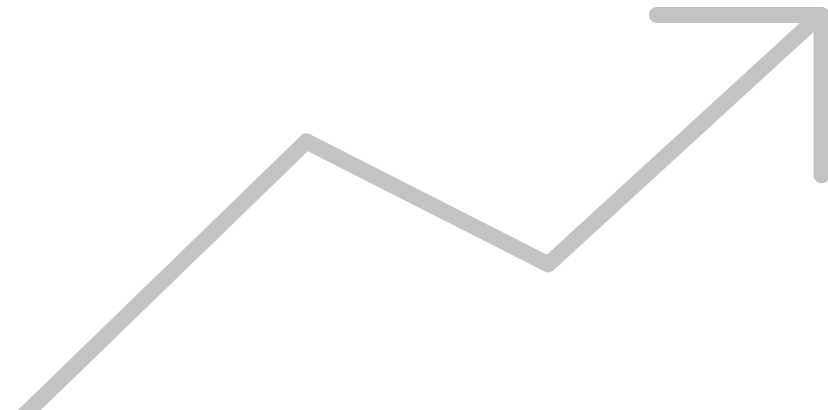


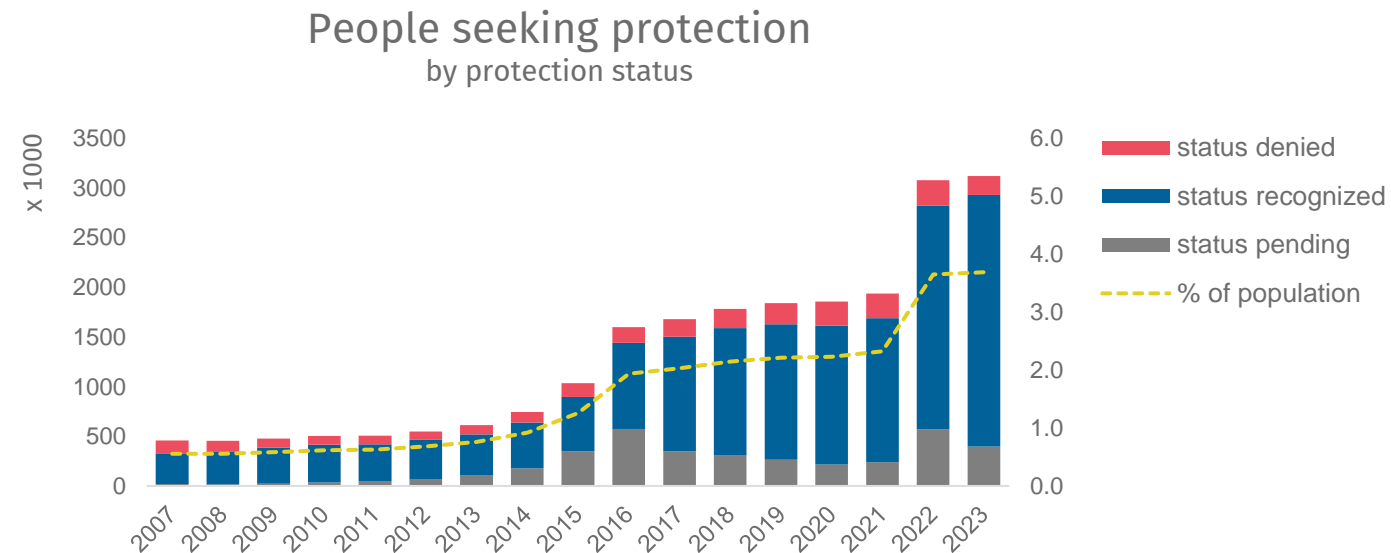
Estimating labour force participation of refugees

by statistical matching of administrative and survey data



Why statistical matching?

