Employment and Skills in the Netherlands

Chapter 4:

Skill Relation to Employment, Unemployment and Non-Participation Rates: An Analysis of the IALS Data for the Netherlands

Sebastian Buhai

Utrecht University College
Postbus 81-081, Postcode 3508 BB,
Utrecht, The Netherlands
Email: sbuhai@ucu.uu.nl

Contents

4.1 Introduction

The relevance of the skill distribution in explaining the employment-population ratios across Europe and US has been widely debated during the last decades. Several studies have reached agreement in that the low-educated people have a higher chance to be unemployed or out of the labour force, compared to the high-educated people. For instance, Stephen Nickell and Brian Bell have pointed out that the unemployment rate of the low-educated group is substantially higher than that of the high-educated group in seven OECD countries[i] using economical data from 1971 till 1993 (Nickell and Bell: 302). In the present paper we will pronounce ourselves with reference to the importance of the skills in the employment and unemployment population ratios of the Netherlands by analysing the IALS data set collected in 1993.

In what follows we will describe the characteristics of the sample used in our analysis. Questions were answered by a sample of 3090 people. Relevant for our analysis are the subject's answers to the labour force status inquiry (variable D1 in the data set) and their skills, assessed following a quantitative test that the subjects had to perform (variable QUANT in the data set). With regards to the labour force status, the subjects had a choice

between six possible outcomes: 'employed', 'retired', 'unemployed', 'student', 'homemake' and 'other'. For our analysis, all the persons with different answers then 'employed' or 'unemployed' are considered 'out of the labour force' (thus 'retired', 'homemake' and 'other' adding to the non-participation category). Concerning the different ways in which the skills were measured, it has been proved in chapter 1 of this research paper that for our purpose the subjects' skills are very well described by their quantitative score (QUANT) and that all other skill indicators would produce similar outcomes.

The IALS economic data set has been given beforehand, thus any responsibility regarding the reliability of the information contained shall be beyond our concern. The analysis has been performed using the statistical software Intercooled Stata 6.0.

4.2 Compared skill distribution between employment, unemployment and non-participation

As mentioned earlier, the sample of 3090 persons has been divided into 3 categories, 'employed', 'unemployed' and 'out of the labour force', according to the subjects' own classification.

We will further analyse the distribution of the skills within each of the above mentioned social categories. Below we present statistical summaries for the distribution of the skills within each of the 3 categories with an associated histogram attached for better illustrating the results.

We notice by looking at the three consecutive tables below that there are 1815 observations for the employed people, the unemployed category has a size of 125 subjects, while 1150 persons classified themselves as out of the labour market (that is retired, student, homemake or other). What is striking from the very beginning is that the number of unemployed people is extremely small compared to that of the employed and that of the out-of-the-labour persons; thus in this sample the unemployment rate is about 4.05%. This result is considerably lower than the reported unemployment rate in the Netherlands in 1993 (only the male unemployment rate stayed at 6.8%). This sample unemployment rate conforms rather to 1998 figures. Nowadays we expect even lower numbers for the unemployment and higher for the out of the labour market, as the unemployment in1999 stayed at 3.5%, but "generous welfare benefits have prompted large numbers to drop out of the labour market", after a CIA factbook estimation [ii]. For our purpose the sample size of only 125 for the unemployed is statistically very low and it might cause low power estimates and possibly even insignificant models.

Quantity Scores Employed People

	Percentiles	Smallest		
1%	177.496	78.30794		
5%	221.4142	80.88482		
10%	244.073	99.84394	Obs	1815
25%	274.7214	121.9001	Sum of Wgt	6454522.15
50%	301.9911		Mean	297.7093
		Largest	Std. Dev.	42.01751
75%	325.3782	404.4946		
90%	345.7752	405.9382	Variance	1765.471
95%	358.2269	407.7414	Skewness	-.8189199
99%	379.1469	411.7614	Kurtosis	4.787906

