
		(-9.77)	(3.30)		(-4.10)	(-0.85)
Overed			-0.111^{***}			-0.0865
			(-4.38)			(-0.75)
Undered			0.00366			-0.0000258
			(0.09)			(-0.00)
Individual characteristics, Level of education (ref. Diplomatura), Educational Program Attributes, Fields of study (ref. Humanities), Job characteristics, Occupation and Industry included.						
N	1230	1212	1199		1245	
chi2/R2	400.4	604.4	0.392		1177.7	
Sig_3	0.3078^{***}					
Rho_12	0.1888					
Rho_13	0.3650^{**}					
Rho_23	-0.0257					

t statistics in parentheses, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

6. Conclusion

We propose a recursive system of equations to explain wage, horizontal match and over-education. Comparing the simultaneous equation estimation with standard probits and OLS regressions we find that the specification of the model matters when studying the effects of educational mismatch on wage. In particular, results show that the wage penalty associated to being over-educated gets insignificant when using the simultaneous equations approach for the general and female sample, and significance decreases in the male sample regression. Results suggest that the over-education wage penalty found in the literature is more the result of an omitted variable problem than over-education causing a wage reduction on itself. Which is the omitted variable? To answer this question more research needs to be done. Our results also show that being horizontally matched is good because it decreases the likelihood of being overeducated and has a positive effect on wage.

References

ALLEN, J. & VELDEN, R. V. D. (2001) Educational Mismatches vs. Skill Mismatches: Effects of Wages, Job Satisfaction and On-the-job Search. *Oxford Economic Papers,* 3**,** 434.

BAUER, T. K. (2002) Educational mismatch and wages: a panel analysis. *Economics of Education Review,* 21**,** 221.

CAMERON, C. A. & TRIVEDI, P. K. (2005) *Microeconometrics: Methods and Applications,* Cambridge University Press.

CHEVALIER, A. (2003) Measuring Over-education. *Economica,* 70.

CHEVALIER, A. & LINDLEY, J. (2009) Overeducation and the skills of UK graduates. *Journal of the Royal Statistical Society,* 172**,** 307-337.

DOLTON, P. J. & SILLES, M. A. (2008) The effects of over-education on earnings in the graduate labour market. *Economics of Education Review,* 27**,** 125-139.

DUNCAN, G. J. & HOFFMAN, S. D. (1981) The Incidence and Wage Effects of Overeducation. *Economics of Education Review,* 1**,** 75.

FRENETTE, M. (2004) The overqualified Canadian graduate: the role of the academic program in the incidence, persistence, and economic returns to overqualification. *Economics of Education Review,* 23**,** 29.

GREEN, F. & MCINTOSH, S. (2007) Is there a genuine under-utilization of skills amongst the over-qualified? *Applied Economics,* 39**,** 427.

HALABY, C. N. (1994) Overeducation and Skill Mismatch. *Sociology of Education,* 67**,** 47.

HARTOG, J. (2000) Over-education and earnings: where are we and where should we go? *Economics of Education Review,* 19**,** 131.

MCGUINNESS, S. (2006) Overeducation in the labor market. *Journal of Economic Surveys,* 20**,** 387.

MCGUINNESS, S. & BENNETT, J. (2007) Overeducation in the graduate labour market: A quantile regression approach. *Economics of Education Review,* 26**,** 521-531.

MCGUINNESS, S. & SLOANE, P. (2009) Labour Market Mismatch Among UK Graduates; An Analysis Using REFLEX Data. *ESRI Working Papers,* 294**,** 1-56.

ROBST, J. (2007) Education and job match: The relatedness of college major and work. *Economics of Education Review,* 26**,** 397.

ROODMAN, D. (2009) Estimating fully observed recursive mixed-process models with cmp. *Center for Global Development Working Paper,* 168, 1-57.

Appendix

Table 3 (complete results). General sample.

