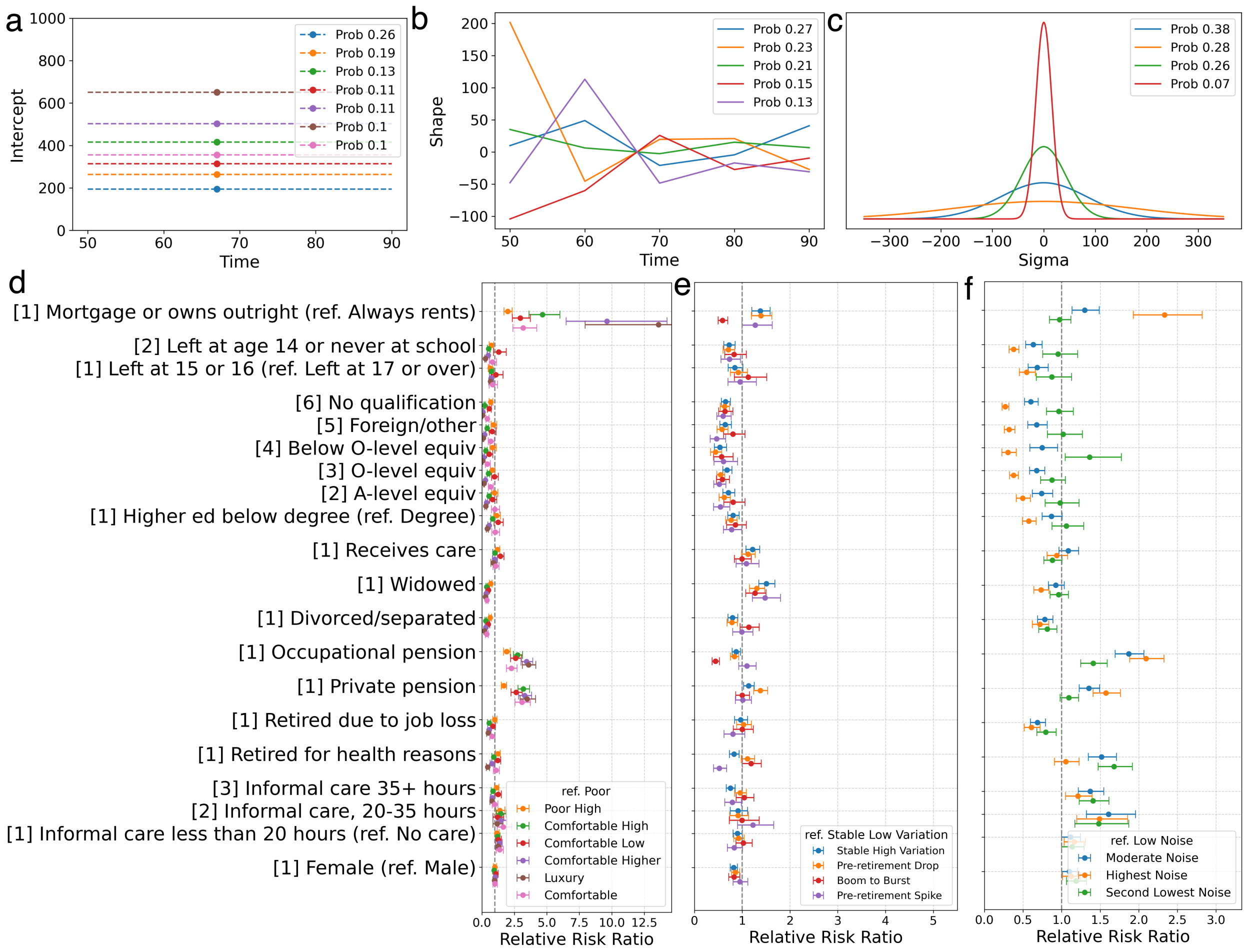

Figure 1: **a)** Seven intercept clusters broadly fall into the categories of “luxury", “comfortable" and “poor", at the retirement age of 67. **b)** Five shape clusters, “pre-retirement spike in income", “stable low variation income", “stable high variation income", “pre-retirement drop in income" and “boom to burst". **c)** Four noise clusters. Relative risk ratios (RRRs) for **d)** intercept **e)** shape and **f)** noise clusters.

