Term Paper Proposal

Sarah Gust *ifo Institute at the University of Munich, gust@ifo.de*
Ann-Christin Kreyer *Max Planck Institute for Innovation and Competition, ann-christin.kreyer@ip.mpg.de*

This is where we put the abstract. . .

Introduction

What is the topic of your dissertation? What were the reasons for choosing this topic? What is your **research hypothesis**? How do you define the central terms of your hypothesis? Why and for whom is it important to answer these questions?

- Motivate (one paragraph).
- Summarize what we know (one paragraph).
- Third paragraph: Tell us what you are doing!
- Describe research design.
- Value added.
- Summarize key findings.
- DO NOT write a "roadmap".
- No subsections in the Introduction

Literature

(We contribute to the literature on digitalization, job tasks, training, and job mobility.)

Computer capital and workers that perform routine tasks are substitutes whereas computer capital and workers that perform non-routine cognitive tasks are complements (Autor et al., 2003). The declining price of computer capital has led to a U-shaped labor demand function (Acemoglu and Autor, 2011; Autor and Dorn, 2013). This indicates that middle-skilled workers were replaced by technology while the demand for high skilled workers and low-skilled workers has grown. Autor and Dorn (2013) show that increased employment at the lower tail of the earnings distribution is mainly due to an increase in service occupations. At the upper tail technological changed led to a college wage premium: wages of college graduates relative to high-school graduates increased

(Acemoglu and Autor, 2011). De La Rica et al. (2020) develop abstract, routine and manual task measures and find that a one-standard-deviation increase in abstract tasks is related to a 3.3-log-point wage premium. For each standard deviation of routine tasks there is a 2.6 to 2.9-log-point wage penalty.

This vast literature on the change of job tasks has mainly focused on wage effects and takes skills of workers as pre-defined. However, workers and firms could also invest in new skills via training.

Becker (1962) distinguished between two kinds of on-the-job training: specific and general. Specific training increases the marginal product of a worker within one specific firm while general training increases her productivity in many other firms. In a perfect labor market workers are paid their marginal product. In such case, firms would not invest into general training of their employees as they could leave the firm and look for a better paid job. Instead, workers would pay for their general training as an investment into higher future wages. Acemoglu and Pischke (1999) argue that firms still invest in general training due to their monopsony power. Wages increase by less than the marginal productivity and firms can profit. Konings and Vanormelingen (2015) find that an increase in the share of trained work-ers by 10 percentage points raises the productivity by 1.7 to 3.2 percent while wages only increase by 1.0 to 1.7 percent.

Evidence on job mobility behaviour of workers is more mixed. Zweimüller et al. (2003) findings support Becker (1962) human capital theory. Workers who received firm specific training quit less often and show less job searching behaviour. Workers who received general training increased their job searching activities and quit more often. Dietz and Zwick (2020) use German employer-employee data and find that training increases the retention probability. These studies focus on on-the-job training.

Lynch (1991) and Lynch (1992) compares on-the-job to off-the-job training. She focuses on young workers that are particularly mobile. She finds that on-the job training tends to be firm specific in the US and thus wage raises cannot be taken along to subsequent employers. Off-the-job training by proprietary institution have little effect on wages in the current employment but raise future expected wages in subsequent employment. Lynch (1991) shows that the probability of leaving an employer varies with respect to race, gender, and educational level. Workers with disabilities, black workers and workers with a high school degree or less increased the probability of leaving the first employer. Working in a job with collective agree-ment or having a college degree decreased their probability of leaving the employer. The effect of training, disability, and education disappears when Lynch (1991) re-estimates the equation only for men, while these effects are particularly strong for women.

Applying a machine learning approach permits us to take a broader approach on this topic. Instead of restricting our estimation to a specific group of workers or countries, we identify the factors that drive the probability of leaving-a-job and job-switching from a set of ## variables.

- mostly effect on wages, inequality

Data

What is the epistemological framework of the dissertation? For empirical studies it should be made clear: Why were the specific methods of data analysis chosen? How was the data acquired?

