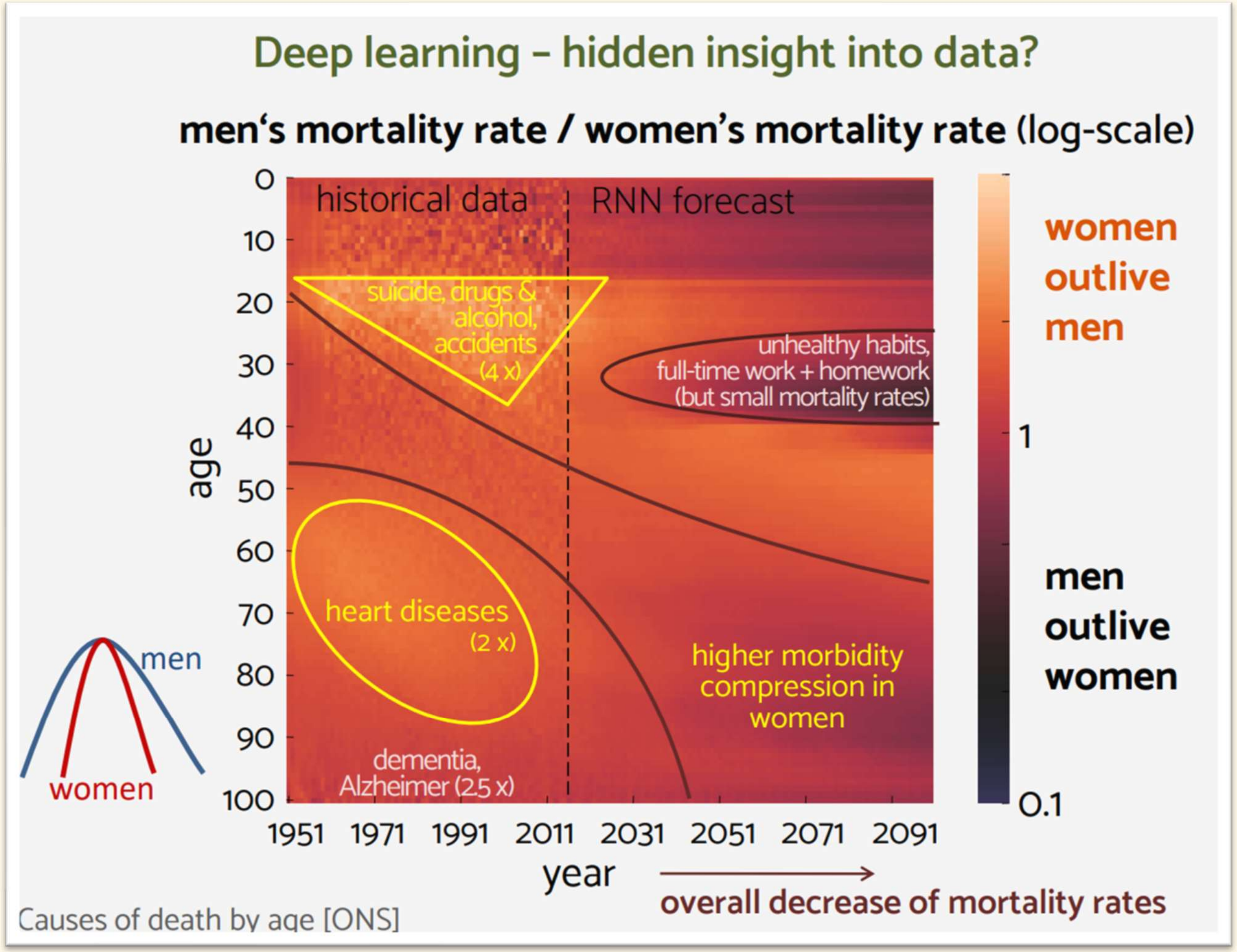
Applicability of Neural Networks in Microsimulations
A Case Study on Fertility Analysis

A. Werpachowska (Averisera Ltd & Centre for Microsimulations and Policy Analysis)

9th IMA World Congress, Wien 2024

Forecasting Mortality Rates with Neural Networks



Forecasting the impact of state pension reforms in post-Brexit England and Wales using microsimulation and deep learning (AECS & IMA2018 Narita, Japan)

Microsimulations of demographic changes in England and Wales under different EU referendum scenarios (IMA 6th World Congress, Moncalieri, Italia 2017)

Microsimulations: pre-defined well-understood relationships between individual attributes and overall population characteristics
- smoking, diet, exercise → incidence/prevalence of diseases
- energy usage, recycling habits → carbon emissions, waste generation
- fertility rate, women in childbearing age → population growth and age structure

Sometimes these relationships are very complicated, context-dependent, contain subtle interactions and dependencies between factors, as well as feedback loops and adversarial effects, which may not be easily captured by traditional analytical approaches or domain-specific knowledge.

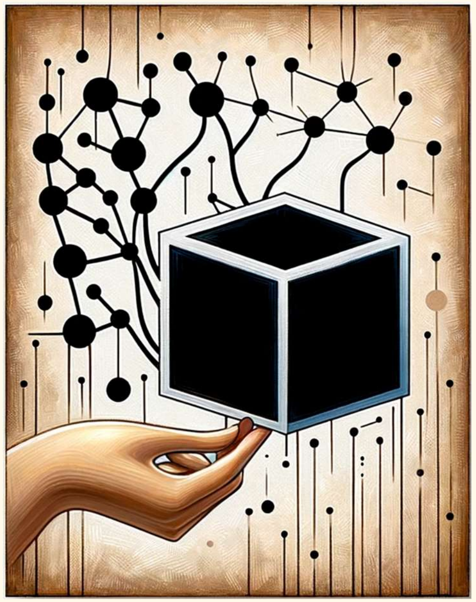
Dangers of oversimplifications:
- introducing prohibition to mitigate social problems → organized crime, bootlegging, lost tax revenue
- promotion of biofuels to reduce greenhouse effect → deforestation, increased food prices, or even higher emissions due to land-use changes and intensive farming
- pro-natalist policies in Europe: financial incentives, maternal leave extensions, childcare support → short-lived and feeble results due to economic uncertainty, youth unemployment, house prices

Microsimulations: pre-defined well-understood relationships between individual attributes and overall population characteristics

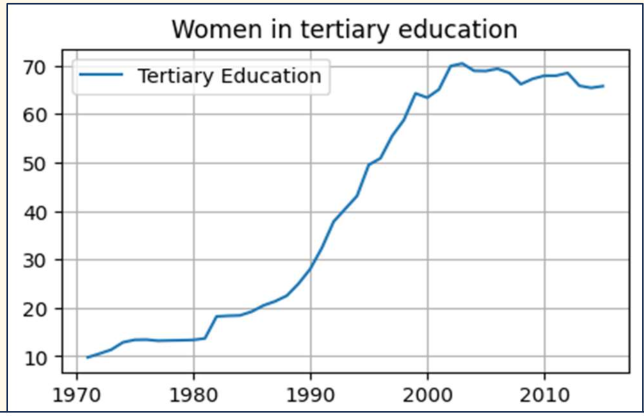
Sometimes these relationships are very complicated, context-dependent, contain subtle interactions and dependencies between factors, as well as feedback loops and adversarial effects, which may not be easily captured by traditional analytical approaches or domain-specific knowledge.

