

A. Online supplementary material

Table A.1: List of predictors

Group	Predictor	Abbrev.
<i>Socio- demo.</i>	Age	
	Education, categorized (6 dummy variables)	
	School education, categorized (7 dummy variables)	
	State of residence (16 dummy variables)	State: x
<i>Labor market history</i>	Number of moves	
	In Employment six weeks before unemployment?	
	Subsidized employment six weeks before unemployment?	
	Long-term unemployment benefits receipt six weeks before unemployment?	Receipt LHG before
	Short-term unemployment benefits receipt six weeks before unemployment?	
	Registered as job-seeking while not unemployed six weeks before unemployment?	
	Registered with PES for other reasons six weeks before unemployment?	
	No information available six weeks before unemployment?	
	Number of employers worked for	
	Number of jobs without any vocational training held	
	Total duration in employment	Emp. tot. dur.
	Mean duration of employment without any vocational training	
	Total duration worked in industry x (14 types of industries)	Ind: x tot. dur.
	Total duration more than one job	
	Total duration in marginal employment	
	Total duration in full-time employment	Full-time tot. dur.
	Total duration in fixed-term employment	
	Total duration in temporary employment	
	Number of ALG II benefits receipt episodes	
	Total duration of ALG II benefits receipt episodes	LHG tot. dur.
	Mean duration of ALG II benefits receipt episodes	LHG mean dur.
	Number of ALG I benefits receipt episodes	
	Total duration of ALG I benefits receipt episodes	
	Mean duration of ALG I benefits receipt episodes	LEH mean dur.
	Number of labor market program participation episodes	
	Total duration of labor market program participation episodes	ALMP tot. dur.
	Mean duration of labor market program participation episodes	
Total duration of subsidized employment episodes		
Number of job seeking episodes	Seeking tot.	
Total duration of job seeking episodes	Seeking tot. dur.	

	Mean duration of job seeking episodes	
	Industry individual worked in for the longest time (14 dummy variables)	
	Days since last employment, categorized (3 dummy variables)	
	Days since last labor market contact, categorized (4 dummy variables)	
	Days since last labor market contact (full-time), categorized (4 dummy variables)	
	Time since last unemployment spell, categorized (6 dummy variables)	
	Maximum of skill-level required for all employment episodes, categorized (4 dummy variables)	
	Total duration of employment episodes, scaled by age	Emp. tot. dur. by age
	Total duration of employment episodes with more than one job, scaled by age	
	Total duration of marginal employment, scaled by age	
	Total duration of full-time employment episodes, scaled by age	Full-time tot. dur. by age
	Total duration of fixed-term employment episodes, scaled by age	Fixed tot. dur. by age
	Total duration of temporary work episodes, scaled by age	
	Total duration of ALG II benefits receipt episodes, scaled by age	LHG tot. dur. by age
	Total duration of ALG I benefits receipt episodes, scaled by age	LEH tot. dur. by age
	Total duration of labor market program participation episodes, scaled by age	
	Total duration of labor market program participation (activation) episodes, scaled by age	
	Total duration of subsidized employment episodes, scaled by age	
	Total duration of job seeking episodes, scaled by age	Seeking tot. dur. by age
<i>Last job</i>	No information about previous jobs available	
	Duration of last job	
	More than one job at last job	
	Inflation-deflated wage of last job	Last job wage
	Type of last job	
	Type of last job missing	
	Last job was part-time	
	Last job part-time missing	
	Skill-level required for last job, categorized (4 dummy variables)	
	Last job was fixed-term	
	Last job was fixed-term, missing	
	Last job was temporary work, missing	
	Last job was temporary work, missing	
	Industry of last job (14 dummy variables)	
	Commuted for last job?	

Commuted for last job, missing
Last employment more than one job

Table A.2: Tuning grids

Model type	Hyperparameter	Values
Penalized Logistic Regression	penalty	11 , 12
	C	0.001, 0.01, 0.1 , 1, 10, 100, 1000
Random Forest	max_features	sqrt , log2
	min_samples_leaf	1, 5, 10
	n_estimators	500, 750
Gradient Boosting Machines	max_depth	3, 5, 7
	max_features	sqrt , log2
	n_estimators	250, 500, 750
	learning_rate	0.01 , 0.025, 0.05
	subsample	0.6, 0.8

Note: `scikit-learn` default settings are used for parameters not listed. Selected parameters for training the final models are marked in bold.