	A			B		
	Hmatch (probit)	Overed (probit)	Lnghwage (OLS)	Hmatch	overed	lnghwage
				Simultaneous equations (cmp)		
Mismatch Dimensions						
Hmatch		-1.256***	0.0913***		-1.290***	0.525***
		(-11.83)	(3.86)		(-4.61)	(10.35)
Overed			-0.117***			-0.0972
			(-5.77)			(-1.20)
Undered			-0.0376			-0.0301
			(-1.22)			(-0.97)
Individual characteristics						
Female	-0.0654	0.237***	-0.0528***	-0.0709	0.235***	-0.0469***
	(-0.74)	(2.82)	(-3.35)	(-0.86)	(2.79)	(-2.70)
Immigrant	-0.168	-0.598	0.243***	-0.0743	-0.600	0.247***
	(-0.56)	(-1.49)	(4.06)	(-0.27)	(-1.48)	(3.88)
Married	-0.0756	0.0650	0.0754***	-0.0802	0.0639	0.0819***
	(-0.91)	(0.83)	(5.12)	(-1.04)	(0.81)	(5.17)
Experience	0.00413	-0.00849***	0.00297***	0.00445*	-0.00837***	0.00263***
	(1.42)	(-3.11)	(5.56)	(1.65)	(-3.02)	(4.47)
Tenure	-0.00138	0.00209	-0.000156	-0.000485	0.00211	-0.0000467
	(-1.06)	(1.57)	(-0.58)	(-0.39)	(1.57)	(-0.16)
Gradsec	0.0309	-0.110**	0.0229***	0.0280	-0.110**	0.0205**
	(0.66)	(-2.48)	(2.78)	(0.65)	(-2.46)	(2.28)
Level of education (ref. Diplomatura)						
Licenciatura	-0.218**	1.176***	0.0918***	-0.173*	1.177***	0.111***
	(-2.11)	(11.21)	(4.53)	(-1.76)	(11.07)	(3.65)
Doctorate	-0.0206	2.849***	0.00251	0.153	2.844***	0.00935
	(-0.07)	(11.97)	(0.05)	(0.54)	(11.95)	(0.12)
Educational Program Attributes						
Prestige	0.271***	-0.147	0.0352**	0.267***	-0.144	0.0155
	(2.78)	(-1.60)	(2.07)	(2.94)	(-1.54)	(0.84)
Vocational	-0.0454	0.0676	0.0210	0.0136	0.0655	0.0278
	(-0.42)	(0.67)	(1.11)	(0.13)	(0.65)	(1.37)
Demanding	0.0514	-0.0363	-0.0275*	0.0695	-0.0359	-0.0314*