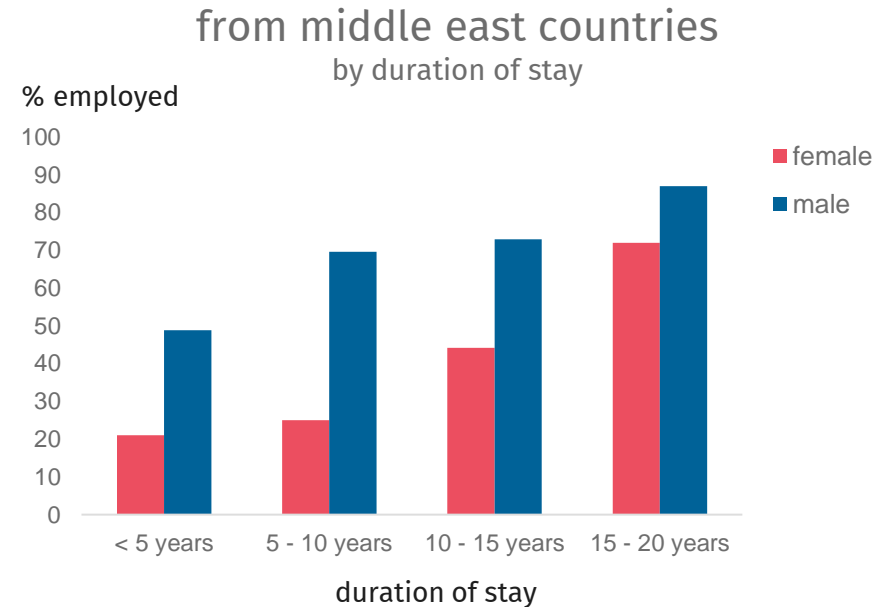
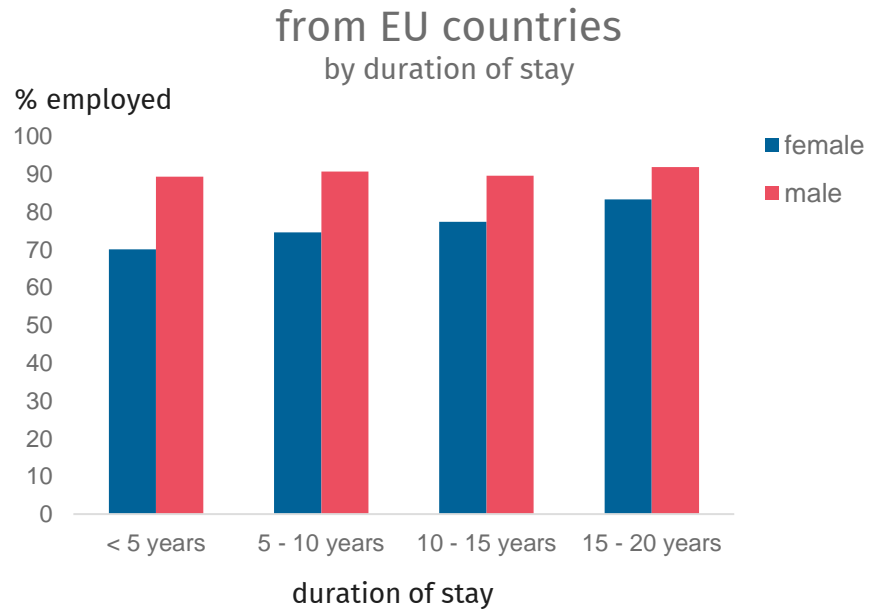
Administrative data on residence status



Source: Central Register of Foreigners (CRF)