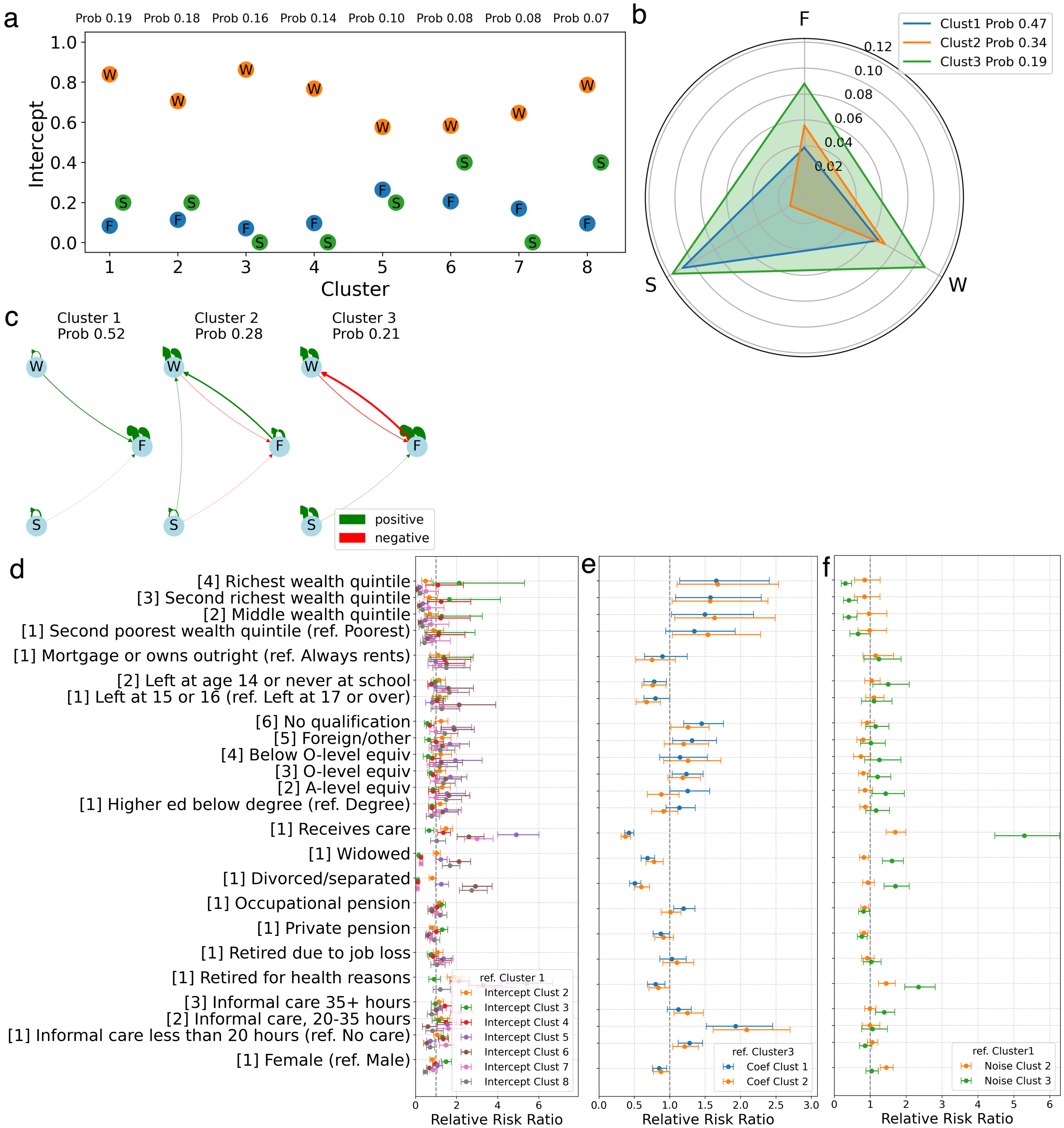

Figure 2: **a)** Eight intercept clusters. **b)** Three noise clusters. **c)** Three coefficient matrix clusters. Relative risk ratios (RRRs) for **d)** intercept **e)** coefficients and **f)** noise clusters.

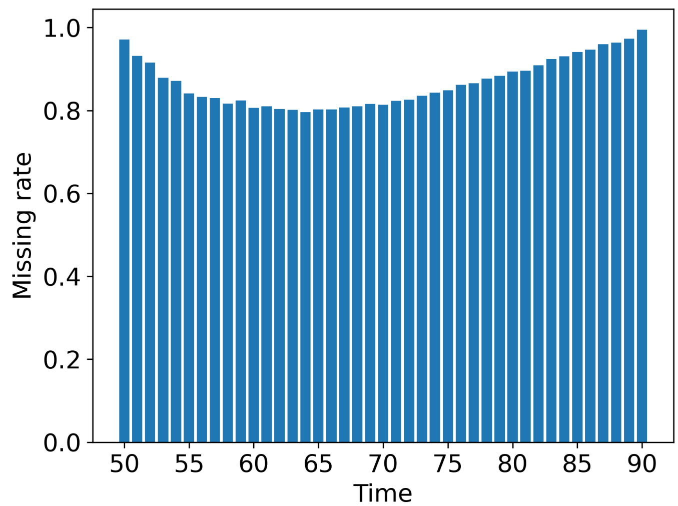

Figure 3: The missing rate of ELSA data for the NLG model at each age time point.

Table 1: Average runtime of the BMF-NLG model using different inference methods.

Size N	HMC	ADVI	MLE
250	1h	3min	30s
300	4h	10min	80s
1200	15h	1h	7min
2400	60h	2h	20min
15000	>10days	>2days	20h

Table 2: Comparison of typical clustering models with ours.

Model	Facet	Data	Method
Our nonparametric Bayesian multi-facet clustering model	Multi-facet	time series	probabilistic: Variational Bayes
Nonparametric Bayesian vector autoregression (https://jmlr.org/papers/v25/22-0717.html)	Multi-facet	time series	probabilistic: Gibbs Sampling
Nonparametric Bayesian model for multiple clustering (https://mlg.eng.cam.ac.uk/pub/pdf/NiuDyGha12.pdf)	Multi-facet	static features	probabilistic: Gibbs Sampling
Multi-facet clustering variational autoencoders (https://arxiv.org/pdf/2106.05241)	Multi-facet	images/ static features	neural network
Distance-based clustering by Dynamic Time Warping	Single-facet	time series	non-probabilistic
Mixture of nonlinear trajectories (https://doi.org/10.1007/s12062-023-09437-2)	Single-facet	time series	probabilistic: MLE