distribution of skills for employed people
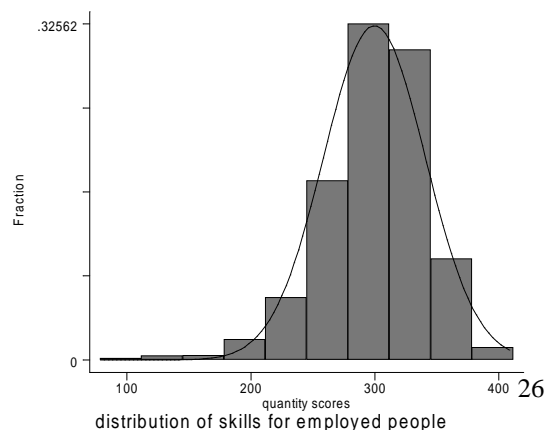
Quantity Scores for the Unemployed				
	Percentiles	Smallest		
1%	105.549	95.71884		
5%	154.1806	105.549		
10%	204.0729	115.0621	Obs	125
25%	249.1764	136.1519	Sum of Wgt.	449351.123
50%	288.203		Mean	273.4121
		Largest	Std. Dev	54.2288
75%	307.2821	358.5934		
90%	328.4919	360.2149	Variance	2940.762
95%	335.3408	378.202	Skewness	-1.141359
99%	380.1462	380.1462	Kurtosis	4.51966

distribution of skills for unemployed people
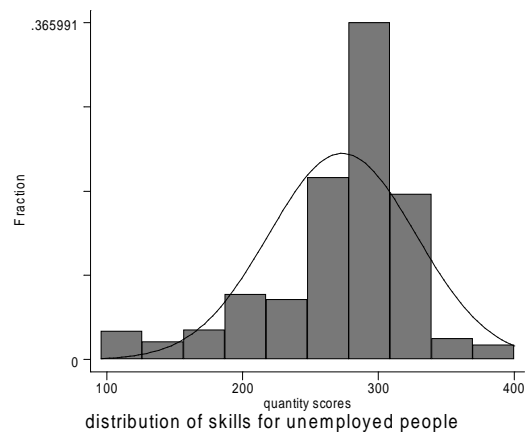
Quantity Scores for the Out-of-Labor People				
	Percentiles	Smallest		
1%	109.6403	60.31372		
5%	172.2921	80.33638		
10%	193.7202	83.08902	Obs	1150
25%	237.5508	87.95364	Sum of Wgt.	4591845.75
50%	270.7953		Mean	265.1075
		Largest	Std. Dev.	51.44604
75%	300.3218	368.0091		
90%	325.1123	368.498	Variance	2646.695
95%	337.5651	369.1324	Skewness	-.7877422
99%	361.5462	372.0411	Kurtosis	3.845681

distribution of skills for out of the labor people
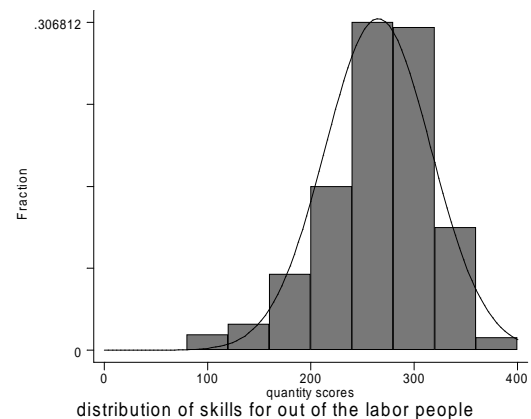
If we carefully analyse each of the statistical summaries and associated histograms printed above, we notice that the way skills are distributed among the three social categories slightly differs, but in general an assumption of normality seems to be respected. Apart from the small negative skewness in each of the cases, indicating a shift to the right of the quantity scores, there are no significant deviations to be signalled for all the three categories. Coming back to the small negative skewness (less than –1.5 in each case), this underlies the fact that in each of the categories both the median and the mean of the observations are above the average (265.107 is the lowest in this sense, the mean for the out of the labour group). Also, except few cases within the unemployed group, people do not present very low skills. This line of reasoning is in accordance with reality, appropriate training within particular EU countries like Germany, the Netherlands and Switzerland being aimed at the lower skilled levels as well and thus raising the overall education levels.

From a high to low median and mean, the 3 categories are employed, unemployed, respectively out of the labour market (see the statistical summaries above) This ranking seems to be observed also from the percentile scores. We have to point out here that the very first percentile of employed people scores higher that the 5th percentile in the other 2 categories. Thus the higher education of the lowest skilled among the employed people compared to the lowest skilled among the out of the labour or unemployed is clear. Another important observation we have to make is that the variance and the standard deviation in case of the unemployment category are higher than for the unemployed than for the other 2

groups, partly underlying the low sample size of the unemployed category (there seem to be also a few outliers within the unemployed category, with far higher skills than everybody else in that group; also a consequence of the very low sample size).