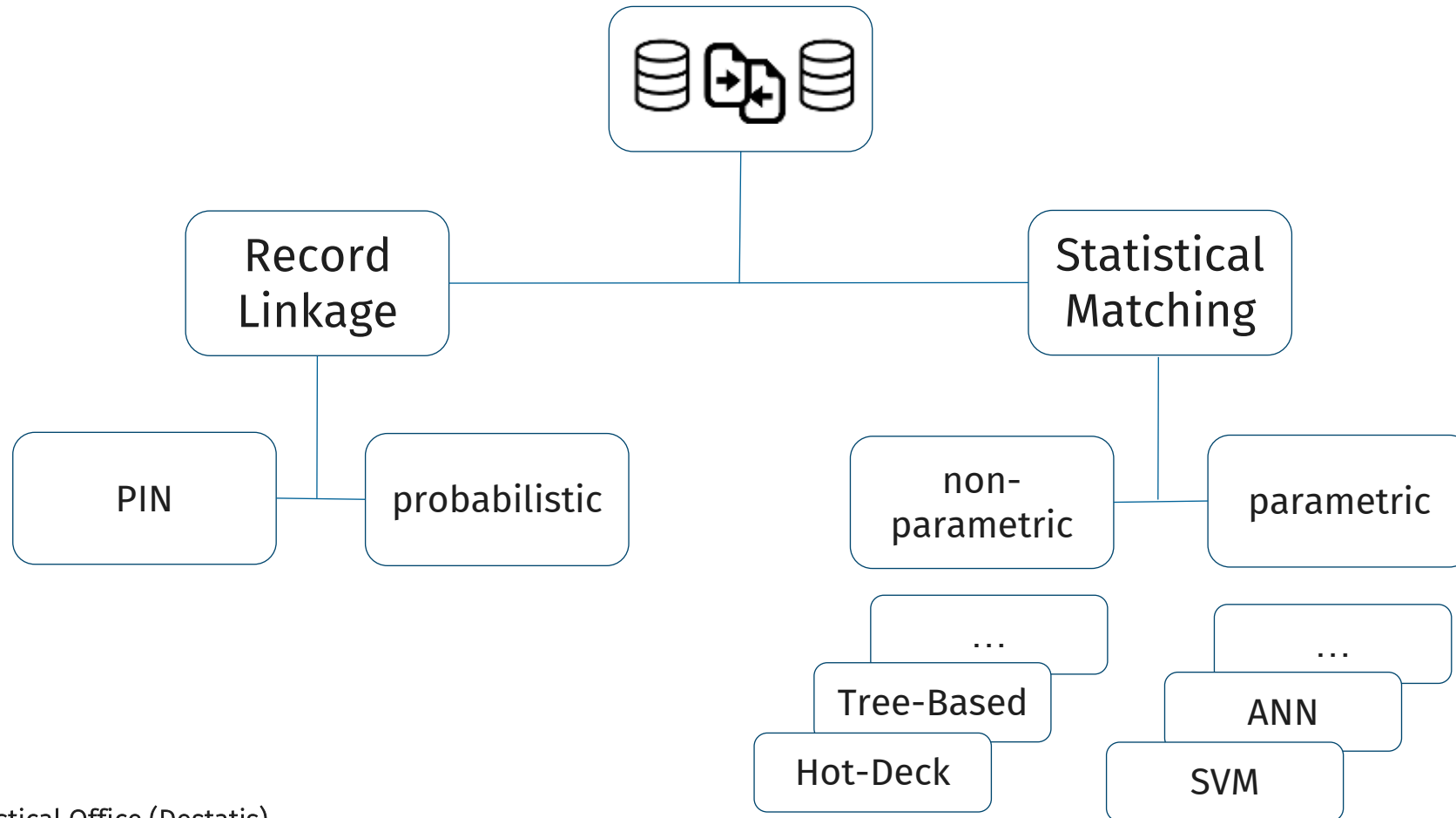
Why statistical matching?