To explore these questions we use the results of the survey of the Programme for the International Assessment of Adult Competencies (PIAAC) (GESIS - Leibniz Institute for the Social Sciences (n.d.)). The survey was conducted by the Organisation for Economic Co-operation and Development (OECD) with the goal to assess which skills adults need to manage challenges and tasks at work as well as in their personal life. The study targeted explicitly the skills in literacy, numeracy and adaptive problem solving while also comprising comprehensive background information on the respondents past and current education, subjective assessments of their skills and job requirements as well as information on migration (GESIS - Leibniz Institute for the Social Sciences (n.d.)). The first cycle consisted of three rounds and began in 2011/12. In the first round 24 countries took part. In the second round nine additional countries participated and in the last round individuals from five different countries were questioned. In total 40 countries participated in the first cycle comprising about 5,000 randomly selected adults who were between 16 and 65 years old. The second cycle started in 2018 and results are to be expected in 2022 (GESIS - Leibniz Institute for the Social Sciences (n.d.)).

For the study at hand the results of the first wave are used in a reduced form. The original dataset comprises 1,460 columns with 230,691 observations of respondents. However, the 'research question of this paper is to analyze the probability of trainings for middle-skilled workers compared to trainings for high-skilled and low-skilled workers in the wake of the increasing polarization of skills following from digitization'. To answer this, the original PIAAC dataset is reduced to 130 variables in total, including indices. For example, information on the various test results conducted in the study are excluded. To achieve comparability across countries, questions which were only answered by respondents living in the United States are also excluded. The final dataset is cross-sectional with one observation representing the answers of one respondent.

The 130 variables we kept in the final dataset comprise information on the individual's background information, her past and ongoing formal or informal education, information on training activities, information on ICT skills and the respective extensive and intensive margin, her subjective job requirements, information on her curent job and information on monthly income. As we do not restrict the dataset in terms of respondents but only in terms of questions answered, our final dataset comprises 230,691 observations of individuals. Of those 230,691 individuals, 122,830 are female and 107,859 are male (see Figure 1). The age of the respondents is evenly distributed between the ages 16 to 64 with a female mean age of 39.95 years and male mean age of 39.38 (see Figure 2).
The classification of the respondents jobs in terms of skills is also evenly distributed across age groups and gender. However, respondents working in semi-skilled white-collar occupations are slightly younger than those working in skilled occupations or semi-skilled blue-collar occupations (see Fiugre 3).

The key variables of this study are the the skill classification of the individual's job and her trainings comprising on-the-job training, seminars or workshops, distance or open training courses as well as private lessons. The simple OLS regressions reveal that there are positive correlations between the high-skilled jobs and the number of trainings respondents participated in. This holds true for on-the-job-training, seminars or workshops, distance or open educational training as well as for private lessons. However, for semi-skilled jobs, the picture is slightly different. Here, the number of seminars or workshops and private lessons are positively correlated with the semi-skilled occupations.

- Name, source, unit, time, structure, number of observations, relevant population.
- Definition of (main) sample.