	(0.58)	(-0.42)	(-1.69)	(0.83)	(-0.42)	(-1.79)
Broad	-0.0524	-0.0487	0.00684	-0.0342	-0.0497	0.0118
	(-0.60)	(-0.59)	(0.44)	(-0.42)	(-0.60)	(0.70)
Freedom	-0.0540	-0.135	0.00804	-0.0458	-0.135	0.0131
	(-0.64)	(-1.64)	(0.52)	(-0.58)	(-1.63)	(0.78)
Empfml	-0.0833	0.0770	0.00654	-0.0999	0.0757	0.0115
	(-1.00)	(0.98)	(0.44)	(-1.30)	(0.96)	(0.72)
Fields of study (ref. Humanities)						
Education	0.0772	0.666***	0.0764**	-0.0294	0.660***	0.0639
	(0.40)	(3.19)	(2.04)	(-0.16)	(3.12)	(1.53)
Social Sc.	0.749***	0.378**	0.00695	0.647***	0.382**	-0.0763**
	(4.87)	(2.21)	(0.23)	(4.25)	(2.11)	(-2.25)
Science	-0.0962	0.552***	-0.0104	-0.0599	0.549***	-0.0161
	(-0.56)	(2.90)	(-0.31)	(-0.35)	(2.88)	(-0.43)
Engineer.	0.450**	0.772***	0.0715*	0.361*	0.772***	0.0271
	(2.20)	(3.79)	(1.94)	(1.84)	(3.74)	(0.66)
Agric. & Vet	-0.0245	0.395	-0.0117	0.0322	0.396	-0.0274
	(-0.10)	(1.41)	(-0.24)	(0.13)	(1.41)	(-0.53)
Health	0.544**	0.330	-0.0367	0.456**	0.333	-0.0886*
	(2.31)	(1.40)	(-0.88)	(2.04)	(1.40)	(-1.94)
Job characteristics						
Public	0.0944	-0.300**	0.111***	0.126	-0.298**	0.107***
	(0.72)	(-2.42)	(4.95)	(1.03)	(-2.40)	(4.39)
Firmsize_2	-0.162	0.126	0.130***	-0.162	0.123	0.145***
	(-1.04)	(0.86)	(4.59)	(-1.12)	(0.84)	(4.73)
Firmsize_3	0.0212	-0.122	0.177***	-0.0443	-0.125	0.180***
	(0.12)	(-0.72)	(5.52)	(-0.26)	(-0.73)	(5.24)
Firmsize_4	-0.235	-0.0481	0.166***	-0.251	-0.0507	0.192***
	(-1.26)	(-0.27)	(5.01)	(-1.47)	(-0.28)	(5.35)
Firmsize_5	-0.142	0.0433	0.170***	-0.149	0.0403	0.185***
	(-0.80)	(0.26)	(5.36)	(-0.90)	(0.24)	(5.40)
Firmsize_6	-0.172	0.151	0.263***	-0.193	0.148	0.281***
	(-1.10)	(1.03)	(9.41)	(-1.36)	(1.00)	(9.32)
Firmtype	-0.145	-0.445***	-0.0574***	-0.111	-0.446***	-0.0451**
	(-1.48)	(-4.70)	(-3.27)	(-1.23)	(-4.70)	(-2.28)
Numemp	-0.0106	0.0182*	-0.00227	-0.0130	0.0182*	-0.00162
	(-0.96)	(1.82)	(-1.00)	(-1.39)	(1.80)	(-0.66)

Supervise			0.0577***			0.0650***
			(3.72)			(4.29)
_cons	1.365*	0.0237	1.663***	1.090	0.0487	1.286***
	(1.86)	(0.04)	(11.74)	(1.59)	(0.07)	(8.04)
Other Controls						
Occupation	Yes	Yes	Yes	Yes	Yes	Yes
Industry	yes	yes	yes	yes	yes	yes
N	2027	2008	1943		2050	
chi2/R2	561.5	900.0	0.373		1749.6	
Sig_3	0.3324***					
Rho_12	0.0188					
Rho_13	-0.721***					
Rho_23	-0.0351					

t statistics in parentheses, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 4 (complete results). Male sample.