Survey data on employment rates of immigrants

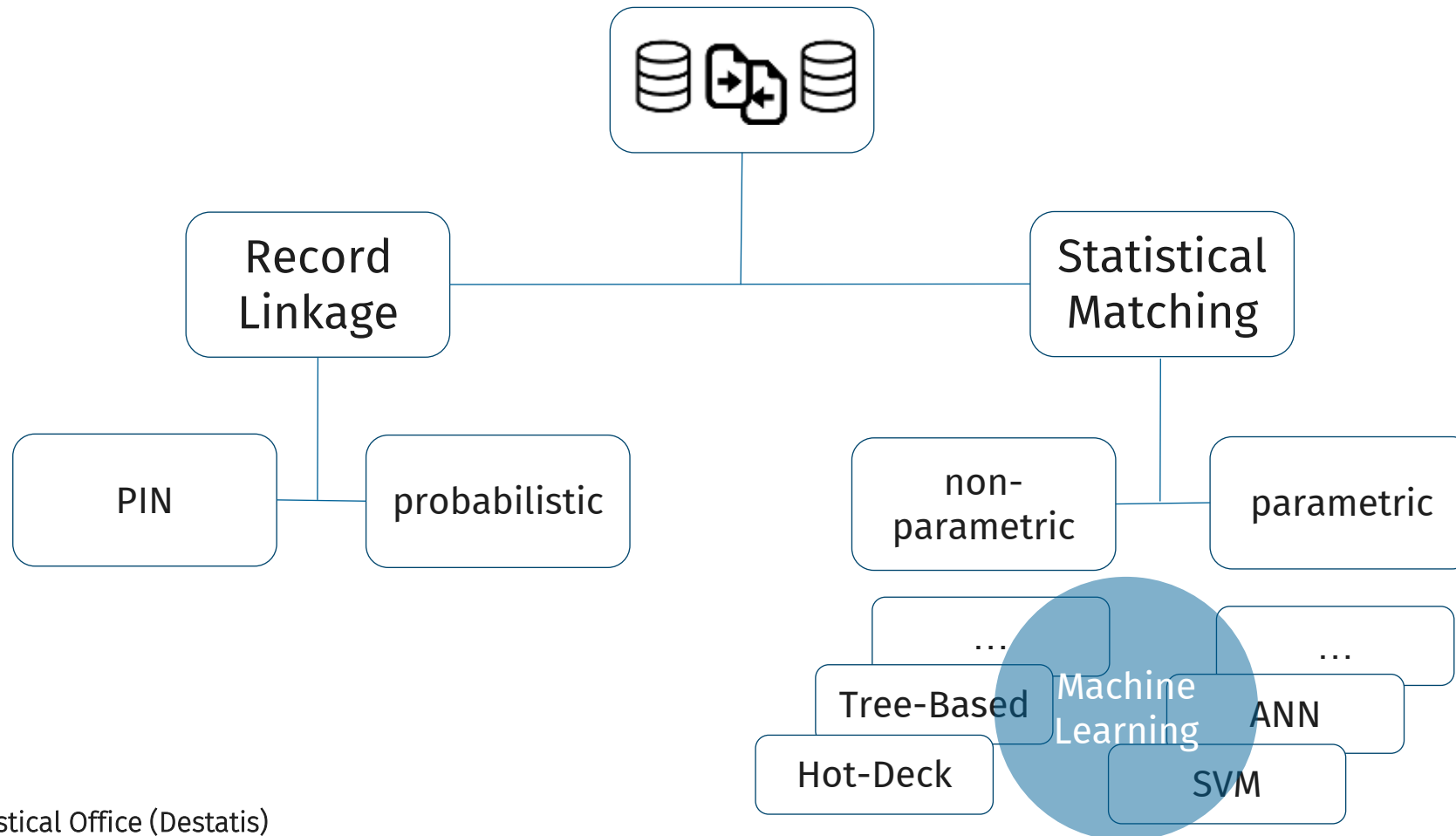


Source: Microcensus (MC)

What is statistical matching?



What is statistical matching?



What is statistical matching?