Figure 1: Distribution of Gender

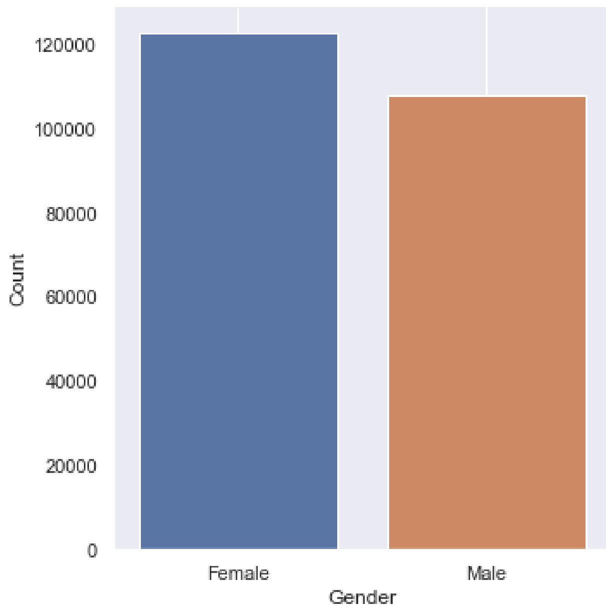


Figure 2: Distribution of Age

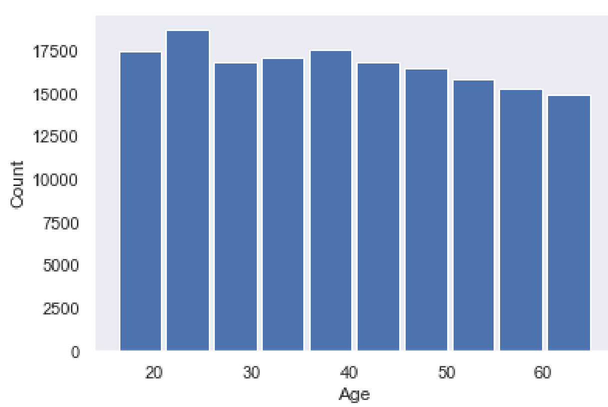
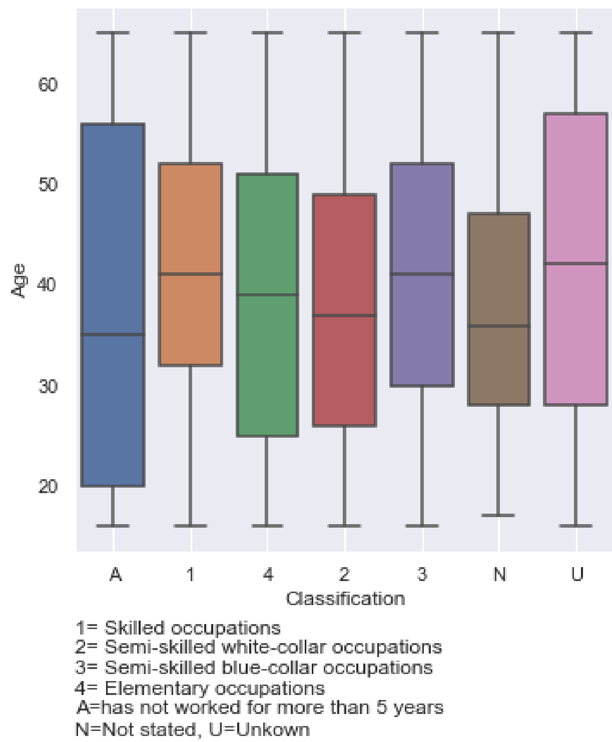


Figure 3: Age and Classification of jobs



- Definition and characteristics of key variables.
- Limitations and potential biases.
- Provide the data and the software code (replication).
- Plot the main empirical associations you want to study!
- Do NOT assume the reader knows anything about these data

Lasso Linear Model

We start with a simple linear regression model $y = X\beta + \varepsilon$, where $y \in \mathbb{R}^N$ is the predicted participation in on-job or off-job training, $X \in \mathbb{R}^{N \times k}$ are the vectors of covariates, and $\varepsilon \in \mathbb{R}^N$ is the residual with the standard assumptions of OLS. To select the set of covariates with the strongest predictive power from our set of $\ldots$ variables, we apply the Least absolute shrinkage and selection operator (Lasso) that was first proposed by Tibshirani (1996). We add the Lasso penalty equal to $\sum_k |\beta_k|$ to our linear model. The Lasso linear estimator $\hat{\beta}$ is then given by

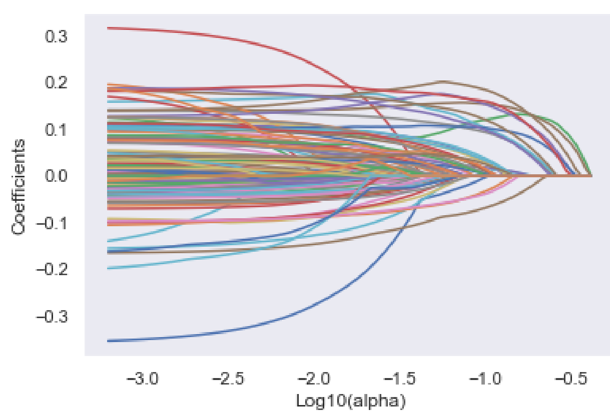
$$\hat{\beta}_\lambda = argmin\{\sum_i (y_i - x_i^T\beta)^2 + \lambda \sum_k |\beta_k|\} \tag{1}$$

That is, we minimize the sum of squared residuals but the Lasso penalty puts a cost at every $\hat{\beta} \neq 0$ and thus, we penalize complexity and avoid over-fitting the model. $\lambda > 0$ is the penalty weight or the *tuning parameter*.