	A			B		
	Hmatch (probit)	Overed (probit)	Lnghwage (OLS)	Hmatch	overed	lnghwage
				Simultaneous equations (cmp)		
Mismatch Dimensions						
Hmatch		-1.239***	0.0678*		-0.933**	0.587***
		(-6.67)	(1.69)		(-2.30)	(9.18)
Overed			-0.116***			-0.174*
			(-3.42)			(-1.87)
Undered			-0.0888*			-0.0566
			(-1.92)			(-1.27)
Individual characteristics						
Immigrant	0.461	-0.104	0.197**	0.321	-0.138	0.154
	(0.74)	(-0.19)	(1.97)	(0.78)	(-0.26)	(1.47)
Married	-0.0241	0.0745	0.106***	-0.109	0.0849	0.108***
	(-0.16)	(0.55)	(4.39)	(-0.86)	(0.63)	(4.12)
Experience	0.00264	-0.0132***	0.00436***	0.00275	-0.0139***	0.00372***
	(0.53)	(-2.96)	(5.07)	(0.64)	(-3.09)	(3.83)
Tenure	0.00261	0.00252	-0.000272	0.00319*	0.00236	-0.000382
	(1.25)	(1.31)	(-0.74)	(1.70)	(1.24)	(-0.96)
Gradsec	0.141*	-0.176**	0.0328**	0.123*	-0.177**	0.0182
	(1.68)	(-2.27)	(2.43)	(1.72)	(-2.28)	(1.22)
Level of education (ref. Diplomatura)						
Licenciatura	-0.0176	1.192***	0.0831**	-0.0774	1.178***	0.120***
	(-0.10)	(6.72)	(2.55)	(-0.49)	(6.61)	(2.91)
Doctorate	-0.0517	3.072***	-0.0266	0.560	3.080***	0.0576
	(-0.10)	(7.36)	(-0.32)	(1.11)	(7.29)	(0.53)
Educational Program Attributes						
Prestige	0.348**	-0.348**	0.0636**	0.348**	-0.371**	0.0308
	(2.13)	(-2.33)	(2.40)	(2.48)	(-2.46)	(1.05)
Vocational	-0.224	0.104	0.0165	-0.132	0.133	0.0428
	(-1.24)	(0.62)	(0.55)	(-0.85)	(0.79)	(1.31)
Demanding	0.00146	0.0117	-0.0250	0.00279	0.0161	-0.0217
	(0.01)	(0.08)	(-0.91)	(0.02)	(0.11)	(-0.72)
Broad	-0.225	-0.101	0.0120	-0.258**	-0.0769	0.0294
	(-1.47)	(-0.73)	(0.48)	(-1.97)	(-0.55)	(1.08)
Freedom	-0.165	-0.128	0.0223	-0.141	-0.117	0.0357

	(-1.14)	(-0.91)	(0.88)	(-1.11)	(-0.83)	(1.31)
Empfml	-0.0569	0.292**	-0.0168	-0.0816	0.294**	-0.0122
	(-0.39)	(2.16)	(-0.69)	(-0.65)	(2.19)	(-0.46)
Fields of study (ref. Humanities)						
Education	-0.253	0.426	0.107	-0.447	0.477	0.152*
	(-0.55)	(0.97)	(1.42)	(-1.12)	(1.09)	(1.84)
Social Sc.	0.262	0.412	0.0404	0.364	0.371	0.00370
	(0.75)	(1.27)	(0.72)	(1.14)	(1.13)	(0.06)
Science	-0.730**	0.439	0.0115	-0.379	0.455	0.0761
	(-2.00)	(1.31)	(0.20)	(-1.15)	(1.36)	(1.20)
Engineer.	-0.196	0.545	0.0817	-0.126	0.541	0.101
	(-0.51)	(1.61)	(1.40)	(-0.38)	(1.61)	(1.58)
Agric. & Vet	-0.0520	-0.367	0.0664	0.344	-0.402	0.0483
	(-0.11)	(-0.74)	(0.87)	(0.74)	(-0.80)	(0.58)
Health	0.963	0.565	-0.0273	0.670	0.489	-0.0834
	(1.36)	(0.88)	(-0.28)	(1.13)	(0.75)	(-0.80)
Job characteristics						
Public	0.0734	-0.184	0.00251	0.403*	-0.189	0.00276
	(0.27)	(-0.74)	(0.06)	(1.81)	(-0.76)	(0.06)
Firmsize_2	0.0884	-0.288	0.0729	0.165	-0.288	0.0636
	(0.31)	(-1.05)	(1.44)	(0.66)	(-1.05)	(1.14)
Firmsize_3	1.010**	-0.235	0.138**	0.767**	-0.251	0.0842
	(2.41)	(-0.78)	(2.49)	(2.27)	(-0.83)	(1.39)
Firmsize_4	0.0682	-0.112	0.195***	0.146	-0.103	0.198***
	(0.20)	(-0.35)	(3.38)	(0.50)	(-0.33)	(3.13)
Firmsize_5	0.232	-0.352	0.107*	0.380	-0.359	0.0845
	(0.72)	(-1.19)	(1.95)	(1.36)	(-1.22)	(1.41)
Firmsize_6	-0.0734	-0.213	0.220***	-0.0617	-0.207	0.228***
	(-0.26)	(-0.82)	(4.45)	(-0.25)	(-0.79)	(4.22)
Firmtype	0.0202	-0.354**	-0.0724**	0.0274	-0.354**	-0.0798**
	(0.12)	(-2.18)	(-2.53)	(0.18)	(-2.18)	(-2.55)
Numemp	-0.00631	0.00984	0.000857	-0.0123	0.0106	0.00117
	(-0.41)	(0.67)	(0.29)	(-1.05)	(0.74)	(0.36)
Supervise			0.0637**			0.0806***
			(2.57)			(3.46)
Other Controls						