Table A.4: Temporal cross-validated prediction performance of selected prediction models

(a) ROC-AUC										
	2011	2012	2013	2014	2015					
LR	0.717	0.694	0.702	0.710	0.714					
PLR	0.743	0.737	0.751	0.762	0.760					
RF	0.763	0.744	0.755	0.765	0.763					
GBM	0.769	0.749	0.763	0.774	0.771					

(b) Precision at top 10% (policy 1a)						(c) Precision at top 25% (policy 1b)				
	2011	2012	2013	2014	2015	2011	2012	2013	2014	2015
LR	0.369	0.379	0.382	0.368	0.351	0.328	0.319	0.318	0.312	0.290
PLR	0.333	0.401	0.403	0.397	0.378	0.312	0.337	0.336	0.334	0.319
RF	0.391	0.406	0.406	0.400	0.388	0.339	0.346	0.338	0.340	0.323
GBM	0.405	0.413	0.416	0.421	0.399	0.347	0.353	0.352	0.347	0.330

(d) Recall at top 10% (policy 1a)						(e) Recall at top 25% (policy 1b)				
	2011	2012	2013	2014	2015	2011	2012	2013	2014	2015
LR	0.238	0.230	0.238	0.237	0.243	0.529	0.484	0.494	0.502	0.501
PLR	0.215	0.243	0.251	0.255	0.262	0.503	0.510	0.523	0.536	0.551
RF	0.252	0.246	0.253	0.257	0.269	0.546	0.523	0.525	0.546	0.559
GBM	0.261	0.250	0.259	0.271	0.276	0.559	0.534	0.548	0.558	0.572

Table A.5: Prediction performance of selected prediction models (in 2016), trained with 2015 data.

(a) Ranking performance

	ROC-AUC	PR-AUC
LR	0.695	0.253
PLR	0.756	0.298
RF	0.757	0.297
GBM	0.763	0.308

(b) Classification performance based on policy 1a

	Accuracy	Precision	Recall
LR	0.838	0.332	0.259
PLR	0.843	0.356	0.278
RF	0.841	0.348	0.272
GBM	0.844	0.360	0.281

(c) Classification performance based on policy 1b

	Accuracy	Precision	Recall
LR	0.742	0.241	0.471
PLR	0.766	0.289	0.563
RF	0.765	0.287	0.559
GBM	0.768	0.292	0.569

Table A.6: Prediction performance of selected prediction models (in 2016), trained with 2010-2015 data and with sensitive attributes.

(a) Ranking performance

	ROC-AUC	PR-AUC
LR	0.703	0.257
PLR	0.760	0.298
RF	0.764	0.312
GBM	0.771	0.326

(b) Classification performance based on policy 1a

	Accuracy	Precision	Recall
LR	0.837	0.324	0.253
PLR	0.842	0.351	0.273
RF	0.845	0.365	0.284
GBM	0.847	0.376	0.293

(c) Classification performance based on policy 1b

	Accuracy	Precision	Recall
LR	0.747	0.250	0.487
PLR	0.767	0.290	0.565
RF	0.767	0.290	0.565
GBM	0.770	0.297	0.580

Table A.7: Arithmetic means of selected features for predicted LTU episodes (in 2016), models trained with sensitive attributes.

