Ca' Foscari University of Venice

Department of Economics

Working Paper

Roberto Roson
Emanuela Ghignoni

A Numerical Simulation of Educational Mismatch in the Italian Labor Market

ISSN: 1827-3580
No. 15/WP/2023

Working Papers
Department of Economics
Ca' Foscari University of Venice
No. 15/WP/2023
ISSN 1827-3580

A Numerical Simulation of Educational Mismatch in the Italian Labor Market

Roberto Roson
Ca' Foscari University of Venice; Loyola Andalusia University

Emanuela Ghignoni
University of Rome 'Sapienza'

Abstract
This paper presents a data set, associating education levels to occupations, and a methodology, which allow estimating how the distribution of the two variables could change, after some exogenous shock affecting the labor market. We assess some implications of the empirical finding that, in response to a weaker demand for labor, sufficiently educated workers would reallocate themselves into lower-ranked occupations, rather than getting unemployed. The exercise is conducted with Italian data, where 37 occupations and 10 education levels are considered. A counterfactual distribution is estimated, using a computable general equilibrium model to simulate the impact on the labor market of a trade disruption crisis with Russia.

Keywords
Skill Mismatch, Education, Overeducation, Unemployment, Labor Market, Computable General Equilibrium Models

JEL Codes
A20, C23, C68, C82, D58, E24, F16, I20, J21, J24, J62, J82

Address for correspondence:
Roberto Roson
Department of Economics
Ca' Foscari University of Venice
Cannaregio 873, Fondamenta S.Giobbe
30121 Venezia - Italy
e-mail: roson@unive.it

The Working Paper Series
is available only on line
(http://www.unive.it/pag/16882/)
For editorial correspondence, please contact:
wp.dse@unive.it

Department of Economics
Ca' Foscari University of Venice
Cannaregio 873, Fondamenta San Giobbe
30121 Venice Italy
Fax: ++39 041 2349210

A Numerical Simulation of Educational Mismatch in the Italian Labor Market

Roberto Roson[1]

Emanuela Ghignoni[2]

ABSTRACT

This paper presents a data set, associating education levels to occupations, and a methodology, which allow estimating how the distribution of the two variables could change, after some exogenous shock affecting the labor market. We assess some implications of the empirical finding that, in response to a weaker demand for labor, sufficiently educated workers would reallocate themselves into lower-ranked occupations, rather than getting unemployed. The exercise is conducted with Italian data, where 37 occupations and 10 education levels are considered. A counterfactual distribution is estimated, using a computable general equilibrium model to simulate the impact on the labor market of a trade disruption crisis with Russia.

KEYWORDS

Skill Mismatch, Education, Overeducation, Unemployment, Labor Market, Computable General Equilibrium Models.

JEL CODES

A20, C23, C68, C82, D58, E24, F16, I20, J21, J24, J62, J82.

1. Introduction

Education is key for personal and collective development. According to UNESCO (2010), One extra year of schooling increases an individual's earnings by up to 10%, and each additional year of schooling, on average for a country, raises its annual gross domestic product (GDP) growth by 0.37%.

However, whereas the importance of education in boosting economic growth, through productivity and human capital resources, has long been recognized, a much less considered aspect is its role of (relative) employment insurance. Indeed, higher educated people not only get easier access to the labor market, but often can avoid getting unemployed, by accepting lower ranked occupations. Vice versa, during positive economic cycles, they can step up to better jobs. Therefore, higher educated workers can be employed more easily, but they also get more stable occupations and income sources.

This issue relates to the matching between the structure of educational attainments in the work force, on one hand, and the structure of available job positions in the economy, on the other hand. There exists a literature, partly reviewed in the following section, which analyzes this match, but mainly from an individual, sectoral,

[1] Ca' Foscari University Venice; Loyola Andalusia University. E-mail: roson@unive.it.

[2] University of Rome 'Sapienza'. E-mail: emanuela.ghignoni@uniroma1.it.

microeconomic perspective. The macroeconomic implications of a good or bad employment/education match are much less known and investigated.

This paper is intended to contribute filling up this knowledge gap. Our approach is not theoretical, but rather focused on available tools for numerical simulation, and assessment, of policies and exogenous shocks. In this area, the applied macroeconomist can choose from a variety of alternative approaches: macro-econometric models, input-output models, computable and dynamic stochastic general equilibrium models, stock-flow consistent models, and others. Very often, these models can say something about the impact on employment levels, but they are silent on the distribution of occupations and, more generally, on the quality of employment.

We propose here a methodology which, by post-processing simulation results or other estimated variations in employment, calculates a counterfactual joint distribution of employment categories by level of educational attainment. Mathematically, this alternative distribution is obtained as a solution of a constrained optimization problem, where a baseline employment by education (EE) matrix is taken as a starting point.

We illustrate the process by taking, as an example, the estimated employment impact of a trade disruption crisis with Russia, obtained from a standard, global CGE model. According to the model output, employment should decrease for all occupations and all education levels. Interestingly, however, after applying our method to get the counterfactual EE matrix, we found that employment increases for specific combinations of (higher level) education and job position. This means that the shock on the labor market generates a redistribution of employment flows, where indeed some workers are reallocated to lower ranked (less paid, less reputed) jobs.

The paper is structured as follows. First, we overview the available literature, dealing with skill mismatch, overeducation, and related topics. We then introduce an employment by education matrix for Italy, considering 37 job categories and 10 education levels. Section four provides some estimates of the education-specific elasticity of unemployment. Section five describes the methodology for the calculation of the alternative EE matrix, and section six illustrates the methodology with an example. We conclude by critically discuss our approach, and by providing some final remarks.

2. Related Literature

The economic literature has long recognized that skill mismatch, in its peculiar form of overeducation, can represent a way out of unemployment for highly skilled workers in response to a negative demand shock even though, in most of the literature about the causes and the effects of skills mismatch, it has often been regarded as a waste of resources for society as a whole, in terms of public investment in education that is not paying-off.

Along these lines, most authors derive overeducation from an excess of highly educated labor supply, when the aggregate amount of highly educated workers exceeds the number of available jobs requiring high levels of education. In this case it is said that demand ‘loses the race’ with the supply of human capital (Caroleo and Pastore, 2015).