Donor		Recipient	Synthetic Dataset		
R	C	S	\hat{R}	C	S

R: Refugee Status

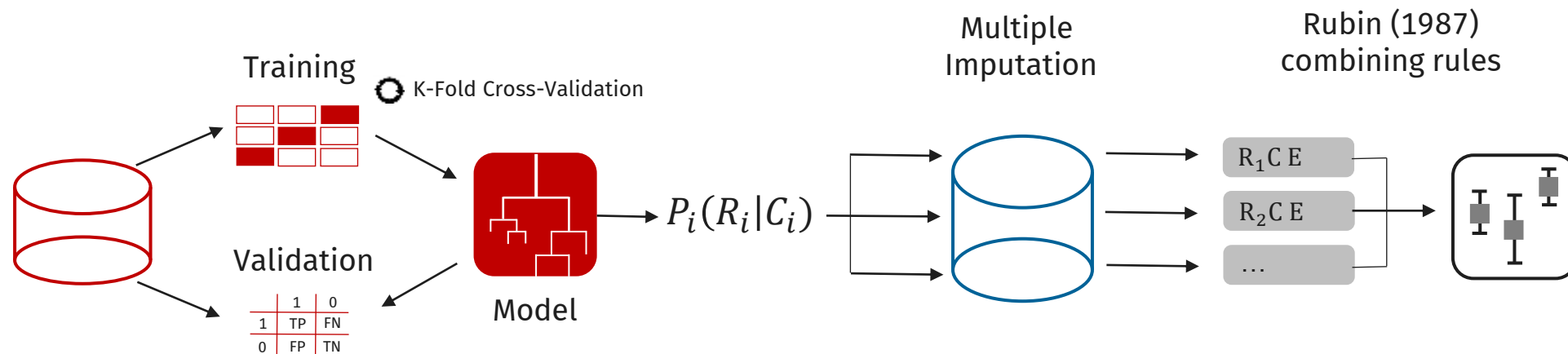
\hat{R} : Imputed Refugee Status

C: Common Variables

S: Socioeconomic Variables

Statistical matching process

- Donor
- Recipient
- Synthetic Dataset



Evaluating the estimation model

- Assumption of conditional independence

$$P(S, R | C) = P(S | C) P(R | C)$$

Evaluating the estimation model

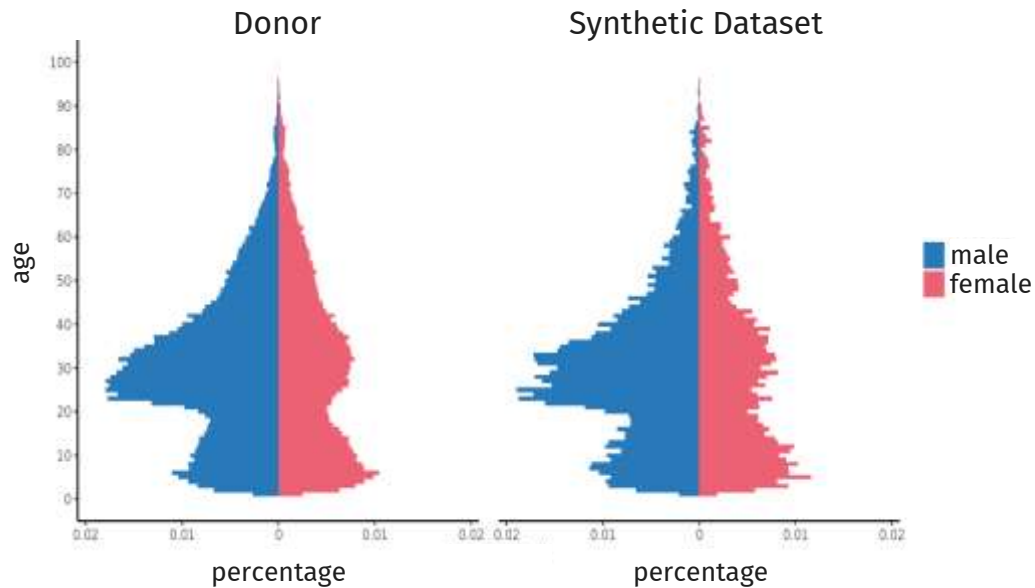
- Model: C5.0 boosted regression tree
- Explanatory power:

Confusion Matrix		True Status	
		refugee	no refugee
Prediction	refugee	141 000	43 000
	no refugee	44 000	935 000