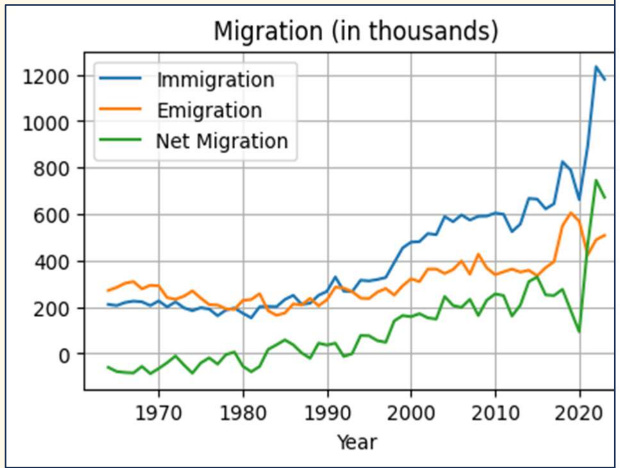
Neural Networks: dynamically learning the relationships from data

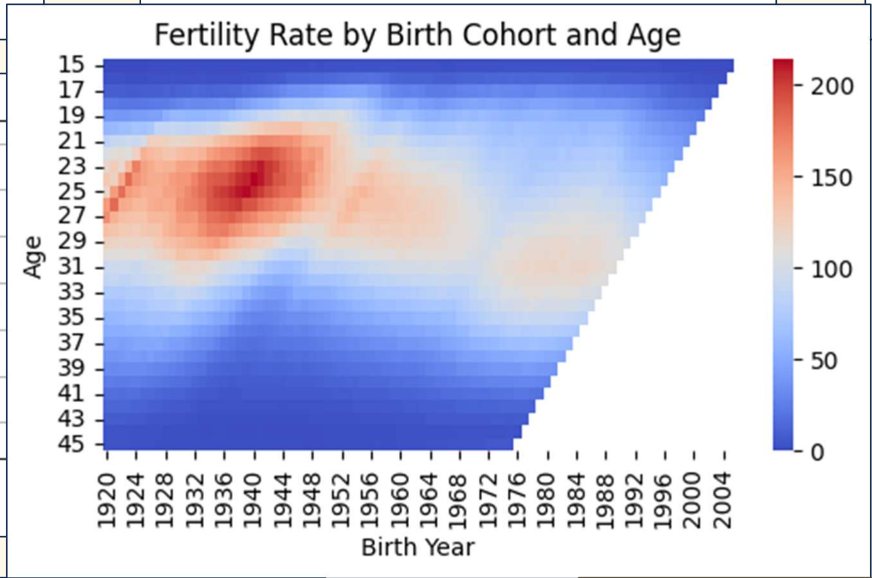
- layered structure enables to learn hierarchical representations capable of modelling the nonlinear and high-dimensional relationships
- capacity to integrate vast amounts of unstructured information (ambiguous, noisy, incomplete) to find emerging patterns
- iteratively adjusting interconnections within their architecture in response to the data's subtleties and new information (adaptivity)
- generalizing the learnt patterns to external data

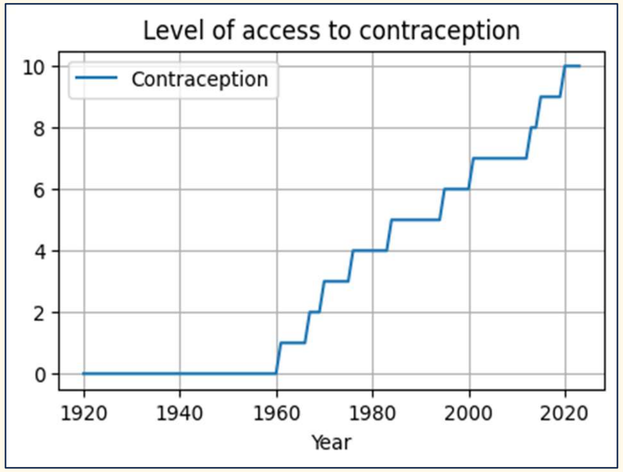


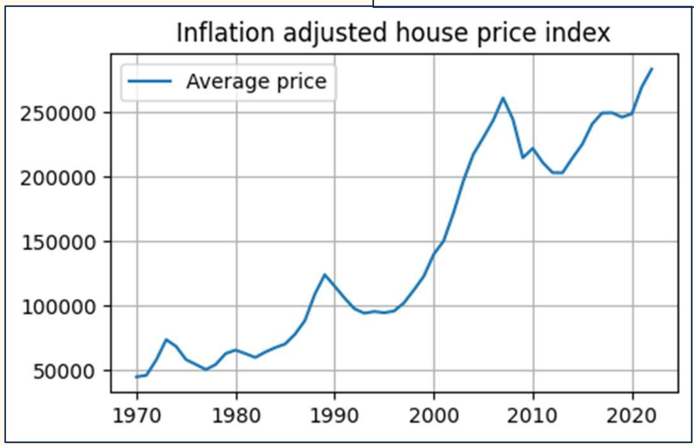
Modelling Age-Specific Fertility Rates with Neural Networks