Training the Model

Figure 4 and Figure 5 show the lasso regularization path of candidate models $\hat{\beta}_1 \ldots \hat{\beta}_t$ that we obtained by minimizing Equation 1 over a sequence of tuning parameters $\lambda_1 < \lambda_2 < ... < \lambda_T$ with on-job training and off-job training as dependen variable respectively. The vertical axis contains different levels of $\hat{\beta}$. The horizontal axis contains different levels for λ. Each vertical section along the horizontal axis, represents one candidate model. Moving from higher to lower λ, the algorithm includes more nonzero $\hat{\beta}_k$ and becomes more complex. To find the optimal value for λ we use *5-fold cross validation*. We split the data in 5 random evenly sized subset and derive the lasso paths $\hat{\beta}_1^k \ldots \hat{\beta}_T^k$ on each of the folds but one fold to train the models. Then we use the left out fold to obtain the out-of-sample error for each candidate model. The best $\hat{\lambda}_t$ minimizes the out-of-sample error.

Figure 4: Lasso path for on-job training



describe optmal λ and p* here

Figure 5: Lasso path for off-job training

Variable Selection

present most important variables here

Out-of-Sample Evaluation Results

present out of sample performance and compare to OLS

Lasso Logistic Model

We now estimate the Lasso model for the binary oucome variables of whether or not a person received on-job training and whether or not a person participated in an off-job training. We estimate the Lasso-regularized logistic model for the probability that a person received one specific training as follows:

$$\hat{\theta}_\lambda = argmin(-l_N(\theta)) + \lambda \sum_k |\theta^k| \tag{2}$$

where $l_N(\theta))$ is the log-likelihood function

$$l_N(\theta)) = \sum_i [y_i x_i \theta - log(1 + e^{x_i \theta})]$$

$\sum_k |\theta^k|$ is the lasso penalty that shrinks coefficients of little explanatory power to zero. $\lambda > 0$ is the penalty weight.

Training the Model

Figure 6 and Figure 7 present the lasso regularization path for the logistic candidate models. The models are ordered from the most penalized to the least penalized model and the algorithm includes more non-zero coefficients in the model. We select λ by a *5-fold cross calidation* process. This yields an optimal Lasso penalty of 0.23 for on-job training and 1.62 for off-job training.

Figure 6: Logistic model: Lasso path for on-job training

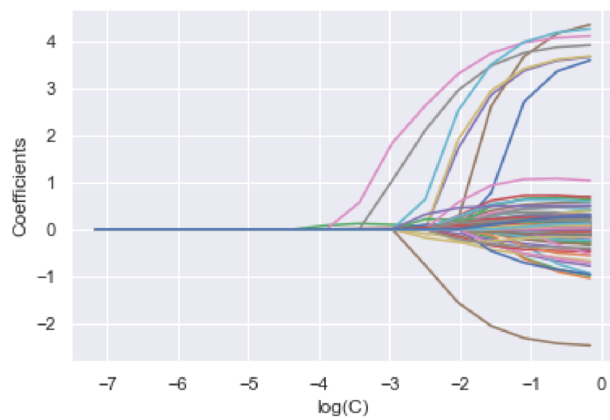
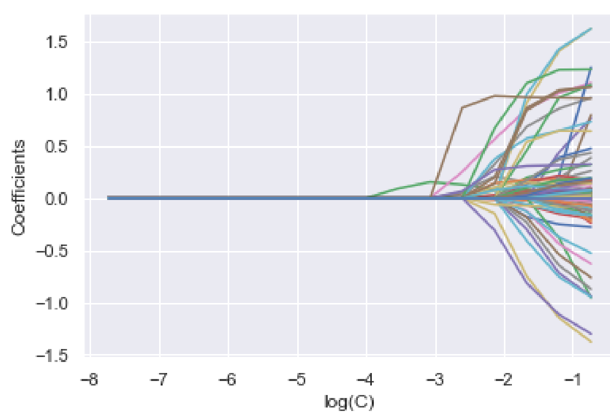


Figure 7: Locistic model: Lasso path for off-job training



Variable Selection

The Lasso logit model indentifies 131 words with nonzero predictive power for on-job training. We present the coefficients in Table . Let us first turn to the reults for the skill level. *skill_4* are elementary occupations and represent the reference group here. *skill_3* is the dummy variable for semi-skilled blue-collar occupations, *skill_2* is the dummy for semi-skilled white-collar occupations, and *skill_1* is the dummy for skilled occupations. All coefficients are nonzero. The coefficients for semi-skilled blue-collar occupations is equal to 0.17. That is, working in a semi-skilled blue-collar occupations compared to an elementary occupation increases the probability of receiving any on-job training by 17 %. The coefficent for semi-skilled white-collar occupations is lower and equal to 0.063. For skilled occupations the coefficient is 0.056. Thus,

Out-of-Sample Evaluation Results

- Describe how the hypothesis is linked to your estimation.
- Describe the estimation using equations.
- Discuss the parameters and variables.
- What are the identifying assumptions, what are (possible) violations and their consequences?
- What will you do about this?
- Tell a story!
- Guide the reader.
- Focus on the key points, not the details.
- Discuss quality and quantity.
- Discuss problems.
- Compare to the literature.