Other explanations do not imply aggregate demand-supply imbalances and highlight labor market imperfections as possible causes of overeducation. Mobility costs and family ties can impact individual spatial flexibility and reduce the number of job opportunities in the local labor market of residence (Büchel and Van Ham, 2003). The impact of the quality of education on overeducation has been tested by Di Pietro and Cutillo (2006) and Ordine and Rose (2009). The authors find a causal relationship between the teaching and research quality of the attended university and the risk of being overeducated.

A long literature, however, has made it clear that the surplus years of education, with respect to occupation requirement, increase individual productivity and are rewarded (even though less than the required ones) in the labor market (McGuinness, 2006; Leuven and Oosterbeek, 2011). More recent research also warns against the interpretation of overeducation as evidence of wasteful investment, confirming that while some individuals have higher qualifications than those required to undertake their job, their additional human capital is nonetheless rewarded by the employers (Johnes, 2019). This seems to imply that employers are not opposed to hiring workers who are over-qualified for their job tasks. Indeed, overeducation has been proven to carry a less negative stigma than being unemployed in employers’ view (Baert and Verhaest, 2019).

In this line, overeducation can be seen as a means of avoiding unemployment and/or reducing unemployment duration for highly skilled workers, as high educated people may be employed in low-skilled sectors and occupations, but not vice versa.

According to some authors, however, labor markets flexibility would play a relevant role on the trade-off between overeducation and unemployment.

Some research analyzed the impact of Employment Protection Legislation (EPL) on overeducation and found contrasting results. On the one hand, Di Pietro (2002) shows that a high level of EPL raises the costs of adjusting the workforce and, therefore, is suspected of discouraging firms from adopting new technologies, thereby increasing both overeducation and unemployment. On the other hand, Verhaest and Van der Velden (2013) do not confirm that the strictness of EPL plays a role in determining overeducation. Instead, they find that the business cycle, as measured by the output gap, is relevant. Indeed, a recession tends to worsen the prospects for workers to be in a well-matched position and, therefore, it may raise the risk of overeducation for high-skill workers who are able to keep their job. This is not at odds with the findings of Quintini (2011) on a large panel of European countries, which show that workers losing their job during a recession face a higher probability of being overeducated in their subsequent employment, if they have a sufficiently high level of education.

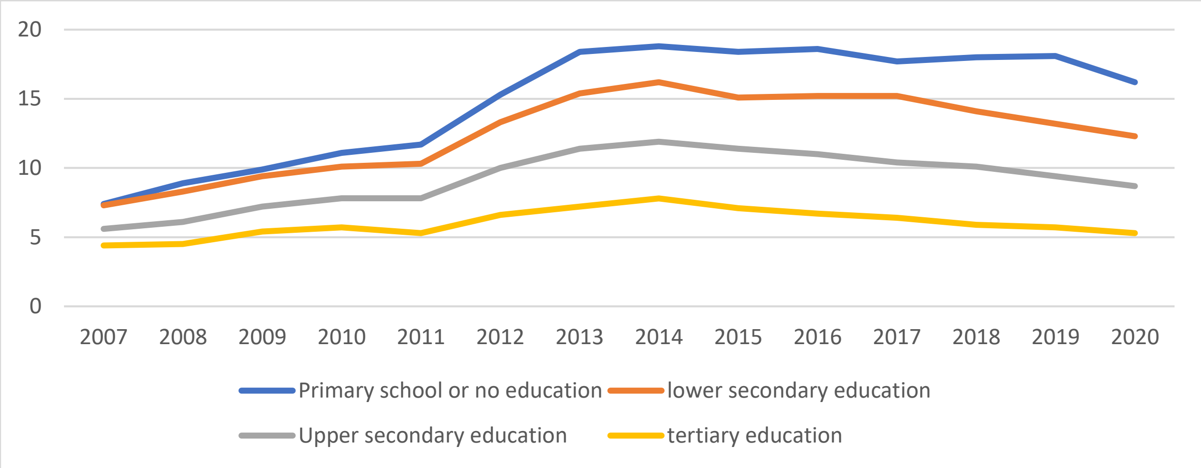
A slightly different strand of literature points to the role of wage flexibility. A well-known paper by Card, Kramarz and Lemieux (1999) argued that the contrast between the United States and Europe illustrates a fundamental trade-off between wage inequality and employment in response to declining demand for less-skilled labor. The authors suggest that when a negative shock affects the relative demand for less-skilled workers, in the US, where labor markets and wages are flexible, the adverse shock will affect the relative wages of low-skilled workers more than low-skill unemployment, whereas in Europe, where wages are relatively inflexible, due to minimum wages, union wage setting, and generous unemployment benefits, the same shocks will mainly affect the relative employment of low-skilled workers. This would imply that in less flexible countries, a negative demand shock would have the effect of excluding low-skilled workers from the labor market, whereas highly skilled workers would not lose their jobs but would risk a downgrade of their tasks.

Italy is a puzzling example in this context. Labor market and wage flexibility have increased in Italy from the late '90s, there is no Statutory Minimum Wage and the strength of Unions and collective bargaining have been decreasing in the last decades. Nevertheless, during the 2008-14 economic and financial crisis the unemployment rate of low-skilled workers has increased much faster than the unemployment rate of high-skilled ones (Figure 1). It can therefore be expected that highly educated workers were able to avoid unemployment even by reassigning themselves to low-skilled jobs, whereas low-educated workers simply lost their jobs.

Indeed, some authors point out that overeducation can arise even with complete flexibility of wages, just because of the decline of the relative wage of skilled labor (Sattinger, 1993; McGuinness, 2006). In practice, when the relative wages of highly skilled workers decrease, educated workers (may be forced or) may prefer to accept low-skilled jobs, if skilled jobs involve more effort or lower employment security. In this case, it is likely that in economies where the pace of technological change is slower than the expansion of skilled labor supply (as it is the case in Italy), a low wage premium is observed for skilled workers, as well as a large share of overeducated workers. On this basis, even if overeducation results in a loss of productivity due to an incomplete use of human capital, it can be considered as a 'voluntary' status, conditioned by the available jobs in the labor market, from an individual point of view.