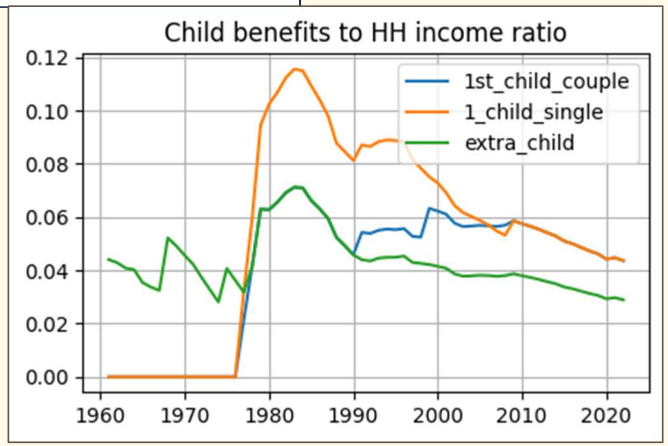




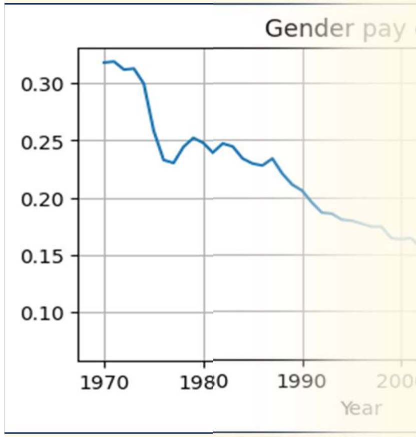


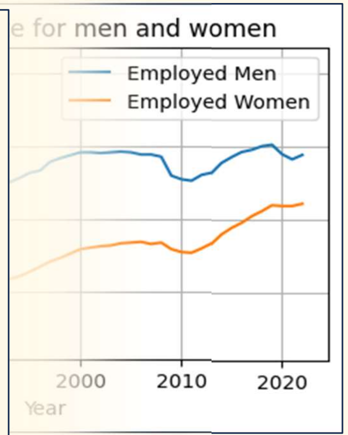






Modelling Age-Specific Fertility Rates with Neural Networks





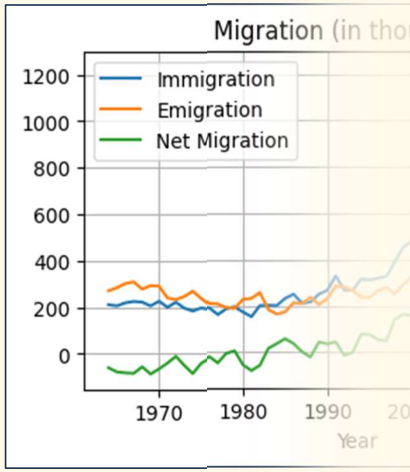
Model: LSTM (Long-Short Term Memory) deep network

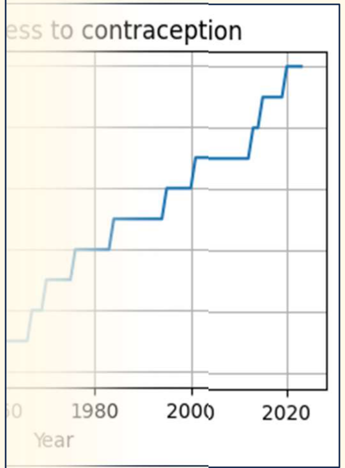
Features:

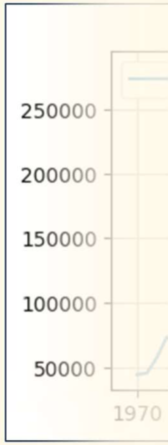
- Age
- Birth year
- Gender pay gap
- Employment rate (men & women)
- Immigration and Emigration
- House Price Index (inflation adj.)
- Tertiary education enrollment
- Child benefits: 1st ch couple, 1 ch single, extra ch
  (ratio to household income after housing cost)
- ...