Further steps

Which results can be expected? What is new? Where lies the progress for science? In what way can scientific discussion proceed / be stimulated by the thesis?

(APPENDIX) Appendix

Appendix: Tables

Table: Lasso logistic regression for on-job training

	Coefficients	Feature
0	-0.0141972	age_r
1	0.0142358	j_q03b
2	-0.000661476	yrsget
3	0.0112553	c_q09
4	0.00519978	c_q10a
5	-0.00181812	readytolearn
6	-2.75888e-07	earnmthallppp
7	-0.0060683	g_q07_Yes
8	0.153469	d_q09_An indefinite contract
9	-0.454806	d_q09_No contract
10	0.0112322	d_q09_Other
11	0.0944897	g_q05a_Every day
12	-0.103928	g_q05a_Never
13	-0.009338	d_q13c_Every day
14	-0.096942	d_q13c_Less than once a month
15	-0.0671665	d_q13c_Less than once a week but at least once a month
16	-0.422257	d_q13c_Never
17	-0.11267	g_q08_Yes
18	0.126734	g_q05h_Less than once a month
19	-0.173587	g_q05h_Never
20	0.0869204	b_q10a_Yes
21	0.288554	d_q06a_11 to 50 people
22	0.585772	d_q06a_251 to 1000 people
23	0.490608	d_q06a_51 to 250 people
24	0.641563	d_q06a_More than 1000 people
25	-0.137498	isic1c_A
26	0.357057	isic1c_B
27	-0.156592	isic1c_C
28	0.105992	isic1c_D

	Coefficients	Feature
29	-0.174957	isic1c_F
30	-0.14426	isic1c_G
31	0.213439	isic1c_H
32	-0.434805	isic1c_I
33	-0.158933	isic1c_J
34	0.375911	isic1c_K
35	-0.276325	isic1c_M
36	0.15145	isic1c_P
37	0.20598	isic1c_Q
38	-0.00645774	isic1c_S
39	-0.381938	isic1c_T
40	-0.0910761	g_q05c_Every day
41	-0.24627	g_q05c_Less than once a month
42	0.0748321	g_q05c_Less than once a week but at least once a month
43	0.0661525	g_q05f_Every day
44	-0.191333	g_q05f_Less than once a month
45	-0.121329	g_q05f_Never
46	-0.134179	cntryid_218.0
47	-0.829642	cntryid_300.0
48	-0.0347516	cntryid_376.0
49	-0.61448	cntryid_398.0
50	-0.236664	cntryid_440.0
51	-0.150328	cntryid_484.0
52	0.0578502	cntryid_218.0
53	-0.425251	cntryid_300.0
54	-1.04733	cntryid_398.0
55	-0.0891423	cntryid_440.0
56	0.191407	cntryid_705.0
57	-0.0281003	cntryid_Belgium
58	1.07484	cntryid_Czech Republic
59	0.333032	cntryid_Denmark
60	-0.661661	cntryid_France
61	0.468554	cntryid_Ireland
62	-0.860989	cntryid_Italy
63	0.0437291	cntryid_Japan