- Accuracy = 0.93
- Precision = 0.77
- Recall = 0.76

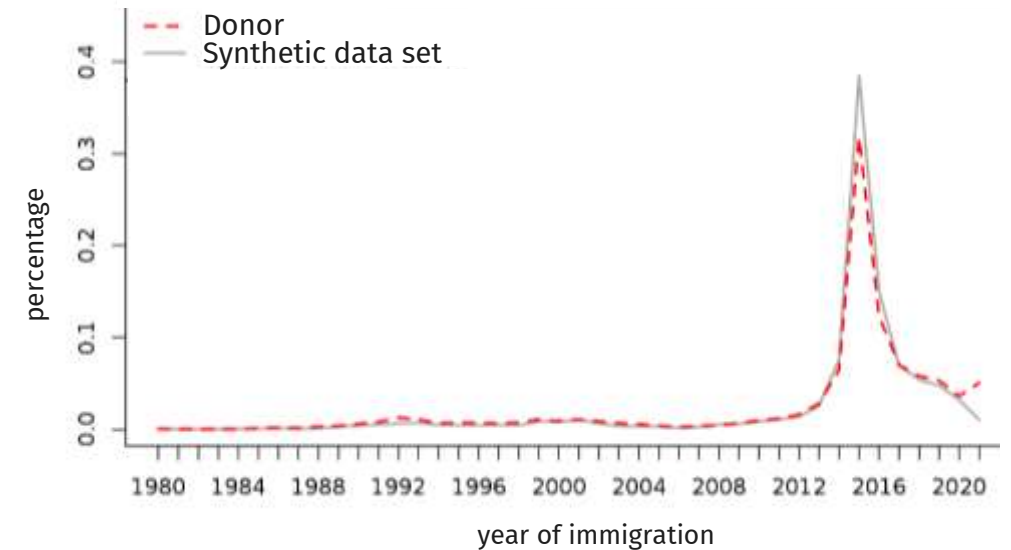
Evaluating the synthetic dataset

Evaluating reproduction of key distributions



Male: Hellinger distance 0.035 and Kolmogorov-Smirnov [p-value] 0.322

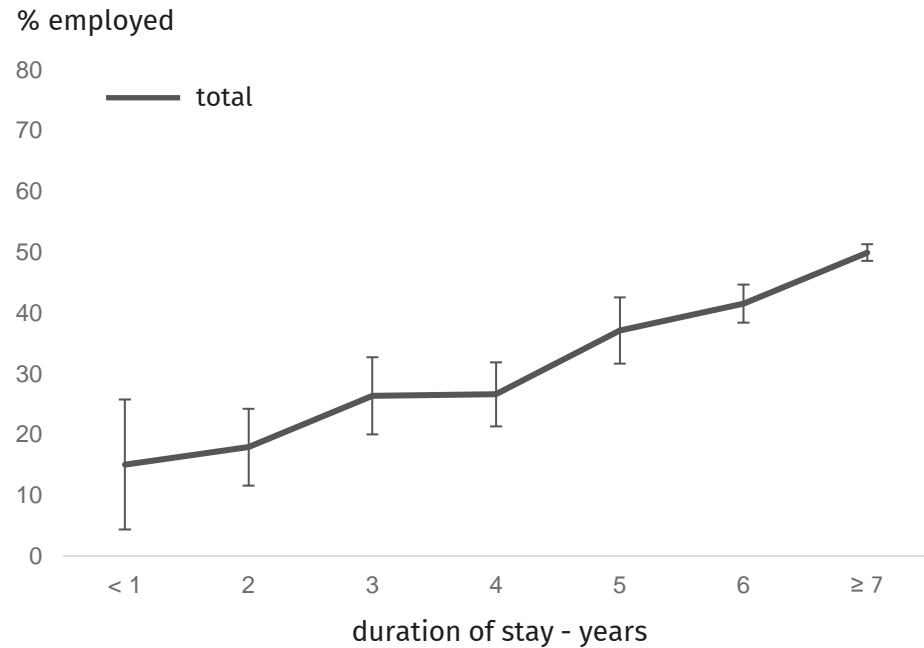
Female: Hellinger distance 0.037 and Kolmogorov-Smirnov [p-value] 0.124