Occupation	Yes	Yes	Yes	Yes	Yes	Yes
Industry	yes	yes	yes	yes	yes	yes
_cons	1.207	0.393	1.711***	1.009	0.137	1.294***
	(1.27)	(0.43)	(9.19)	(1.18)	(0.15)	(6.19)
N	789	794	744	805		
chi2/R2	146.8	245.7	0.370	676.4		
Sig_3	0.3328***					
Rho_12	-0.1676					
Rho_13	-0.8927***					
Rho_23	0.1662					

t statistics in parentheses, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 5 (complete results). Female sample.

	A			B		
	Hmatch (probit)	Overed (probit)	Lnghwage (OLS)	Hmatch	overed	lnghwage
				Simultaneous equations (cmp)		
		Mismatch Dimensions				
Hmatch		-1.342***	0.0985***		-1.670***	-0.0857
		(-9.77)	(3.30)		(-4.10)	(-0.85)
Overed			-0.111***			-0.0865
			(-4.38)			(-0.75)
Undered			0.00366			-0.0000258
			(0.09)			(-0.00)
		Individual characteristics				
Immigrant	-0.492	-0.918	0.278***	-0.578	-0.945	0.272***
	(-1.33)	(-1.44)	(3.68)	(-1.61)	(-1.47)	(3.55)
Married	-0.137	0.0300	0.0598***	-0.122	0.0213	0.0557***
	(-1.28)	(0.30)	(3.20)	(-1.16)	(0.21)	(2.97)
Experience	0.00568	-0.00559	0.00209***	0.00451	-0.00509	0.00230***
	(1.45)	(-1.55)	(3.01)	(1.13)	(-1.39)	(3.26)
Tenure	-0.00455**	0.00172	0.0000829	-0.00486**	0.00146	-0.0000704
	(-2.32)	(0.87)	(0.20)	(-2.47)	(0.72)	(-0.17)
Gradsec	0.0194	-0.0706	0.0216**	0.0170	-0.0669	0.0225**
	(0.32)	(-1.24)	(2.03)	(0.28)	(-1.16)	(2.08)
		Level of education (ref. Diplomatura)				
Licenciatura	-0.320**	1.233***	0.0906***	-0.276**	1.206***	0.0704*
	(-2.29)	(8.97)	(3.41)	(-1.96)	(8.34)	(1.69)
Doctorate	0.0667	2.851***	-0.00278	0.140	2.819***	-0.0244
	(0.16)	(9.32)	(-0.04)	(0.34)	(9.11)	(-0.25)
		Educational Program Attributes				
Prestige	0.195	-0.0510	0.0147	0.176	-0.0383	0.0217
	(1.49)	(-0.42)	(0.65)	(1.36)	(-0.31)	(0.95)
Vocational	0.0861	0.0450	0.0348	0.110	0.0483	0.0369
	(0.60)	(0.34)	(1.40)	(0.77)	(0.37)	(1.49)
Demanding	0.0215	-0.0537	-0.0320	-0.00379	-0.0491	-0.0292
	(0.19)	(-0.49)	(-1.56)	(-0.03)	(-0.45)	(-1.43)
Broad	0.0426	-0.0267	-0.00179	0.0398	-0.0212	-0.000549