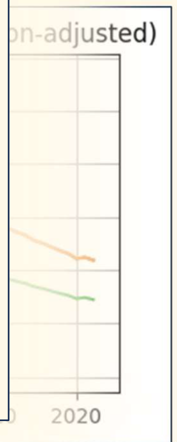
Target: Age-specific fertility rate

Sources: ONS, BoE, IFS, inflation.eu, CEIC, cosmopolitan.com

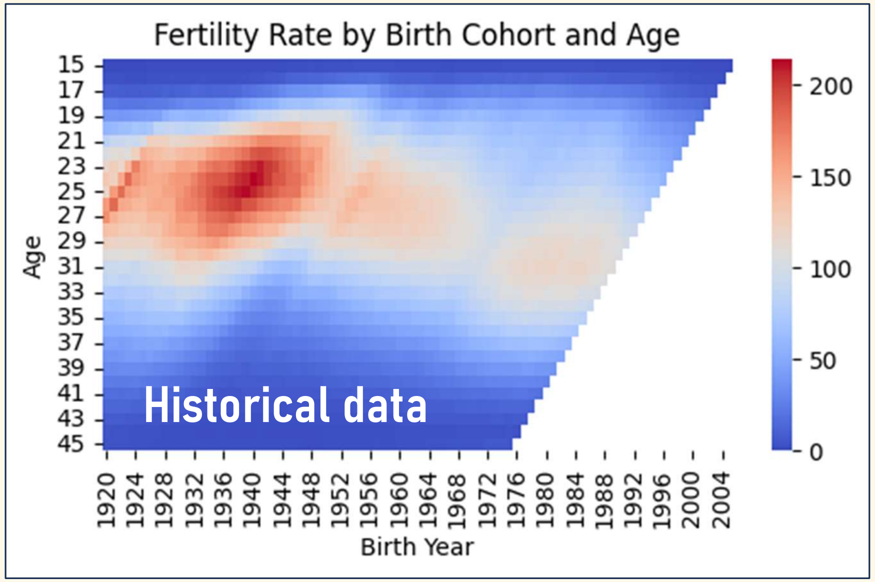


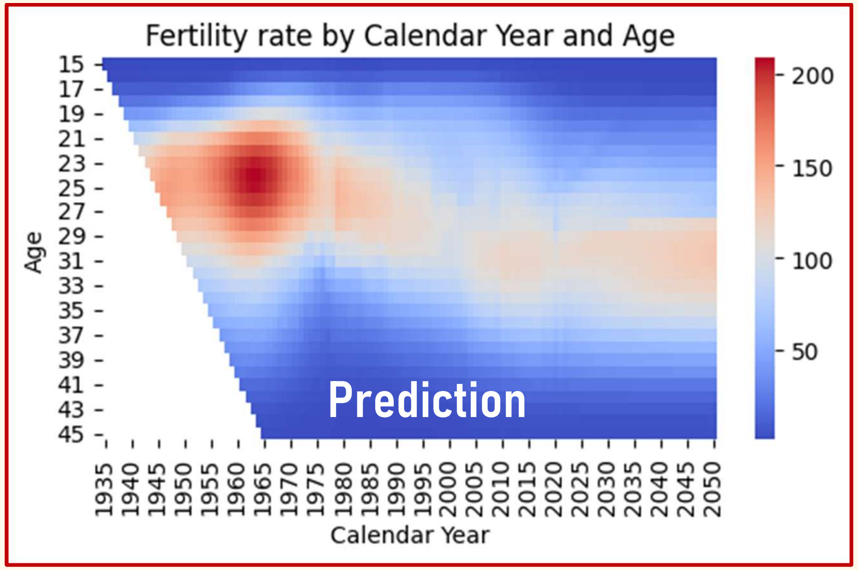


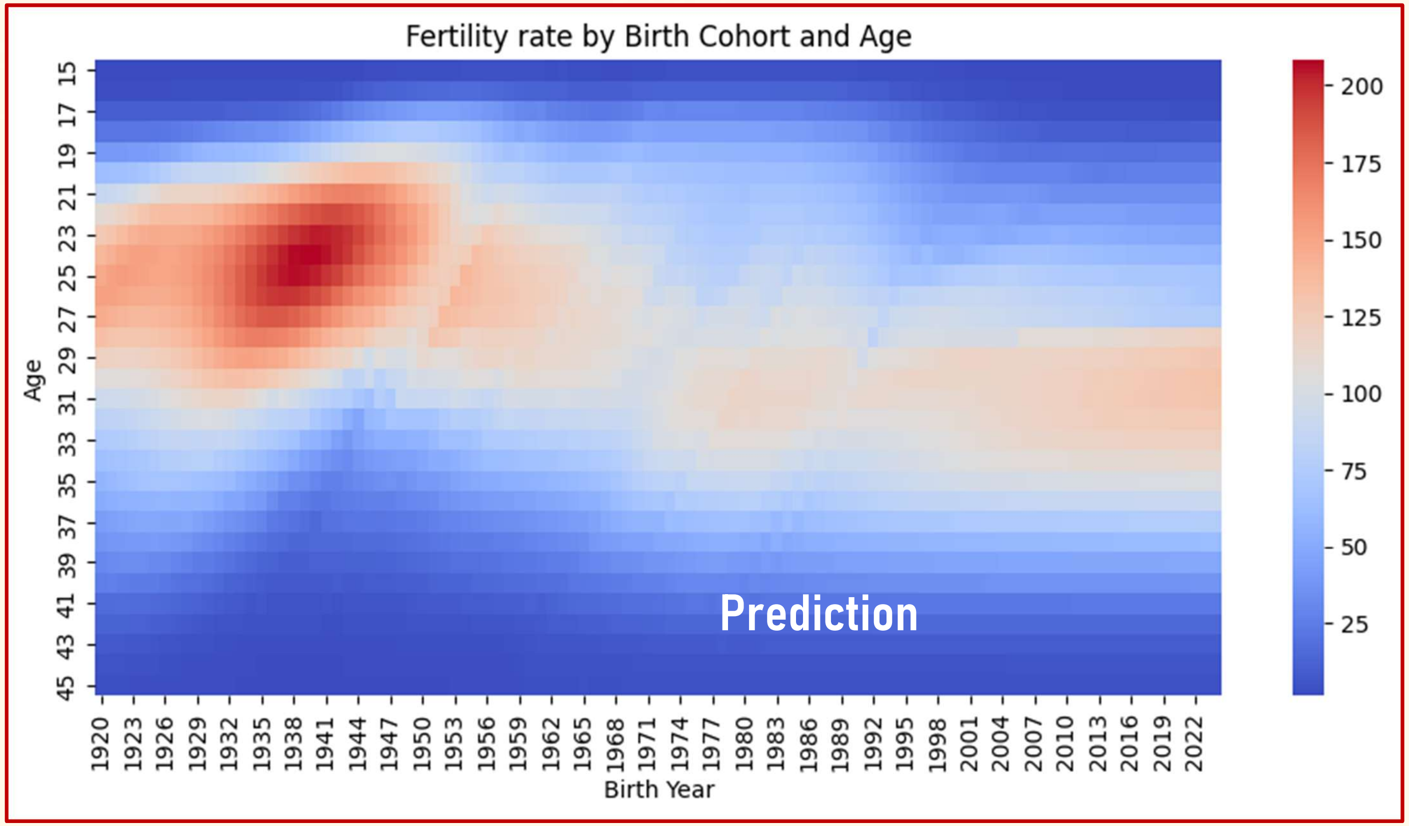




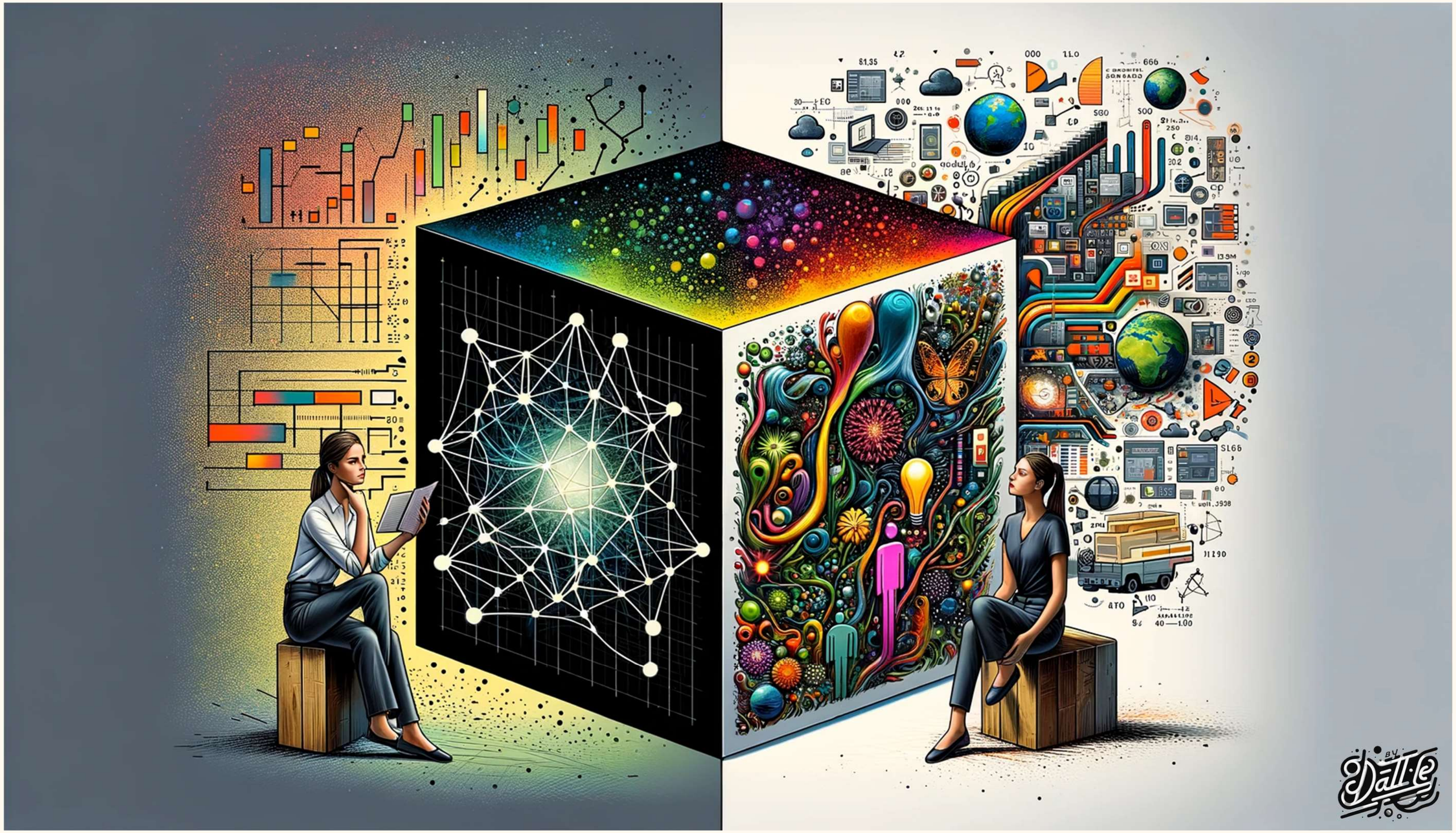
Modelling Age-Specific Fertility Rates with Neural Networks







Can neural networks, adept at capturing intricate trends and patterns, be used to discern and understand such complexities?



Explainability and interpretability of NN models

Explainability and interpretability of NN models

NNs (especially Deep Learning models) increasingly used in critical applications, like healthcare, finance or autonomous vehicles.

The need to understand how they make decisions and explain it in human terms has become crucial for validating results, ensuring fairness, transparency and users' trust, as well as fulfilling legal obligations (GDPR's right to explanation).

SHAP (SHapley Additive exPlanations): based on a simple game theoretical concept, it assigns each feature an importance value for a particular prediction/sample (as compared to an average prediction). Provides insights into how each feature contributes to the model's decision.