	Coefficients	Feature
64	0.617292	cntryid_Korea
65	0.657386	cntryid_Netherlands
66	-0.374	cntryid_Norway
67	-0.0983759	cntryid_Poland
68	-0.507435	cntryid_Russian Federation
69	-0.129192	cntryid_Slovak Republic
70	0.189526	cntryid_Spain
71	0.574385	cntryid_United Kingdom
72	0.122832	gender_r_Male
73	-0.0128679	pared_At least one parent has attained tertiary
74	-0.0772567	pared_Neither parent has attained upper secondary
75	0.00628131	d_q12b_A lower level would be sufficient
76	-0.0366687	d_q12b_This level is necessary
77	-0.0194887	b_q10c_Very useful
78	-0.18683	leaver1624_Not in education, did not complete ISCED 3, aged 16 to 24
79	-2.43266	b_q14a_Yes
80	0.132303	f_q07b_Yes
81	-0.00655557	g_q05e_Every day
82	0.0316508	g_q05e_Less than once a month
83	0.0709008	g_q05e_Less than once a week but at least once a month
84	-0.0343146	g_q05e_Never
85	-0.161526	edcat8_Post-secondary, non-tertiary (ISCED 4A-B-C)
86	-0.00231736	edcat8_Primary or less (ISCED 1 or less)
87	0.0651408	edcat8_Tertiary – bachelor degree (ISCED 5A)
88	-0.0951216	edcat8_Tertiary – master degree (ISCED 5A)
89	0.169503	edcat8_Tertiary – professional degree (ISCED 5B)
90	-0.461169	edcat8_Tertiary – research degree (ISCED 6)
91	0.0406221	edcat8_Upper secondary (ISCED 3A-B, C long)
92	0.47191	b_q26a_t_Yes
93	0.0473652	d_q14_Extremely dissatisfied
94	0.00734626	d_q14_Extremely satisfied
95	-0.0450603	d_q14_Neither satisfied nor dissatisfied
96	0.00769783	g_q05g_Less than once a week but at least once a month
97	-0.250609	d_q03_The private sector (for example a company)
98	-0.015961	d_q04_t_Employee, supervising fewer than 5 people

	Coefficients	Feature
99	0.181613	d_q04_t_Employee, supervising more than 5 people
100	0.0945043	g_q06_Moderate
101	0.146899	g_q06_Straightforward
102	3.49931	b_q14b_Other
103	3.93199	b_q14b_To be less likely to lose my job
104	4.09844	b_q14b_To do my job better and/or improve career prospects
105	3.8749	b_q14b_To increase my knowledge or skills on a subject that interests me
106	3.57337	b_q14b_To increase my possibilities of getting a job, or changing a job or profession
107	4.16043	b_q14b_To obtain a certificate
108	3.47846	b_q14b_To start my own business
109	0.120491	vet_True
110	0.0777629	b_q01b_Engineering, manufacturing and construction
111	-0.0851857	b_q01b_General programmes
112	0.149636	b_q01b_Health and welfare
113	-0.0966087	b_q01b_Humanities, languages and arts
114	-0.0233597	b_q01b_Science, mathematics and computing
115	-0.0316738	b_q01b_Services
116	0.168981	b_q01b_Teacher training and education science
117	-0.0055555	d_q12c_1 to 6 months
118	0.0174142	d_q12c_3 years or more
119	-0.00150785	d_q12c_7 to 11 months
120	-0.127243	d_q12c_Less than 1 month
121	0.0527821	d_q12c_None
122	-0.166524	g_q05d_Every day
123	-0.0971816	g_q05d_Less than once a week but at least once a month
124	0.0816741	g_q05d_Never
125	0.424704	g_q04_Yes
126	-0.0648053	d_q06b_Increased
127	-0.170789	d_q06b_Stayed more or less the same
128	-0.0370546	f_q07a_Yes
129	0.0555898	skill_1
130	0.0631216	skill_2
131	0.172198	skill_3

Table: Lasso logistic regression for off-job training

	Coefficients	Feature
0	-0.0134318	age_r
1	-0.00793444	j_q03b
2	0.0188773	yrsget
3	0.0122814	c_q09
4	0.000457247	c_q10a
5	0.0541885	readytolearn
6	6.40219e-09	earnmthallppp
7	0.177043	g_q07_Yes
8	0.149206	d_q09_A temporary employment agency contract
9	-0.0939051	d_q09_An apprenticeship or other training scheme
10	0.0103396	d_q09_An indefinite contract
11	-0.0909423	d_q09_No contract
12	0.260859	d_q09_Other
13	0.0379812	g_q05a_Every day
14	0.147427	g_q05a_Less than once a week but at least once a month
15	0.00179648	g_q05a_Never
16	0.0306883	d_q13c_Every day
17	-0.0891336	d_q13c_Less than once a month
18	-0.12513	d_q13c_Less than once a week but at least once a month
19	-0.256723	d_q13c_Never
20	-0.137872	g_q08_Yes
21	0.0175554	g_q05h_Every day
22	-0.0401979	g_q05h_Less than once a month
23	-0.0782347	g_q05h_Less than once a week but at least once a month
24	-0.323316	g_q05h_Never
25	0.721732	b_q10a_Yes
26	0.0970483	d_q06a_11 to 50 people
27	0.215533	d_q06a_251 to 1000 people
28	0.178385	d_q06a_51 to 250 people
29	0.242926	d_q06a_More than 1000 people
30	-0.352446	isic1c_A
31	0.165571	isic1c_B
32	-0.181055	isic1c_C
33	0.254614	isic1c_D
34	0.141058	isic1c_E