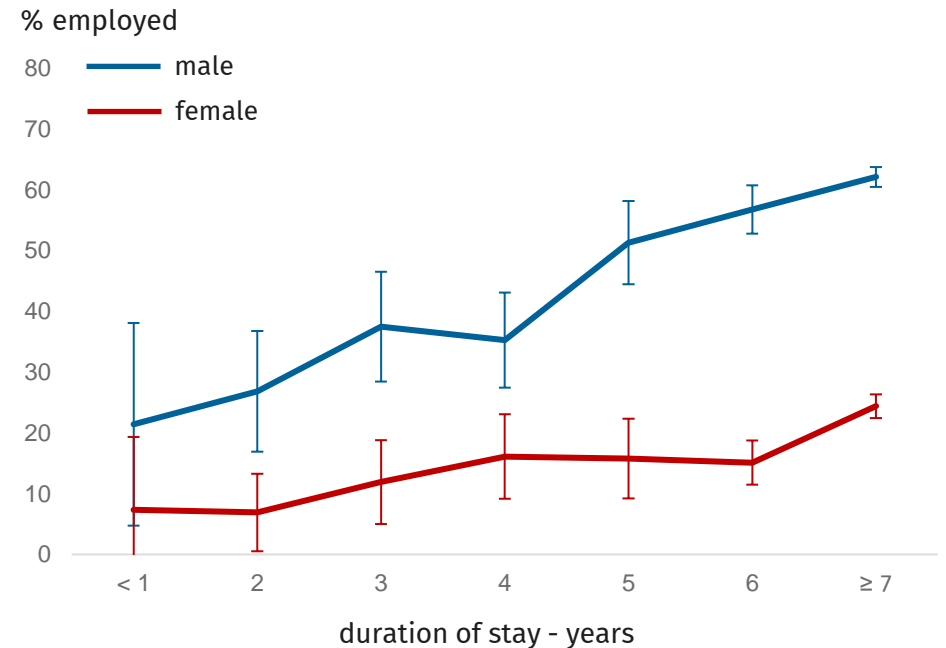
Hellinger distance 0.114 and Kolmogorov-Smirnov [p-value] 0.292

External Validation

Employment rates for refugees



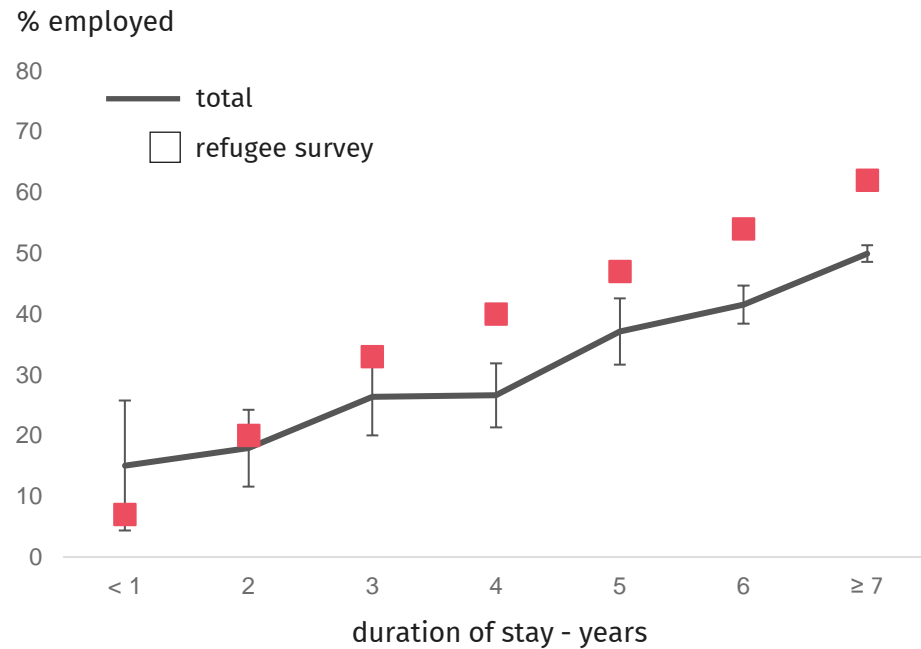
Note: The graph depicts the predicted percentage of people seeking humanitarian protection between age 18 and 65 who are employed according to the ILO-concept of employment. Error bars indicate 95 % confidence intervals. The results of the refugee survey refer to Brücker et al. (2023) available [online](#).