Layer Activation Visualization and Feature Attribution: simple methods for understanding how neural networks process inputs. Show on which part of data the model is focused.

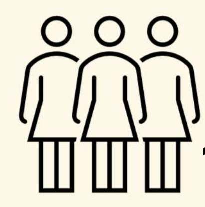
Attention Mechanisms: To which parts of the input the model pays attention.

LIME (Local Interpretable Model-agnostic Explanations): approximates the model locally with an interpretable one and explains individual predictions.

Counterfactual Explanations: How the input would need to change to obtain a different prediction. Testing hypothetical scenarios.

Model-Agnostic Methods (Partial Dependence Plots / PDP, Individual Conditional Expectation / ICE): How changes in feature values impact the predictions, regardless of the model used.

SHAP (SHapley Additive exPlanations)



1.7

37

15%

MSc

Yes

£300

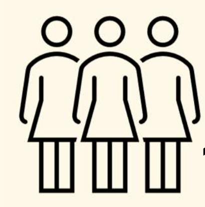
How much each feature value contributed to the 0.2 difference?

1.5

SHAP (SHapley Additive exPlanations)

		1	2	3	4	5	...
	37	✓	✓	✓	✓	✓	...
	15%	✓	×	×	✓	✓	...
	MSc	✓	✓	×	×	✓	...
	Yes	✓	✓	✓	×	✓	...
	£300	✓	×	✓	✓	×	...

All combinations (coalitions)