To sum up and conclude, while the skill seems to be rather normally distributed in each of the categories, a comparison between the groups shows that the average employed people score higher than the average unemployed or out of the labour people. Also, there are few people scoring very low in the employed category than in the out of the labour market and the unemployed category. The extremely high QUANT scores seem to be attained by equal proportions in each of the three group (about 1%), nevertheless the higher QUANT scores are easier attainable by employed and out of the labour people than by the unemployed category. It seems, as a whole, that the socio-economical distribution is not independent from the skills that people have. We will elaborate in detail on this statement in the following section.

4.3 Employment, unemployment and non-participation rates by skill

In order to assess an eventual relation between skills and employment-population ratios we will construct the following new variables. A 'percentile quantity' (QUANT50) variable containing the highest QUANT scores attained in 50 percentile groups of the whole sample will constitute the independent variable in our further analysis. For each of the 3 social categories we will construct a category-to-population ratio variable with 50 values, one for every of the 50 skill percentile groups. Thus 'emprap' will contain 50 observations of the percentages of employed people in each of the 50 equally numbered groups distinguished earlier (each of the 50 skill percentile groups has 62 or 61 subjects). For illustrating this technique, the first value of emprap will be 16.12, as the employment to population ratio in the lowest skills group is 16.12 %, while the last (50th) value of emprap will be 90.32, as the same employment-population ration in the highest skills group is 90.32%. We will run the following linear regressions: employment-population ratio variable (emprap) to QUANT50, unemployment-population ratio (unemprap) to QUANT50, respectively out of the labour-population ratio to QUANT50. It is noteworthy that a Pearson chi-square test is not impossible in this case, predicting a significant p value, nonetheless failing to indicate the exact relation between skills and the social categories[iii]

The hypothesis we support is that skills contribute to these ratios in the following way: the lower the skills, the lower the chance to be employed and consequently the higher the chance to be unemployed and out of the labour. By this hypothesis we would verify the conclusion that Nickell and Bell reached in an analysis mentioned earlier. In addition, in this particular analysis concerning the Netherlands, we need to underline the importance of the 'out of the labour people', given the size of this category compared to the declared unemployed people. As a matter a fact we suspect that many of the people that have declared themselves 'student' or 'other', thus contributing to the 'out of the labour force', could normally add to the unemployment in other circumstances. Corroborating to our assumption, official economic indicators show that in the Netherlands the present registered unemployment tends to go down significantly because of latest developments concerning increased welfare aid. Moreover seasonal effects can matter, during late summer for instance, because of school leavers adding to the job-market, the unemployment rate usually reaches higher figures[iv].

Hence it makes sense to wonder when the data was gathered and what are the implications of that specific season for the differences between unemployment size and out of the labour size. The continuous inflow and outflow between these two categories are highly relevant in the Netherlands. We will come back to this point when analysing the particular results for the unemployed category.

Starting with the regression emprap= f (QUANT50), we try to find whether the employment population ratios could be a function of the employed people's skills. The result of the regression is plotted below.

coef = .35137058, se = .01889792, t = 18.59

Source	SS	df	MS		Number of obs =	50
					F(1, 48) =	345.70
Model	12993.4969	1	12993.4969		Prob > F =	0.0000
Residual	1804.11581	48	37.585746		R-squared =	0.8781
					Adj R-squared =	0.8755
Total	14797.6127	49	301.992095		Root MSE =	6.1307

emprap	Coef.	Std. Err.	t	P>t	[95% Conf. Interval]	
QUANT50	.3513706	.0188979	18.593	0.000	3133738	.3893674
_cons	-43.1688	5.548417	-7.780	0.000	-54.32464	-32.01296
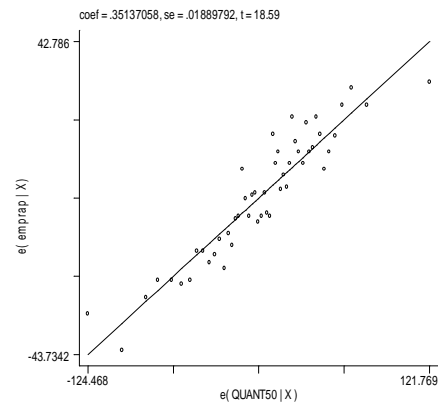
The model is significant, as a first observation (p = 0), thus a relation between the two variables has been found. We see that the coefficient of the regression is positive, therefore the function is increasing. The power estimate of the regression is very high (almost 88%), therefore also its explanation capacity is beyond doubt. In the right we see indeed that the regression line fits. To summarise our findings, we have obtained emprap=-43.16+0.35 *QUANT50, that is the employment-population ratio increases with the increase in the skills. In other words, the more skilled the people are, the lower the chance that they are unemployed and equivalently, the higher the likelihood that they are employed.