To confirm this intuition, McGuinness and Sloane (2011) find that overeducated workers suffer a heavy wage penalty but tend to report a less severe loss of job satisfaction. This rather surprising finding can be explained by assuming that workers trade off earnings for other job attributes, such as an easier work-life balance or higher employment security, according to their preferences. This is also consistent with the assumption of Gottshalk and Hansen (2003), who claim that workers choose between a good or a bad match according to relative wage and other job characteristics.

Figure 1 – Unemployment rates by education level, Italy, 2007-2020



Source: ISTAT, various years

In all these cases, overeducation would be a deliberate and rational response by highly educated workers to the job shortage, easily accepted by employers.

3. The Employment by Education Matrix

The original Employment by Education (EE) Matrix has been drawn from ISTAT Labor Force microdata for 2014.

The classification of occupation used by ISTAT in the 2014 Labor Force Survey is CP2011 (in place until 31/12/2022) at 2 digits[3]. The structure of the CP2011 is based on the logic of the ISCO (International Standard Classification of Occupations), with which it is cross-linkable, and consist of 37 occupations.

Different levels of education are classified according to the International Standard Classification of Education[4] (ISCED 2011) that is widely used to compare education systems around the world.

This leads to the 37x10 matrix set out in Table 1.

Not surprisingly, the bulk of Italian employees concentrate on low and middle levels of education, with 32% having at most a lower secondary education level and 79% having attained at most an upper secondary level. Even so, only 20% of employment is concentrated in elementary and non-qualified occupations (CP2011 codes 7 and 8, see Table A2). It should also be noted that 39% of employees in non-qualified occupations have a level of education at or above upper secondary level.

The EE matrix reported in Table 1 is the base for the numerical simulations described in Section 5.

[3] See Classificazione delle professioni (istat.it) and Table A1 for details.

[4] See international-standard-classification-of-education-isced-2011-en.pdf (unesco.org) and Table A2 for details.

Table 1 – EE Matrix, Italy, 2014

	EARLYED	PRIMAED	LOW2EDU	UP2EDUL	UP2EDUH	POST2ED	SHORT3E	BACHELR	MASTERL	DOCTORL	*TOT*
PUBEXEC	24	77	1933	790	12946	661		2627	61995	1802	**82855**
ADLARGE		3425	10161	3365	58548	432		2690	75016	1570	**155207**
ADSMALL	918	17083	94793	19383	160533	1473		5715	49702	260	**349862**
SPMATHC		13	2102	902	66047	3758		13967	108995	7056	**202840**
ENGARCH		143	188	570	7148	471		14805	296943	8594	**328861**
SPLIFESC					2579	115		1729	115187	5183	**124792**
HEALTHS	173		149	259	2495	69		3264	279948	5761	**292117**
SPHUMAN	141	659	15336	8627	209681	6121		36247	661959	13111	**951883**
SPEDUCR	291	181	4360	15089	375469	11640		69389	658483	60179	**1195080**
TCSCIEN	935	4892	107445	55653	801553	22213		41130	159103	4856	**1197780**
TCHEALT	255	449	35384	69328	202794	32423		275966	86082	1250	**703931**
TCADMIN	240	4215	138209	76798	1020454	19768		53631	332457	2686	**1648458**
TCPUBLS	229	651	32436	16966	192427	7970		50857	94168	748	**396453**
EMSECRO	41	5144	139488	87448	718559	15389		39542	169474	2197	**1158424**
EMMONCA		1859	50971	26082	290851	5635		22062	66368	245	**464072**
EMADMIN	625	3958	130932	58932	389578	9208		19235	54050	1359	**667877**
EMDOCUM	262	1276	44944	18780	133810	3155		3542	25524	83	**231378**
SKCOMMA	6318	52952	602376	147856	871289	15016		41159	77424	165	**1863791**
SKHOSPF	10784	42890	431167	119882	420918	6553		20323	36005	1332	**1047428**
SKHEALS		964	78311	29657	72054	10611	70	4665	6760		**203093**
SKPERSS	16092	35817	386151	140638	427740	11955	35	27202	58421	456	**1104507**
SWEXCON	14234	85663	594259	126744	256352	981		841	9017		**1088090**
SWELECT	4367	54749	519992	177689	327448	3225	134	3535	4810		**1095948**
SWPREPR	1139	5182	87724	28127	61588	760		1049	3404		**188972**
SWAGRIC	8730	76913	208858	49166	118016	1299		3119	13595	478	**480177**
SWOTHIN	6785	51419	338295	61110	140602	1164		2308	5986		**607669**
OPINDPL	1700	13987	157243	41721	99498	1737		1028	4167	86	**321167**
OPFIXMA	2757	31716	331805	100968	190274	1943		2363	6741		**668567**
OPFMAGR	333	3241	36027	11065	27725	573		1917	567		**81450**
OPVEHIC	4801	50372	414501	76702	186470	2413		2986	5909		**744153**
UNSERVC	23779	107711	761670	143939	355073	3630		8647	27197	217	**1431864**
UNDOMEC	15771	37502	207691	49577	149424	1418		8605	29088	132	**499208**
UNAGRIC	14205	50112	180424	20789	68912	154		776	2347	147	**337865**
UNMANUF	4297	16237	101394	23273	41117	286		689	1565		**188859**
MILIOFF			2345	467	16723			2497	6939		**28971**
MILISER		243	18172	6090	56379	178		10022	6851		**97935**
MILIPER			28676	4415	72640	106		2273	2243		**110352**
TOT	**140227**	**761695**	**6295917**	**1818848**	**8605712**	**204501**	**240**	**802400**	**3604487**	**119953**	*22341934*

Source: Elaborations on ISTAT Labor Force microdata.

4. The Education Elasticity of (Un)employment

When total employment varies, how does the educational mix of employed workers vary? Or: how do unemployment rates vary when they refer to education levels? Intuitively, higher education should provide more employment opportunities, so that educated workers should get employed more easily, as well as stay employed during economic downturns.