1.7

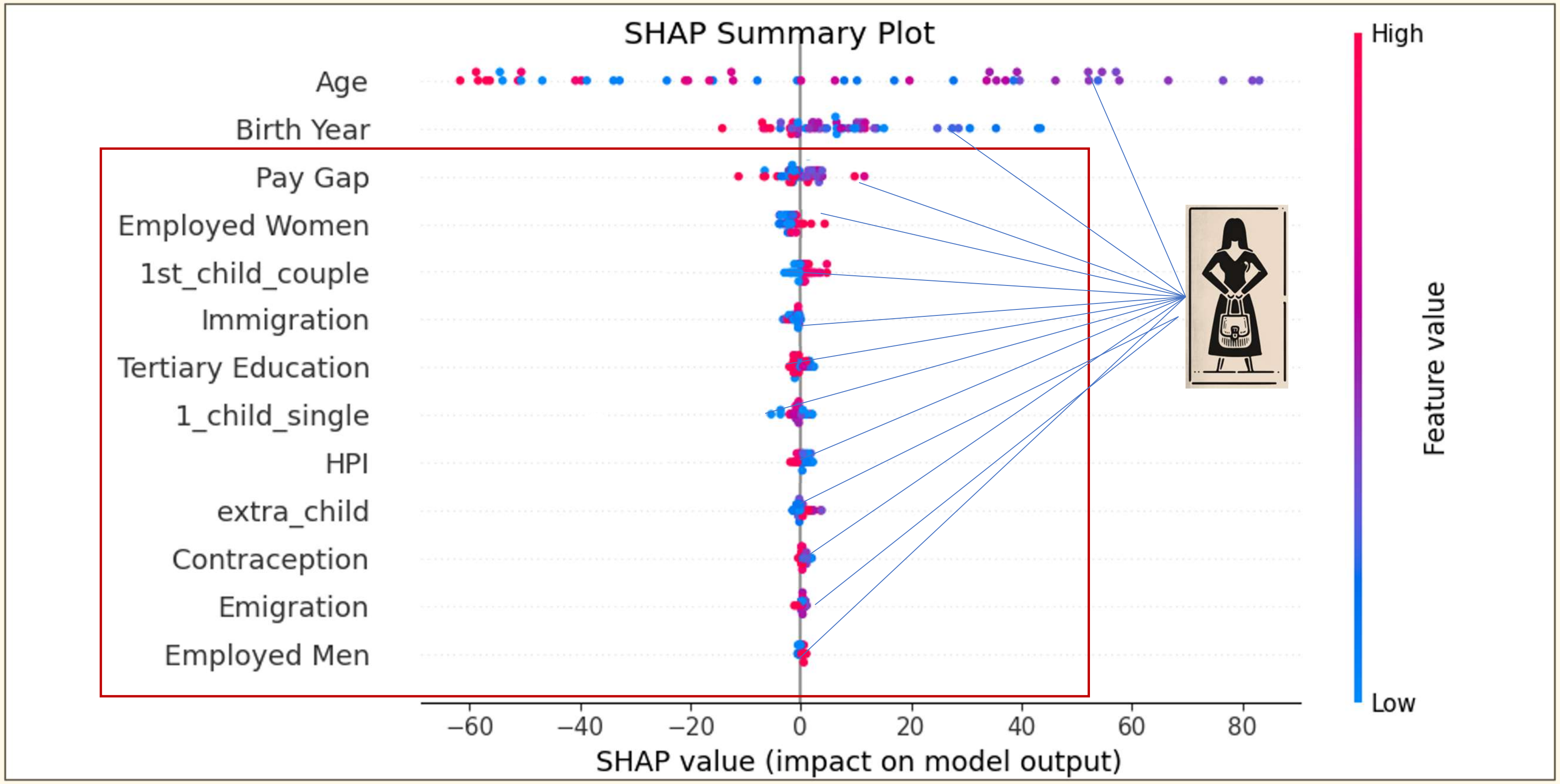
1.5

Predictions for each coalition with & w/o child benefits:

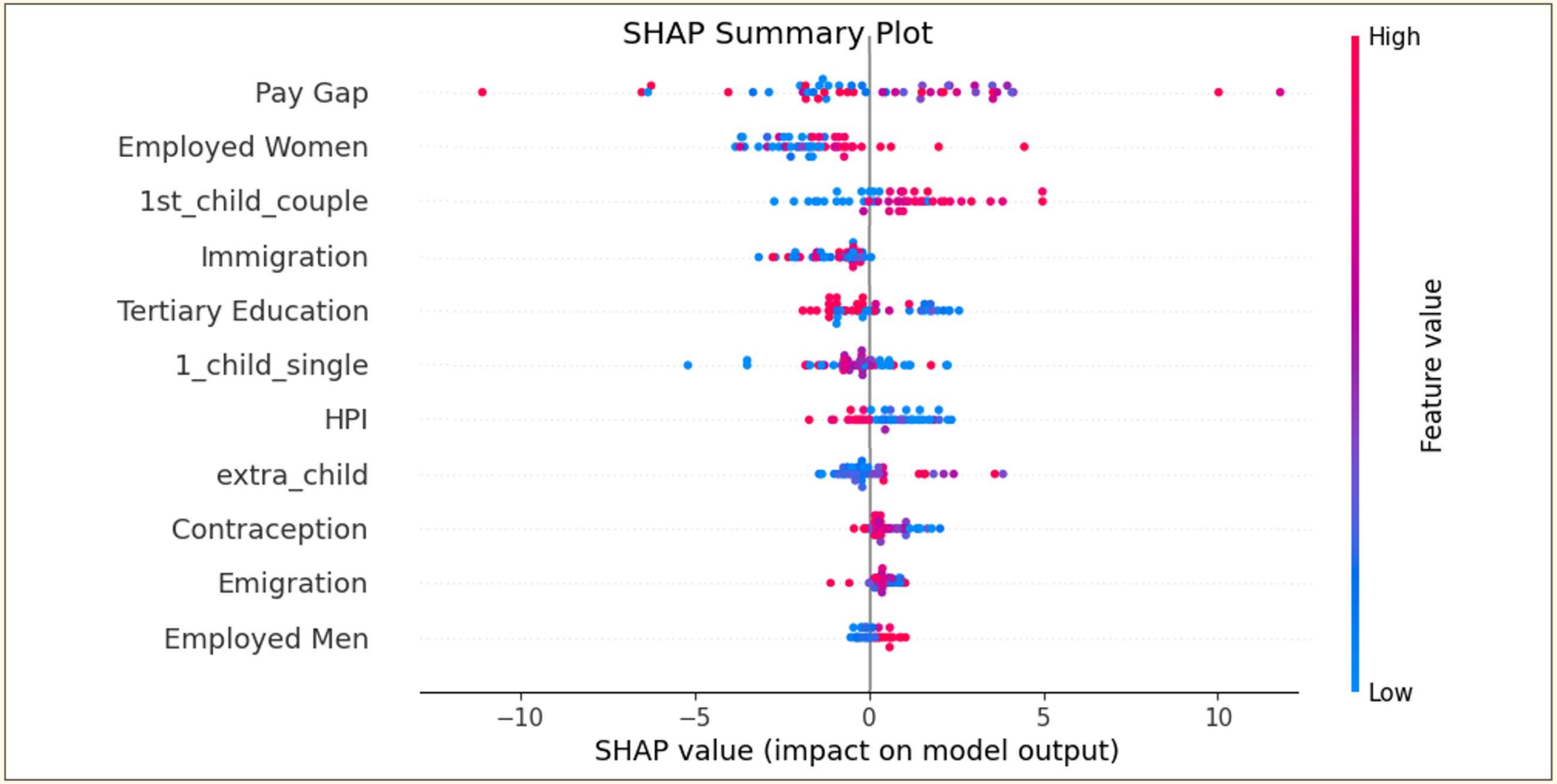
(1✓ - 1×) = 1.9-1.8
(2✓ - 2×) = 1.7-1.5
(3✓ - 3×) = ...