After this interesting result we have obtained for the employment-population ratio, we further regress the unemployment-population ratio to the skill percentile variable.

coef = -.01074506, se = .01063493, t = -1.01

Source	SS	df	MS		Number of obs =	50
					F(1, 48) =	1.02
Model	12.151025	1	12.1510		Prob > F =	0.3174
Residual	571.354699	48	11.9032229		R-squared =	0.0208
					Adj R-squared =	0.0004
Total	583.505724	49	11.9082801		Root MSE =	3.4501

| unemprap	Coef.	Std. Err.	t	P>t	[95% Conf. Interval]	
QUANT50	-.0107451	0106349	-1.010	0.317	-. 032128	.0106379
_cons	7.163053	3.122409	2.294	0.026	.8850292	13.44108
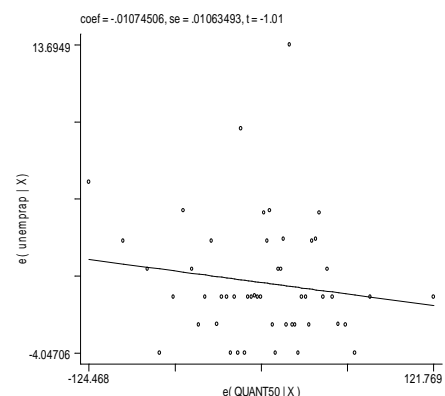
We obtained a model that is not significant (p=0.317>0.05), so apparently there is no relation between the unemployment and the skills. Nonetheless we know this is an aberration, as we have just found the result between the employment-population ratio and the skills. The failure

to obtain a significant result can be explained by the fact that we have a sample size of unemployed people that is too low, first of all. Secondly, dividing this already relatively low sample in other 50 groups reduces the sample size of each of the groups to extremely low figures (3-4 people per group). Thus the estimation is likely not to be significant and not to contribute with too much explanation power (indeed R-squared is only 0.02). Thirdly, in the previous section we have analysed the skill distribution within each of the categories and we have found that the variance and the standard deviation are highest among the 3 categories (see Section 4.2). Thus we expect a less 'good' prediction than in the other 2 cases. Nonetheless we anticipate obtaining a monotonically decreasing function. Ignoring for a moment the p value, we see that the coefficient is negative (-0.01) and the plot shows a decreasing function as the regression fit line.

As stated in the beginning of this section, we suspect that many of the 'out of the labour' people could add to the unemployment level. Inflows and outflows between the two socio-economical categories are particularly important for Netherlands, a country providing for more than reasonable welfare benefits. The present employment rate of the Netherlands is among the lowest in Europe and is still falling down still, albeit slower than in the beginning of this year[v]. This does not belittle at all the relevance of the 'steady' out-of-labour people; among the 1150 people characterised as out of the labour more than 700, thus more than 60%, have declared themselves as 'homemake'. About 75% of these homemake persons are house women. Unlike other categories within the out of the labour force, this group is unlikely to change their labour force status. Thus the heavy nucleus of the non-participation group is fixed and inflows and outflows only affect considerably the unemployment ratio (it is self-understood that the employment level as such is also more or less fixed in the Netherlands). Thus it makes sense to consider the regression outrap= f (QUANT50). Results are presented below.

Source	SS	df	MS		
				Number of obs = 50	
				F(1, 48) = 307.05	
Model	12210.955	1	12210.955	Prob > F= 0.0000	
Residual	1908.89954	48	39.7687403	R-squared= 0.8648	
				Adj R-squared = 0.8620	
Total	14119.8546	49	288.160297	Root MSE = 6.3062	

outrap	Coef.	Std. Err.	t	P>t	[95% Conf. Interval]	
QUANT50	-.3406255	.019439	-17.523	0.000	-.3797102	-.3015409
_cons	136.0058	5.70727	23.830	0.000	124.5305	147.481