	Coefficients	Feature
35	-0.168568	isic1c_F
36	0.0277366	isic1c_G
37	0.0381906	isic1c_H
38	-0.106975	isic1c_I
39	0.0391378	isic1c_J
40	0.387801	isic1c_K
41	0.111583	isic1c_L
42	-0.346841	isic1c_M
43	0.0992696	isic1c_N
44	0.0414967	isic1c_O
45	0.211418	isic1c_P
46	0.017236	isic1c_Q
47	-0.0455511	isic1c_R
48	-0.277199	isic1c_S
49	-0.109209	isic1c_T
50	0.274715	isic1c_U
51	0.0479809	g_q05c_Every day
52	-0.0625536	g_q05c_Less than once a month
53	-0.0940339	g_q05c_Less than once a week but at least once a month
54	-0.162329	g_q05c_Never
55	0.0603814	g_q05f_Every day
56	0.0186714	g_q05f_Less than once a month
57	-0.0784906	g_q05f_Less than once a week but at least once a month
58	-0.143191	cntryid_218.0
59	-1.80903	cntryid_300.0
60	-0.626979	cntryid_376.0
61	0.0767087	cntryid_398.0
62	0.851111	cntryid_440.0
63	-0.0737347	cntryid_484.0
64	0.274422	cntryid_152.0
65	0.0747696	cntryid_218.0
66	-0.417076	cntryid_300.0
67	-0.220985	cntryid_376.0
68	-0.332184	cntryid_398.0
69	0.843111	cntryid_440.0

	Coefficients	Feature
70	-0.0645607	cntryid_484.0
71	-1.26758	cntryid_705.0
72	-1.03884	cntryid_Belgium
73	-0.957643	cntryid_Czech Republic
74	-0.397515	cntryid_Denmark
75	-1.60505	cntryid_France
76	-1.33687	cntryid_Ireland
77	-0.323038	cntryid_Italy
78	-0.506968	cntryid_Japan
79	1.03183	cntryid_Korea
80	-0.0971461	cntryid_Netherlands
81	-1.66399	cntryid_Norway
82	0.773827	cntryid_Poland
83	-0.139928	cntryid_Russian Federation
84	-1.31597	cntryid_Slovak Republic
85	0.425831	cntryid_Spain
86	-0.888405	cntryid_United Kingdom
87	-0.190223	j_q04a_Yes
88	-0.0447	gender_r_Male
89	-0.11241	pared_At least one parent has attained tertiary
90	-0.039335	pared_Neither parent has attained upper secondary
91	-0.0808025	d_q12b_A lower level would be sufficient
92	-0.158813	d_q12b_This level is necessary
93	0.355935	b_q10c_Somewhat useful
94	-0.0315921	b_q10c_Very useful
95	-0.541456	leaver1624_Not in education, did not complete ISCED 3, aged 16 to 24
96	0.958175	b_q14a_Yes
97	0.131533	f_q07b_Yes
98	0.125695	g_q05e_Every day
99	0.0840193	g_q05e_Less than once a month
100	0.0749374	g_q05e_Less than once a week but at least once a month
101	0.11463	g_q05e_Never
102	0.267044	edcat8_Post-secondary, non-tertiary (ISCED 4A-B-C)
103	-0.0983834	edcat8_Primary or less (ISCED 1 or less)
104	0.0782849	edcat8_Tertiary - bachelor/master/research degree (ISCED 5A/6)