In this section, we look for some empirical evidence. To this purpose, we use unemployment rates data from EUROSTAT (2023), from where we derived an unbalanced panel dataset with 20 time periods, 22 countries, and three education levels (low, middle, high), for a total of 423 observations.

We regressed, with fixed effects, the annual change in education-specific unemployment rates (*d_UR_low/mid/high*), against the change in total unemployment (*d_UR_tot*). Results are shown here below:

Table 2 – Changes in total unemployment and changes in education-specific unemployment rates

Dependent variable: d_UR_low

	Coeff.	*S.E.*	*t*	*p-value*
const	0.167	0.058	2.881	0.004
d_UR_tot	1.515	0.038	39.460	0.000

Dependent variable: d_UR_mid

	Coeff.	*S.E.*	*t*	*p-value*
const	0.020	0.018	1.088	0.277
d_UR_tot	1.130	0.012	92.820	0.000

Dependent variable: d_UR_high

	Coeff.	*S.E.*	*t*	*p-value*
const	0.046	0.023	1.959	0.051
d_UR_tot	0.531	0.015	34.280	0.000

Source: Elaborations on EUROSTAT data

Our findings reveal that, when total unemployment rate increases by 1%, the corresponding unemployment rate of highly educated workers only grows by 0.53%, whereas the variations for lower education levels are much larger. We also detect a positive parameter for the constant (*const*), which is statistically significant only for low education, though. This can be interpreted as an historical trend, in the European countries, of increment in the unemployment rate of poorly educated workers.

The results are consistent with those of Murphy and Topel (1997). In their work, they regressed employment levels against average wages in remuneration percentiles, finding that labor elasticity decreases at higher wages. This implies that, if aggregate labor demand declines, the response in the lower categories will be a significant reduction of employment (rather than wage), whereas the opposite would occur at the top end. This is indeed coherent with our findings as, of course, years of education and wages are generally correlated, as discussed in the literature on returns to education (see, e.g., Carneiro, Heckman, and Vytlacil, 2011).

5. Simulating Overeducation with a General Equilibrium Model

We present in this section a methodology for the estimation of a counterfactual EE matrix, consistent with some results of a numerical simulation experiment, which provides some alternative figures for total employment. To this end, we employ a standard, global computable general equilibrium model (GTAP). However, we could well have utilized, as an input for our calculations, estimates from other models.

The GTAP standard model (Corong et al., 2017) considers five different categories of workers, namely:

- Technicians and associated professionals
- Clerks
- Service and shop workers

- Officials and managers
- Agricultural and low-skilled

Structural parameters of the model are calibrated on data from a social accounting matrix (SAM) of the world economy (Aguiar et al. 2019). Simulations are comparative-static: changes in parameters and exogenous variables are assumed, then an alternative general equilibrium state is calculated. Different closure hypotheses can be applied. Specifically, here it assumed that, in the various regional labor markets, real wages are exogenous and fixed, whereas employment levels (by labor category) are endogenous.

Six regions are considered (Italy, Rest of EU, USA, China, Russia, Rest of the World), and ten industries, including "Extraction". The latter comprehends coal, oil, gas, and minerals. The simulation exercise is based on a hypothetic reduction by 90% of imports of extraction products from Russia to EU and USA.[5]

We do not aim at illustrating here the complete set of results from the numerical exercise.[6] We just present the computed variations in employment levels for Italy, and for the five workers categories:

Table 3 – Estimated variations in employment levels

Technicians and associated professionals	-0.83%
Clerks	-0.86%
Service and shop workers	-0.78%
Officials and managers	-0.80%
Agricultural and low-skilled	-0.94%

We use these results to get new row and column totals in the Italian EE matrix. This task is undertaken in two stages. First, the percentage variations in Table 2 are applied to the baseline row totals. Since there are 37 worker categories in the matrix, but only five in the GTAP model, the same (relative) variation is applied to multiple items. In some cases, when it is not possible to directly associate two groups, a weighted average is used instead.[7]

The sum of all row totals gives, of course, total employment. Therefore, we also know how much total employment has supposedly changed from the initial level. Using the econometric estimates described in the previous section (Table 2), it is then possible to gauge the variation in education-specific level of employment. Of course, this is possible only for the three classes considered (low, middle, and highly educated), so here another one-to-many correspondence is applied.[8]

With new row and column marginals, the baseline EE matrix can be interpreted as an a-priori distribution in a maximum likelihood estimation (Roson, 2022). In other words, the problem is modifying the original matrix, such that the new totals are respected. Those who have worked with estimation of blocks in input-output or social accounting matrices may notice the strong resemblance of the problem here with the estimation of sub-matrices trough RAS or entropy-maximizing techniques (see, e.g., Robinson, Cattaneo, and El-Said, 2001).

One way to formulate the problem is defining it as a minimum distance optimization problem:

$$\min \quad \sum_{i,j} \left(\frac{x_{i,j}}{\widetilde{x_{i,j}}} - 1 \right)^2 \tag{1}$$

[5] Technically, some trade flow volumes are imposed, whereas the model determines the corresponding level of non-tariff trade barriers.

[6] They are available on request. Also, the interested reader may find articles in the literature, where the issue is investigated, by means of CGE models (e.g., Chepeliev., Hertel, and van der Mensbrugghe, 2022).

[7] Mapping information is available from the authors on request.

[8] Here the association of our ten education levels to the three aggregated categories is straightforward. When new column totals are computed, some rescaling may be necessary, to ensure that the sum of column totals gives the same employment level obtained by summing row totals.

s.t.

$$\sum_j x_{i,j} = \overline{X_{i.}} \qquad (2)$$

$$\sum_i x_{i,j} = \overline{X_{.j}} \qquad (3)$$

Where: $x_{i,j}$ refer to cells in the new matrix, $\widetilde{x_{i,j}}$ those in the old matrix, $\overline{X_{i.}}$ are constrained row totals, $\overline{X_{.j}}$ are constrained column totals.