+0.1 children

SHAP analysis of fertility rates



SHAP analysis of fertility rates

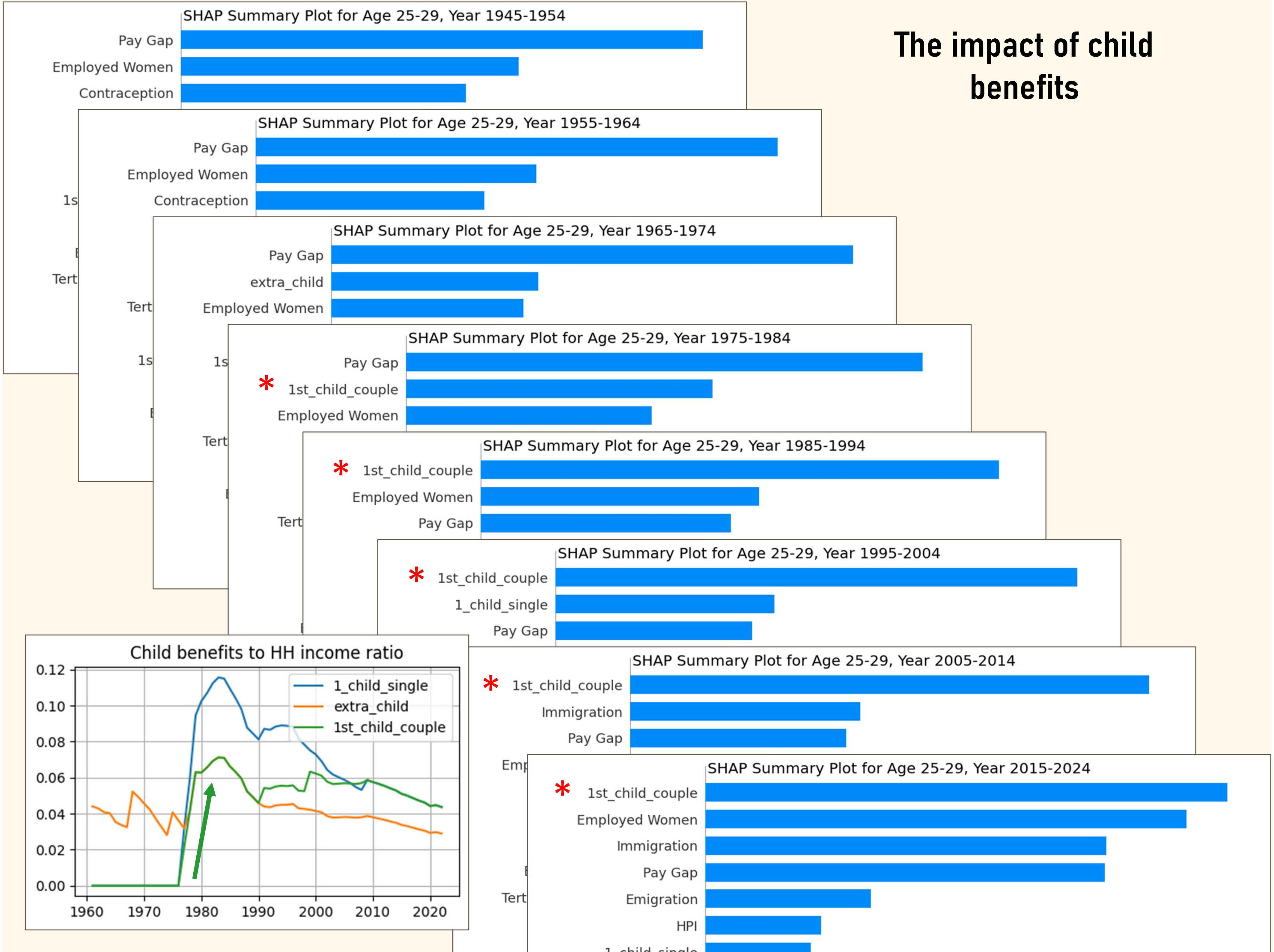


Context-dependence:

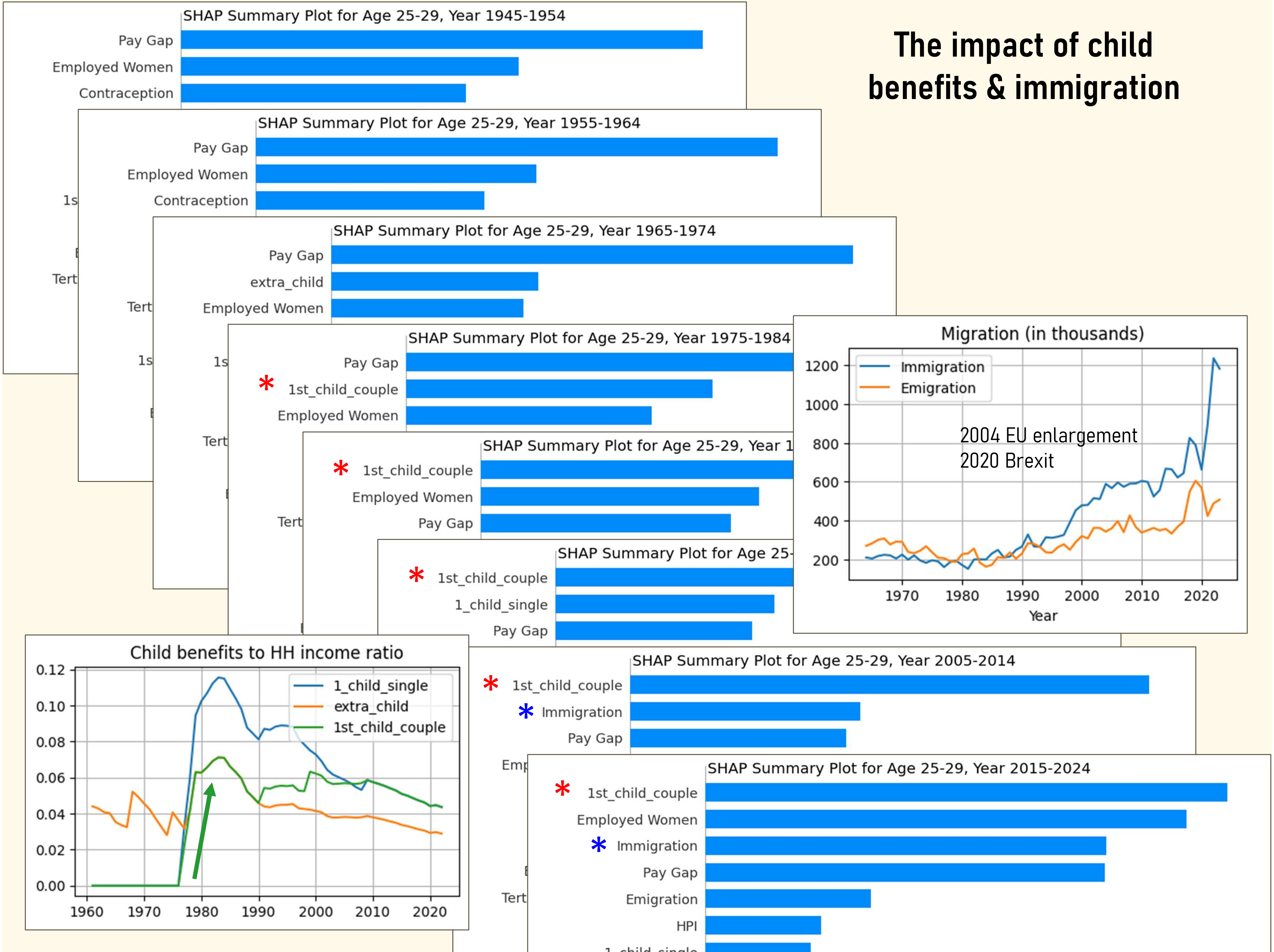
negative/positive pay gap: the lower/higher pay gap - the lower/higher fertility
negative/positive w. employment: the lower/higher - the lower/higher fertility
negative/positive child benefits: the lower/higher - the lower/higher fertility
negative/positive education: the higher/lower education - the lower/higher fertility
negative/positive house prices: the higher/lower HPI - the lower/higher fertility