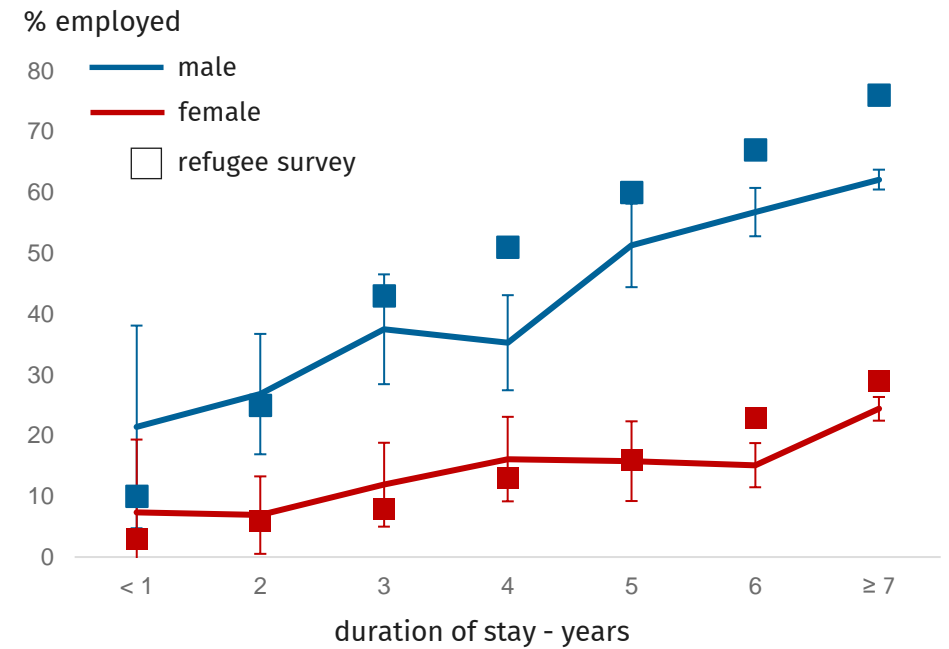
Note: The graph depicts the predicted percentage of people seeking humanitarian protection between age 18 and 65 who are employed according to the ILO-concept of employment. Error bars indicate 95 % confidence intervals. The results of the refugee survey refer to Brücker et al. (2023) available [online](#).

External Validation

Employment rates for refugees



Note: The graph depicts the predicted percentage of people seeking humanitarian protection between age 18 and 65 who are employed according to the ILO-concept of employment. Error bars indicate 95 % confidence intervals. The results of the refugee survey refer to Brücker et al. (2023) available [online](#).



Note: The graph depicts the predicted percentage of people seeking humanitarian protection between age 18 and 65 who are employed according to the ILO-concept of employment. Error bars indicate 95 % confidence intervals. The results of the refugee survey refer to Brücker et al. (2023) available [online](#).

Conclusion

- Statistical matching produces results ...
 - quickly, at low costs and without response burden
 - but with additional uncertainty
- Publication as experimental statistic



Contact

Marieke Smilde-Becker

Marieke.Smilde-Becker@Destatis.de

Jan Eberle

Jan.Eberle@Destatis.de