The measure of distance utilized in (1) is the square of the relative (percentage change) variation. This kind of measure allows preserving as much as possible the relative proportions between flows in the matrix.

If the objective function is interpreted in terms of disutility, or adjustment costs, then its choice also carries an implicit behavioral assumption. Indeed, this would amount to assuming that joining a work category, where there are already other workers with the same education level, would be less unpleasant than joining a group without a sizeable presence of peers.

The constrained minimization problem (1) gives raise to three sets of first order conditions. In addition to (2) and (3):

$$\left(\frac{x_{i,j}}{\widetilde{x_{i,j}}} - 1\right)\frac{1}{\widetilde{x_{i,j}}} + \lambda_{i.} + \lambda_{.j} = 0 \qquad (4)$$

Which can also be written as:

$$\frac{x_{i,j}}{\widetilde{x_{i,j}}} + \widetilde{x_{i,j}}\lambda_{i.} + \widetilde{x_{i,j}}\lambda_{.j} = 1$$

Taken together, (2) (3) and (4) defines a linear system, whose solution identifies the cells in the new matrix $x_{i,j}$, as well as the Lagrange multipliers λ. Being a linear system, it can also be expressed as a single matrix equation, like:[9]

$$\begin{bmatrix} \widetilde{x_{1,1}}^{-1} & 0 & 0 & 0 & \widetilde{x_{1,1}} & 0 & \widetilde{x_{1,1}} \\ 0 & \widetilde{x_{1,2}}^{-1} & 0 & 0 & \widetilde{x_{1,2}} & 0 & 0 \\ 0 & 0 & \widetilde{x_{2,1}}^{-1} & 0 & 0 & \widetilde{x_{2,1}} & \widetilde{x_{2,1}} \\ 0 & 0 & 0 & \widetilde{x_{2,2}}^{-1} & 0 & \widetilde{x_{2,2}} & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_{1,1} \\ x_{1,2} \\ x_{2,1} \\ x_{2,2} \\ \lambda_{1.} \\ \lambda_{2.} \\ \lambda_{.1} \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ \overline{X_{1.}} \\ \overline{X_{2.}} \\ \overline{X_{.1}} \end{bmatrix} \qquad (5)$$

We apply this methodology for the estimation of a counterfactual EE matrix of Italy, using the matrix presented in Section 3 as a starting point (Table 1), and we present in Table 4 the absolute differences between cells in the two matrices.

One category, which is especially noticeable, is that of workers with a master or equivalent degree which, in Italy, broadly corresponds to that of "laureati" (graduates).[10] In our simulation exercise, it is estimated that, because of the trade crisis with Russia, there is an overall loss of 12,248 work units. The occupations where losses are highest are: "Specialists in humanities, social sciences, arts, and management" (SPHUMAN, -5,417),

[9] This would correspond to a 2x2 matrix with only one constraint at the first column.

[10] Individuals getting "laurea magistrale" or "laurea" before the higher education system reform.

“Engineers, architects, and related professions” (ENGARCH, -2,705), “Specialists in education and research” (SPEDUCR, -2,613).

Nonetheless, some occupations display an increase in employment. The three highest gains are in: “Technical professions in organization, administration, and financial and commercial activities” (TCADMIN, +2,979), “Employees responsible for secretarial and office machine functions” (EMSECRO, +684), “Technical professions in the scientific, engineering, and production fields” (TCSCIEN, +673).

Another interesting category is high upper secondary education, corresponding in Italy to “diplomati”. There is a total loss there of 71,968 workplaces, especially concentrated in TCADMIN and TCSCIEN (the same occupations where more graduates are absorbed, thereby replacing those with just a high school degree), but with some limited +777 and +227 compensatory increment in UNSERVC (“Unskilled professions in commerce and services”) and SWEXCON (“Artisans and specialized workers in the extractive industry, construction, and building maintenance”), respectively.

The redistribution of workers is a move towards professions with lower knowledge content and possibly retribution. To better appreciate this point, it possible to construct a professional “education intensity index”, by contrasting the shares of educated workers in the employment category and in the total, and making a weighted sum of the differences:

$$EII_p = \sum_i \left(SH_p^i - SH_{tot}^i\right) Y^i \tag{6}$$

Where *EII* is the index associated with the profession *p*, *SH* are the employment shares in the profession *p* and in the total of employed (*tot*), and *Y* is the number of years needed to achieve the education level *i.*

The education indexes of the three losing categories SPHUMAN, ENGARCH, and SPEDUCR are, respectively: 5.51, 7.19, 5.03. The education indexes of TCADMIN, EMSECRO, and TCSCIEN are, instead: 1.55, 0.97, 1.06. Therefore, the few professions where an increase in employment of educated workers is noted are also those where the educational intensity is lower. Analogously, for workers with upper secondary education, the occupational gains are observed in jobs with even lower ranking: UNSERVC (-2.92) and SWEXCON (-3.08). This would mean that some workers possessing a university or high school degree may avoid becoming unemployed, by accepting overeducation.

From a mathematical perspective, what drives the emergence of positive variations in some cells of the EE matrix? This is related to the initial distribution of flows in the baseline matrix, but much more so to the degree of asymmetry in the changes of row and column totals. To see this, notice from equation (4) that an increase in one cell content requires the sum of the corresponding row and column constraint Lagrange marginals being negative:

$$x_{i,j} - \widetilde{x_{i,j}} > 0 \xrightarrow{yields} \lambda_{i.} + \lambda_{.j} < 0 \tag{7}$$

In the simulation illustrated above, the row marginals are all positive. This depends on the imposed reductions estimated by the CGE model, being not too divergent. On the other hand, there are three negative column constraint marginals, associated with the education levels UP2EDUH, BACHELR, and MASTERL. This reflects the more differentiated response to the demand shock in the education categories. For each combination of row and column, then, positive variations emerge whenever the negative column marginal is, in absolute value, larger than the positive row marginal.