coef = -.34062553, se = .01943897, t = -17.52
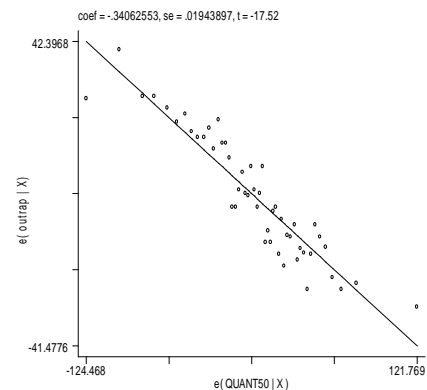
We can see from the regression summary above and the associated plot of the regression line that it made sense considering this regression. The model is significant (p=0) and explains 86% ($R^2 = 0.86$). The coefficient of the regression is –0.34. Thus, the relation we have found is outrap= 136-0.34* QUANT50. The plot in the right of the table above clearly shows that the negative relation. In other words we have found that the higher the skills the lower the chances of being out of the labour force. Indeed, it is unlikely that a highly trained person chooses to be out of the labour force. Now we can also explain why the mean and median for the skill distribution within the non-participating group was lower than that of the unemployed. Within unemployment most of the people are looking for a job, thus have a minimal incentive to work and probably some adequate minimal training. Within the out of

the labour market group most of the people choose to be out of the labour market and usually do not need high education level or particular training.

Obviously, in an equivalent manner to the spotted increasing relation skills-employment level we can refer to the decreasing relation between the skills and the overall non-employment level (that is unemployment together with out of labour force), without performing another regression. As observed, we chose to handle in this paper both unemployment and out of the labour group separately and thus obtain more insight in the relation between skills and socio-economic categories as such.

4.4 Conclusions

We have shown in this paper that the education level, informally referred to as skills, and the employment within the Netherlands are not at all independent of each other. The statistical results show that the unemployment rate of the low-education group is substantially higher than that of the high-education group. Thus one conclusion that we can immediately infer is that higher skilled people are preferred to low skilled people. It is worth to mention that the possible causes of these preference, albeit a shift in the demand against the unskilled or not[vi], are beyond our purpose here. It seems that several previous studied investigating economic data for other OECD states reached similar results in this sense, thus the Netherlands does not exhibit a very particular behaviour from this point of view.

A second interesting conclusion that we can draw is that, in particular for the Netherlands, the relevance of the non-participant persons to the labour market. On the background of a very low unemployment rate (still decreasing, as official economic indicators point out), the size of the out of the labour category is impressive. Out of the labour figures are almost comparable in size to the employment figures. To the extent that this sample is representative for the whole population in the Netherlands, the huge out of the labour market group can be explained on the one hand by the peculiar welfare benefit addressed to some of these categories. To be more concrete, here we would count the students and the retired persons. On the other hand, the rigid and majority part of the out of the labour market, composed of housemakes, in majority either low or average skilled (in accordance with the negative regression relation found), seems to simply choose not to register as unemployed and thus to leave from own resources. It is without any doubt interesting to find the reasons for this phenomenon, nonetheless a second paper could be written in that purpose. For us it might be interesting to pinpoint that such a high ratio of the out of the labor market persons to the population might indicate after all shows after all self-satisfaction. With or without high skills, people seem to be satisfied with their social-economic condition in the Netherlands.

References

Nickell, S.J. and Bell, B. (1996) 'Changes in the Distribution of Wages and Unemployment in the OECD countries' in American Economic Review, 86, Papers and Proceedings, 302-308

Endnotes

---

[i] The 8 OECD countries analysed by Nickell and Bell were Germany, Netherlands, UK, Spain, Italy, Sweden, Canada and US. Italy was not found to exhibit the mentioned relation between education and employment.

[ii] The source of the information is the CIA world factbook site; the exact quotation has been taken from the URL: http://www.odci.gov/cia/publications/factbook/fields/unemployment_rate.html

[iii] A cross-tabulation between two 3 by 3 variables, one containing the number of people in each category, the other containing the quantity scores in 3 percentile categories, has been performed. Although the p-value of the Pearson chi-square test was 0, the explanation power was only 45%. In other words, a relation was found, nonetheless an obvious pattern could not be discerned in this way.

[iv] The paragraph was paraphrased after the findings of the Centraal Bureau voor Statistiek in Netherlands. The source and relevant information about seasonal increase or decrease in the registered unemployment as well as other relevant economical facts at http://www.cbs.nl/en/figures/economic-indicators/p-102-09.htm

[v] ibidem, the information is taken from CBS sources

[vi] See Nickell and Bell (the reference article) for a detailed analysis of this assumption