(a) Classification based on policy 1a

	Seeking Total dur.	Employment Total dur.	Last job Wage	Age	Edu.: High	Non- German	Female	Obs.
LR	2768.84	3082.21	24.58	48.35	0.07	0.09	0.49	8972
PLR	2404.80	2390.22	24.87	45.04	0.07	0.10	0.50	8972
RF	2446.43	2663.99	28.62	45.63	0.09	0.11	0.51	8972
GBM	2288.64	2923.58	28.82	45.74	0.09	0.10	0.50	8972

(b) Classification based on policy 1b

	Seeking Total dur.	Employment Total dur.	Last job Wage	Age	Edu.: High	Non- German	Female	Obs.
LR	1738.64	2749.90	28.35	41.43	0.08	0.19	0.46	22428
PLR	1721.83	2842.33	30.51	44.75	0.09	0.15	0.49	22428
RF	1833.35	2918.22	30.88	44.10	0.09	0.12	0.49	22428
GBM	1761.22	2855.15	30.30	43.83	0.10	0.13	0.49	22428

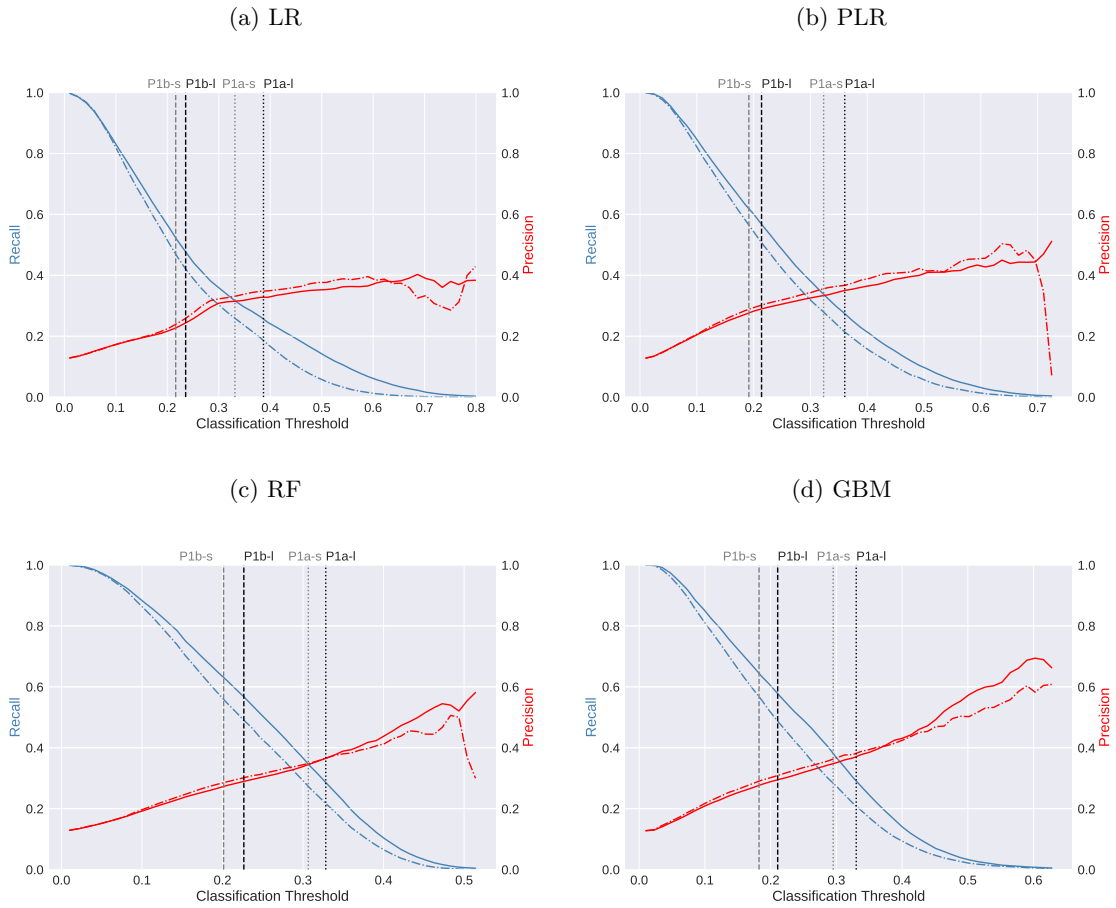


Figure A.1: Precision and recall versus threshold curves of selected prediction models (in 2016). Solid line: models trained with 2010 - 2015 data, dashed-dotted: models trained with 2015 data.