Table 4 – Estimated changes in the EE matrix

	EARLYED	PRIMAED	LOW2EDU	UP2EDUL	UP2EDUH	POST2ED	SHORT3E	BACHELR	MASTERL	DOCTORL	*TOT*
PUBEXEC	0	0	-1	0	-34	0	0	-1	-625	-1	**-663**
ADLARGE	0	-4	-19	-3	-548	0	0	-1	-666	-1	**-1242**
ADSMALL	-1	-77	-829	-54	-1769	-1	0	-2	-67	0	**-2800**
SPMATHC	0	0	-1	0	-574	-9	0	-22	-1069	-9	**-1684**
ENGARCH	0	0	0	0	-4	0	0	-12	-2705	-9	**-2729**
SPLIFESC	0	0	0	0	-1	0	0	0	-1030	-5	**-1036**
HEALTHS	0	0	0	0	0	0	0	-1	-2419	-4	**-2424**
SPHUMAN	0	0	-18	-10	-2371	-21	0	-45	-5417	-18	**-7900**
SPEDUCR	0	0	-1	-28	-6708	-76	0	-135	-2613	-358	**-9919**
TCSCIEN	-1	-5	-446	-276	-9631	-261	0	**8**	**673**	-2	**-9941**
TCHEALT	0	0	-93	-601	-2094	-594	0	-2390	-69	0	**-5842**
TCADMIN	0	-4	-730	-524	-15210	-207	0	**14**	**2979**	0	**-13681**
TCPUBLS	0	0	-101	-42	-2695	-37	0	-138	-277	0	**-3291**
EMSECRO	0	-6	-805	-703	-9170	-126	0	**3**	**684**	0	**-10123**
EMMONCA	0	-1	-175	-80	-3696	-18	0	-12	-9	0	**-3991**
EMADMIN	0	-4	-909	-360	-4457	-46	0	-4	**36**	0	**-5744**
EMDOCUM	0	0	-235	-59	-1656	-6	0	-1	-33	0	**-1990**
SKCOMMA	-37	-561	-9829	-1697	-2623	-117	0	**27**	**228**	0	**-14608**
SKHOSPF	-107	-376	-5801	-1175	-1341	-22	0	**5**	**44**	0	**-8774**
SKHEALS	0	0	-823	-162	-574	-70	0	-2	-3	0	**-1635**
SKPERSS	-239	-265	-4981	-1660	-1787	-75	0	**7**	**108**	0	**-8891**
SWEXCON	-185	-1418	-7122	-1136	**227**	0	0	0	**4**	0	**-9630**
SWELECT	-17	-592	-6623	-2369	-92	-5	-1	0	**1**	0	**-9698**
SWPREPR	-1	-8	-1072	-150	-439	0	0	0	-1	0	**-1672**
SWAGRIC	-72	-1326	-2227	-246	-381	-1	0	0	**3**	0	**-4249**
SWOTHIN	-43	-555	-4197	-326	-258	-1	0	0	**1**	0	**-5378**
OPINDPL	-3	-50	-2066	-233	-593	-2	0	0	0	0	**-2948**
OPFIXMA	-7	-214	-4401	-923	-591	-2	0	0	**1**	0	**-6137**
OPFMAGR	0	-5	-447	-48	-246	0	0	-1	0	0	**-748**
OPVEHIC	-21	-520	-5508	-486	-293	-3	0	0	**1**	0	**-6830**
UNSERVC	-514	-2211	-10133	-1409	**777**	-7	0	**2**	**35**	0	**-13459**
UNDOMEC	-238	-337	-2868	-288	-957	-1	0	-2	-1	0	**-4692**
UNAGRIC	-193	-599	-2134	-50	-199	0	0	0	0	0	**-3176**
UNMANUF	-19	-81	-1387	-100	-188	0	0	0	0	0	**-1776**
MILIOFF	0	0	-4	0	-196	0	0	-4	-32	0	**-236**
MILISER	0	0	-82	-11	-720	0	0	-21	-9	0	**-843**
MILIPER	0	0	-156	-5	-876	0	0	-1	-1	0	**-1038**
TOT	**-1698**	**-9221**	**-76224**	**-15213**	**-71968**	**-1710**	**-1**	**-2727**	**-12248**	**-407**	*-191418*

6. Discussion

How robust are our findings to alternative definitions of the matrix distance? Of course, objective functions different from (1) could have been adopted. For example, in a similar issue, Golan et al. (1994) use a cross entropy formulation to estimate the coefficients in an input-output table. Following their approach, the objective function (1) would be replaced here by:

$$\min \quad \sum_{i,j} x_{i,j} \cdot ln\frac{x_{i,j}}{\widetilde{x_{i,j}}} \tag{8}$$

However, first order conditions for the solution of the associated constrained optimization problem would give raise to a non-linear system of equations, for which could only be (approximately) solved numerically. Furthermore, our preferred formulation (1) keeps the relative proportions between cells in the matrix as much as possible close to the ones in the baseline matrix. This is because the solution tends, while respecting row and column constraints, to get the absolute percentage variation for the various cells as uniform as possible. We regard this property as desirable, also because it allows interpreting (1) as adjustment or disutility costs.

As a sensitivity test, we nonetheless did estimate the counterfactual EE matrix, using the objective function (8) above, and we found that this would not carry out any relevant qualitative difference in the results. For instance, the two distributions of occupations for workers with a master's degree only differ, on average, by 0.33%.[11]

Instead of setting a given objective function, could it be possible to get information about occupational mobility from historical data? As far as we know, EE matrices are not regularly produced and published, so that no time

[11] On absolute value. The standard deviation is 0.18%.

series of this kind are available. Whenever two matrices, for two (not too distant) different years would be available, then it could be possible to test various specifications of distance, or to estimate parameters in the same function. For example, instead of the squared differences in (1), it would be possible to place a generic power parameter, whose value would then be computed to get the best fitting, when contrasting the second estimated matrix with the observed one.

Another possible variant is given by the inclusion of additional constraints which, however, should be empirically justified. Suppose, for instance, that some recent legislation requires a minimum level of education, to access to specific occupations. This would be easily introduced as an extra constraint in the optimization problem. However, its introduction would be necessary only if such a constraint could be active for the labor market in the aggregate, and not for only particular job positions.