	Coefficients	Feature
105	0.259089	edcat8_Tertiary – bachelor degree (ISCED 5A)
106	0.184096	edcat8_Tertiary – master degree (ISCED 5A)
107	0.119397	edcat8_Tertiary – professional degree (ISCED 5B)
108	-0.256567	edcat8_Tertiary – research degree (ISCED 6)
109	0.0858208	edcat8_Upper secondary (ISCED 3A-B, C long)
110	0.295653	b_q26a_t_Yes
111	-0.20226	d_q14_Extremely dissatisfied
112	-0.0202635	d_q14_Extremely satisfied
113	-0.088988	d_q14_Neither satisfied nor dissatisfied
114	-0.0775716	d_q14_Satisfied
115	-0.0198438	g_q05g_Every day
116	0.14809	g_q05g_Less than once a month
117	0.166102	g_q05g_Less than once a week but at least once a month
118	-0.0147709	g_q05g_Never
119	-0.13778	d_q03_The private sector (for example a company)
120	0.0323726	d_q04_t_Employee, supervising more than 5 people
121	0.0596865	g_q06_Moderate
122	1.02678	b_q14b_Other
123	1.07695	b_q14b_To be less likely to lose my job
124	1.15018	b_q14b_To do my job better and/or improve career prospects
125	1.04051	b_q14b_To increase my knowledge or skills on a subject that interests me
126	1.63777	b_q14b_To increase my possibilities of getting a job, or changing a job or profession
127	1.6722	b_q14b_To obtain a certificate
128	1.29948	b_q14b_To start my own business
129	-0.00714804	vet_True
130	-0.00876664	b_q01b_Engineering, manufacturing and construction
131	-0.0890238	b_q01b_General programmes
132	0.108674	b_q01b_Health and welfare
133	-0.0865545	b_q01b_Humanities, languages and arts
134	0.0430168	b_q01b_Science, mathematics and computing
135	0.00634779	b_q01b_Social sciences, business and law
136	0.0556589	b_q01b_Teacher training and education science
137	0.0106996	d_q12c_1 to 6 months
138	-0.0562169	d_q12c_3 years or more
139	-0.055038	d_q12c_7 to 11 months

	Coefficients	Feature
140	-0.215726	d_q12c_Less than 1 month
141	0.119527	d_q12c_None
142	0.126311	g_q05d_Every day
143	-0.147083	g_q05d_Less than once a month
144	0.0695134	g_q05d_Less than once a week but at least once a month
145	0.043234	g_q05d_Never
146	0.330605	g_q04_Yes
147	0.022516	d_q06b_Stayed more or less the same
148	0.0393714	f_q07a_Yes
149	0.322029	computerexperience_Yes
150	0.00969342	skill_1
151	0.0132645	skill_2
152	-0.032086	skill_3

References

Acemoglu, D., Autor, D.H., 2011. Skills, tasks and technologies: Implications for employment and earnings, in: Handbook of Labor Economics. Elsevier, pp. 1043–1171.

Acemoglu, D., Pischke, J.-S., 1999. Beyond becker: Training in imperfect labour markets. The economic journal 109, 112–142.

Autor, D.H., Dorn, D., 2013. The growth of low-skill service jobs and the polarization of the us labor market. American Economic Review 103, 1553–97.

Autor, D.H., Levy, F., Murnane, R.J., 2003. The skill content of recent technological change: An empirical exploration. The Quarterly journal of economics 118, 1279–1333.

Becker, G.S., 1962. Investment in human capital: A theoretical analysis. Journal of political economy 70, 9–49.

De La Rica, S., Gortazar, L., Lewandowski, P., 2020. Job tasks and wages in developed countries: Evidence from piaac. Labour Economics 65, 101845.

Dietz, D., Zwick, T., 2020. The retention effect of training: Portability, visibility, and credibility1. The International Journal of Human Resource Management 1–32.

GESIS - Leibniz Institute for the Social Sciences, n.d. Programme for the international assessment of adult competencies (piaac).

Konings, J., Vanormelingen, S., 2015. The impact of training on productivity and wages: Firm-level evidence. Review of Economics and Statistics 97, 485–497.

Lynch, L.M., 1992. Private-sector training and the earnings of young workers. The American Economic Review 82, 299–312.

Lynch, L.M., 1991. The role of off-the-job vs. On-the-job training for the mobility of women workers. The American Economic Review 81, 151–156.

Tibshirani, R., 1996. Regression shrinkage and selection via the lasso. Journal of the Royal Statistical Society: Series B (Methodological) 58, 267–288.

Zweimüller, J., Winter-Ebmer, R., others, 2003. On-the-job-training, job search and job mobility. REVUE SUISSE D ECONOMIE ET DE STATISTIQUE 139, 563–576.