	(0.38)	(-0.25)	(-0.09)	(0.35)	(-0.20)	(-0.03)
Freedom	-0.00809	-0.140	0.00365	0.00353	-0.140	0.00448
	(-0.07)	(-1.33)	(0.18)	(0.03)	(-1.34)	(0.22)
Empfml	-0.115	-0.0182	0.0107	-0.0784	-0.0268	0.00720
	(-1.07)	(-0.18)	(0.56)	(-0.73)	(-0.26)	(0.38)
Fields of study (ref. Humanities)						
Education	0.164	0.820***	0.0654	0.221	0.839***	0.0728
	(0.70)	(3.23)	(1.45)	(0.95)	(3.28)	(1.48)
Social Sc.	0.961***	0.443**	0.0117	1.018***	0.531**	0.0587
	(5.08)	(2.09)	(0.31)	(5.52)	(2.26)	(1.33)
Science	0.104	0.665***	0.0111	0.187	0.679***	0.0219
	(0.46)	(2.67)	(0.25)	(0.83)	(2.70)	(0.48)
Engineer.	0.820***	1.142***	0.0941*	0.870***	1.198***	0.125**
	(2.67)	(4.08)	(1.82)	(2.91)	(4.19)	(2.10)
Agric. & Vet	-0.150	0.828**	-0.0581	-0.125	0.828**	-0.0569
	(-0.46)	(2.28)	(-0.90)	(-0.40)	(2.30)	(-0.86)
Health	0.526*	0.430	-0.0421	0.603**	0.483*	-0.0173
	(1.91)	(1.57)	(-0.86)	(2.21)	(1.72)	(-0.33)
Job characteristics						
Public	0.135	-0.340**	0.163***	0.155	-0.328**	0.168***
	(0.84)	(-2.28)	(6.03)	(0.96)	(-2.19)	(6.03)
Firmsize_2	-0.215	0.353*	0.151***	-0.198	0.325*	0.139***
	(-1.08)	(1.96)	(4.37)	(-1.01)	(1.77)	(3.86)
Firmsize_3	-0.307	-0.130	0.192***	-0.253	-0.153	0.181***
	(-1.37)	(-0.59)	(4.77)	(-1.14)	(-0.70)	(4.47)
Firmsize_4	-0.331	0.0488	0.134***	-0.276	0.0176	0.118***
	(-1.41)	(0.22)	(3.24)	(-1.18)	(0.08)	(2.83)
Firmsize_5	-0.289	0.270	0.210***	-0.265	0.240	0.197***
	(-1.28)	(1.27)	(5.25)	(-1.18)	(1.12)	(4.80)
Firmsize_6	-0.145	0.330*	0.271***	-0.118	0.308*	0.261***
	(-0.72)	(1.78)	(7.88)	(-0.59)	(1.65)	(7.38)
Firmtype	-0.276**	-0.532***	-0.0388*	-0.266**	-0.544***	-0.0455*
	(-2.20)	(-4.37)	(-1.72)	(-2.13)	(-4.47)	(-1.77)
Numemp	-0.0107	0.0247	-0.00533	-0.0107	0.0237	-0.00601*
	(-0.61)	(1.57)	(-1.49)	(-0.62)	(1.51)	(-1.66)
Supervise			0.0482**			0.0458**
			(2.38)			(2.32)

			Other Controls			
Occupation	Yes	Yes	Yes	Yes	Yes	Yes
Industry	yes	yes	yes	yes	yes	yes
_cons	6.825***	0.529	1.473***	0.885	0.763	1.598***
	(9.09)	(0.49)	(6.23)	(1.30)	(0.69)	(6.30)
N	1230	1212	1199		1245	
chi2/R2	400.4	604.4	0.392		1177.7	
Sig_3	0.3078***					
Rho_12	0.1888					
Rho_13	0.3650**					
Rho_23	-0.0257					

t statistics in parentheses, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$