7. Concluding Remarks

Our numerical simulation exercise suggests that, when the structure of an employment by education matrix is taken as a starting point, and a new matrix can be estimated by minimizing a concept of total distance from the base one, then a redistribution of workplaces is obtained, where – when labor demand gets weaker – some educated workers accommodate into jobs with lower knowledge, productivity and possibly wages. This is consistent with the empirical finding that, in response to a decreasing labor demand, some educated workers can avoid getting unemployed, by accepting jobs in lower-ranked occupations.

Overqualification is, therefore, one form of underemployment, but not the only one. Bell and Blanchflower (2013, 2018a-b, 2020) have investigated another kind of underemployment, the one related to the time worked, and its implications. They propose a method to compute underemployment as the difference between the hours one person would like to work at constant wage rate, and the actual numbers of hours she works. Having estimated underemployment measures for twenty-five European countries, they show that, in contrast to the unemployment rate, underemployment in most countries has not returned to its prerecession levels after the financial crisis of 2008-2009.

Reduction of unemployment has always been one of the key targets of economic policy. In Murphy and Topel (1997) words: “the unemployment rate is a summary statistic for the overall state of the economy and for the success or failure of economic policy”. Surprisingly, underemployment, in its various dimensions, has not received the same degree of attention, despite its raising relevance. Indeed, jobs quality rather than quantity is getting more important, most notably in the developed economies (Findlay, Kalleberg, and Warhurst, 2013).

Numerical macroeconomic models are available for the simulation and assessment of economic policies, as well as of external shocks (like wars, pandemics, climate change, earthquakes, etc.). They include computable or dynamic stochastic general equilibrium models, macro-econometric models, input-output models, and others. Often, these models can evaluate the impact on employment but, to the best of our knowledge, no information can be provided in terms of overeducation effects. In this paper, we propose a methodology to partly fill this gap, focusing on the relationship between educational attainment and occupation. As an example, we undertook a relatively standard comparative-static simulation with a CGE model, which provides estimates of changes in labor demand, and subsequently use these results to generate a counterfactual EE matrix for Italy, from where one can deduct how the distribution of workers among the various occupations could vary.

Our proposed methodology, however, does not necessarily need to be employed in conjunction with a macroeconomic model. For instance, Albanesi et al. (2023) examine the link between labor market developments and new technologies, such as artificial intelligence. Using occupational measures of artificial intelligence exposure, they find that AI-enabled automation in Europe is associated with employment increases, and that his positive relationship is mostly driven by occupations with relatively higher proportion of skilled workers. Using their estimates, it would then be possible to estimate the employment impact of AI technology diffusion, and our method to infer the associated effect on the educational match.

Of course, much work remains to be done. Microeconomic analysis could complement the macroeconomic simulations and provide more solid basis for the behavioral assumption employed therein. Other forms of underemployment could be considered and modelled, and we need to better understand how the various types of underemployment interact.

We can conclude that the phenomenon deserves more consideration in applied and theoretical research, but also that it should be made central in the economic policy debate. For the latter objective, an important role could be played by the availability of a new generation of numerical macroeconomic models, assessing impacts not just on macroeconomic variables, like GDP, consumption levels, trade flows, employment, but also on the extent of underemployment, overqualification, and skill mismatch. The methodology presented in this paper can therefore be regarded as a step in this direction.

References

Aguiar, A., Chepeliev, M., Corong, E. L., McDougall, R., van Der Mensbrugghe, D. (2019). The GTAP data base: version 10. *Journal of Global Economic Analysis*, *4*(1), 1-27.

Albanesi, S., da Silva, A. D., Jimeno, J. F., Lamo, A., Wabitsch, A. (2023). *New Technologies and Jobs in Europe* (No. w31357). National Bureau of Economic Research.

Baert, S., Verhaest, D. (2019). Unemployment or overeducation: Which is a worse signal to employers? *De Economist*, 167, 1–21.

Bell, D. N., Blanchflower, D. G. (2013). Underemployment in the UK revisited. *National Institute Economic Review*, *224*, F8-F22.

Bell, D. N., Blanchflower, D. G. (2018a). The lack of wage growth and the falling NAIRU. *National Institute Economic Review*, *245*, R40-R55.

Bell, D. N., Blanchflower, D. G. (2018b). *Underemployment in the US and Europe* (No. w24927). National Bureau of Economic Research.

Bell, D. N., Blanchflower, D. G. (2020). US and UK labour markets before and during the Covid-19 crash. *National Institute Economic Review*, *252*, R52-R69.

Büchel, F., Van Ham, M. (2003). Overeducation, regional labour markets, and spatial flexibility. *Journal of Urban Economic*s, 53, 482–493.

Card, D., Kramarz, F., Lemieux, T. (1999). Changes in the relative structure of wages and employment: A comparison of the United States, Canada, and France. *The Canadian Journal of Economics / Revue Canadienne d'Economique*, 32(4), 843-877.

Carneiro, P., Heckman, J. J., Vytlacil, E. J. (2011). Estimating marginal returns to education. *American Economic Review*, *101*(6), 2754-2781.

Caroleo, F.E., Pastore, F. (2015). *Overeducation: a disease of the school-to-work transition system*. IZA Discussion Papers, 9049, Institute of Labor Economics (IZA).

Chepeliev, M., Hertel, T., and van der Mensbrugghe, D. (2022). Cutting Russia's fossil fuel exports: Short-term economic pain for long-term environmental gain. *The World Economy*, *45*(11), 3314-3343.

Corong, E. L., Hertel, T. W., McDougall, R., Tsigas, M. E., van der Mensbrugghe, D. (2017). The standard GTAP model, version 7. *Journal of Global Economic Analysis*, *2*(1), 1-119.

Di Pietro, G. (2002). Technological change, labor markets, and 'low-skill, low-technology traps. *Technological forecasting & social change*, 69, 885-895.

Di Pietro, G., Cutillo, A. (2006). University quality and labour market outcomes in Italy. *Labour*, 20(1), 37–62.

EUROSTAT (2023). Unemployment by sex, age and educational attainment (1998-2020) - annual data. https://ec.europa.eu/eurostat/web/products-datasets/-/une_educ_a_h. Accessed January 23, 2023.

Findlay, P., Kalleberg, A. L., Warhurst, C. (2013). The challenge of job quality. *Human Relations*, *66*(4), 441-451.

Golan, A., Judge, G., Robinson, S. (1994). Recovering information from incomplete or partial multisectoral economic data. *The Review of Economics and Statistics*, 541-549.

Gootschalk, P., Hansen, M. (2003). Is the proportion of college workers in noncollege jobs increasing? *Journal of Labor Economics, 21*, 449-471.

Johnes, G. (2019). The incidence of and returns to 'overeducation': PIAAC evidence on the G7. *Minerva*, 57, 85–107.

Leuven, E., Oosterbeek, H. (2011). *Overeducation and mismatch in the labor market*. IZA Discussion Papers, 5523, Institute of Labor Economics (IZA).

McGuinness, S. (2006). Overeducation in the labour market. *Journal of Economic Surveys*, 20, 387-418.

McGuinness, S., Sloane, PJ. (2011). Labour market mismatch among UK graduates: An analysis using REFLEX data. *Economics of Education Review, 30,* 130-145.

Murphy, K. M., Topel, R. (1997). Unemployment and nonemployment. *The American Economic Review*, *87*(2), 295-300.

Ordine, P., Rose, G. (2009). Overeducation and instructional quality: A theoretical model and some facts. *Journal of Human Capital*, 3(1), 73–105.

Quintini, G. (2011). Right for the job: *Over-qualified or under-skilled?* OECD Social, Employment and Migration Working Papers 120, OECD Publishing, Paris.

Robinson, S., Cattaneo, A., El-Said, M. (2001). Updating and estimating a social accounting matrix using cross entropy methods. *Economic Systems Research*, 13(1), 47-64.

Roson, R. (2022). *Underemployment in a Computable General Equilibrium Model*. Dept. of Economics Working Paper 17/2022, Ca'Foscari University, Venice.

Sattinger, M. (1993). Assignment models of the distribution of earnings. *Journal of Economic Literature*, 31(2), 831–880.

UNESCO (2010). Education counts: towards the Millennium Development Goals. Available at: https://unesdoc.unesco.org/ark:/48223/pf0000190214.

Verhaest, D., Van Der Velden, R. (2013). Cross-country differences in graduate overeducation and its persistence. European Sociological Review, 29, 642-653

Appendix

ACRONYMS DESCRIPTION

Table A1 - Classification of occupations

CP2011 code	Abbr.	Description
1.1	PUBEXEC	Members of legislative and governmental bodies, executives, and equivalent positions in public administration, judiciary, healthcare, education, research, and national and supranational organizations.
1.2	ADLARGE	Entrepreneurs, administrators, and directors of large companies.
1.3	ADSMALL	Entrepreneurs and managers of small businesses.
2.1	SPMATHC	Specialists in mathematical, computer, chemical, physical, and natural sciences.
2.2	ENGARCH	Engineers, architects, and related professions.
2.3	SPLIFESC	Specialists in life sciences.
2.4	HEALTHS	Health specialists.
2.5	SPHUMAN	Specialists in humanities, social sciences, arts, and management.
2.6	SPEDUCR	Specialists in education and research.
3.1	TCSCIEN	Technical professions in the scientific, engineering, and production fields.
3.2	TCHEALT	Technical professions in healthcare and life sciences.
3.3	TCADMIN	Technical professions in organization, administration, and financial and commercial activities.
3.4	TCPUBLS	Technical professions in public services and for people.
4.1	EMSECRO	Employees responsible for secretarial and office machine functions.
4.2	EMMONCA	Employees responsible for money movements and customer assistance.
4.3	EMADMIN	Employees responsible for administrative, accounting, and financial management.
4.4	EMDOCUM	Employees responsible for collecting, checking, preserving, and delivering documentation.
5.1	SKCOMMA	Skilled professions in commercial activities.
5.2	SKHOSPF	Skilled professions in hospitality and food service.
5.3	SKHEALS	Skilled professions in healthcare and social services.
5.4	SKPERSS	Skilled professions in cultural, security, cleaning, and personal services.
6.1	SWEXCON	Artisans and specialized workers in the extractive industry, construction, and building maintenance.
6.2	SWELECT	Specialized metalworkers, installers, and maintainers of electrical and electronic equipment.
6.3	SWPREPR	Specialized artisans and workers in precision mechanics, artistic crafts, printing, and related fields.
6.4	SWAGRIC	Farmers and specialized workers in agriculture, forestry, animal husbandry, fishing, and hunting.
6.5	SWOTHIN	Artisans and specialized workers in food processing, wood, textile, clothing, leather, and entertainment industries.
7.1	OPINDPL	Industrial plant operators.
7.2	OPFIXMA	Semi-skilled operators of fixed machinery for mass production and assembly line workers.
7.3	OPFMAGR	Fixed machinery operators in agriculture and the food industry.
7.4	OPVEHIC	Vehicle, mobile machinery, and lifting equipment operators.
8.1	UNSERVC	Unskilled professions in commerce and services.
8.2	UNDOMEC	Unskilled professions in domestic, recreational, and cultural activities.
8.3	UNAGRIC	Unskilled professions in agriculture, green maintenance, animal husbandry, forestry, and fishing.
8.4	UNMANUF	Unskilled professions in manufacturing, mineral extraction, and construction.
9.1	MILIOFF	Military officers.
9.2	MILISER	Non-commissioned officers and sergeants of the armed forces.
9.3	MILIPER	Military personnel.

Table A2 – Levels of Education

ISCED level	Abbr.	Description
0	EARLYED	Early childhood education
1	PRIMAED	Primary education
2	LOW2EDU	Lower secondary education
3_3	UP2EDUL	Upper secondary education (L)
3_4	UP2EDUH	Upper secondary education (H)
4	POST2ED	Post-secondary non-tertiary education
5	SHORT3E	Short-cycle tertiary education
6	BACHELR	Bachelor’s or equivalent level
7	MASTERL	Master’s or equivalent level
8	DOCTORL	Doctoral or equivalent level