Correlation, not causation!

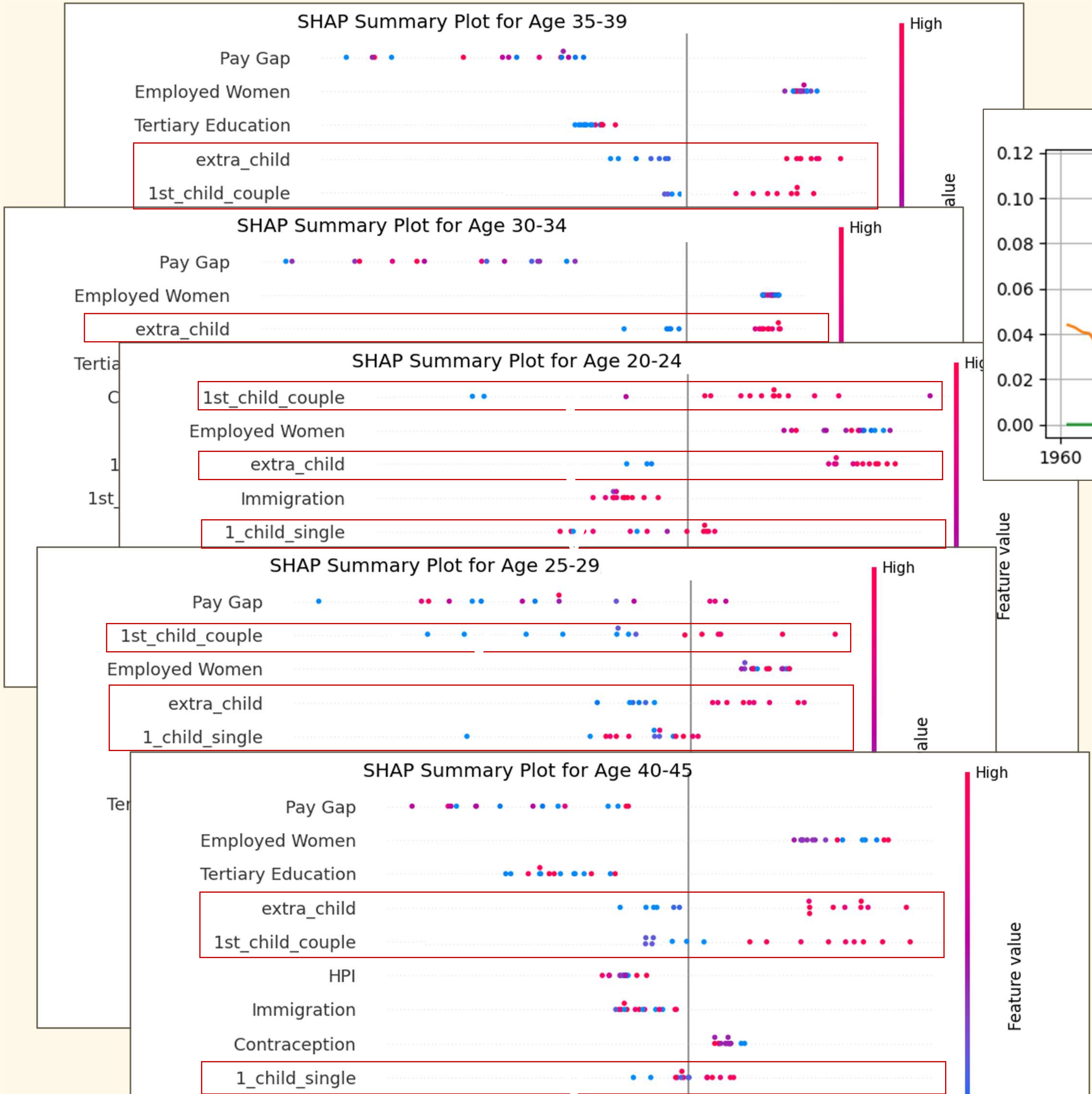
The impact of child benefits

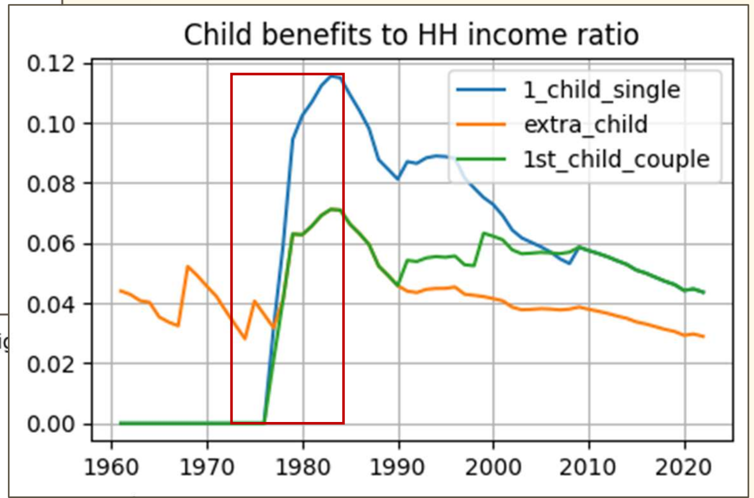


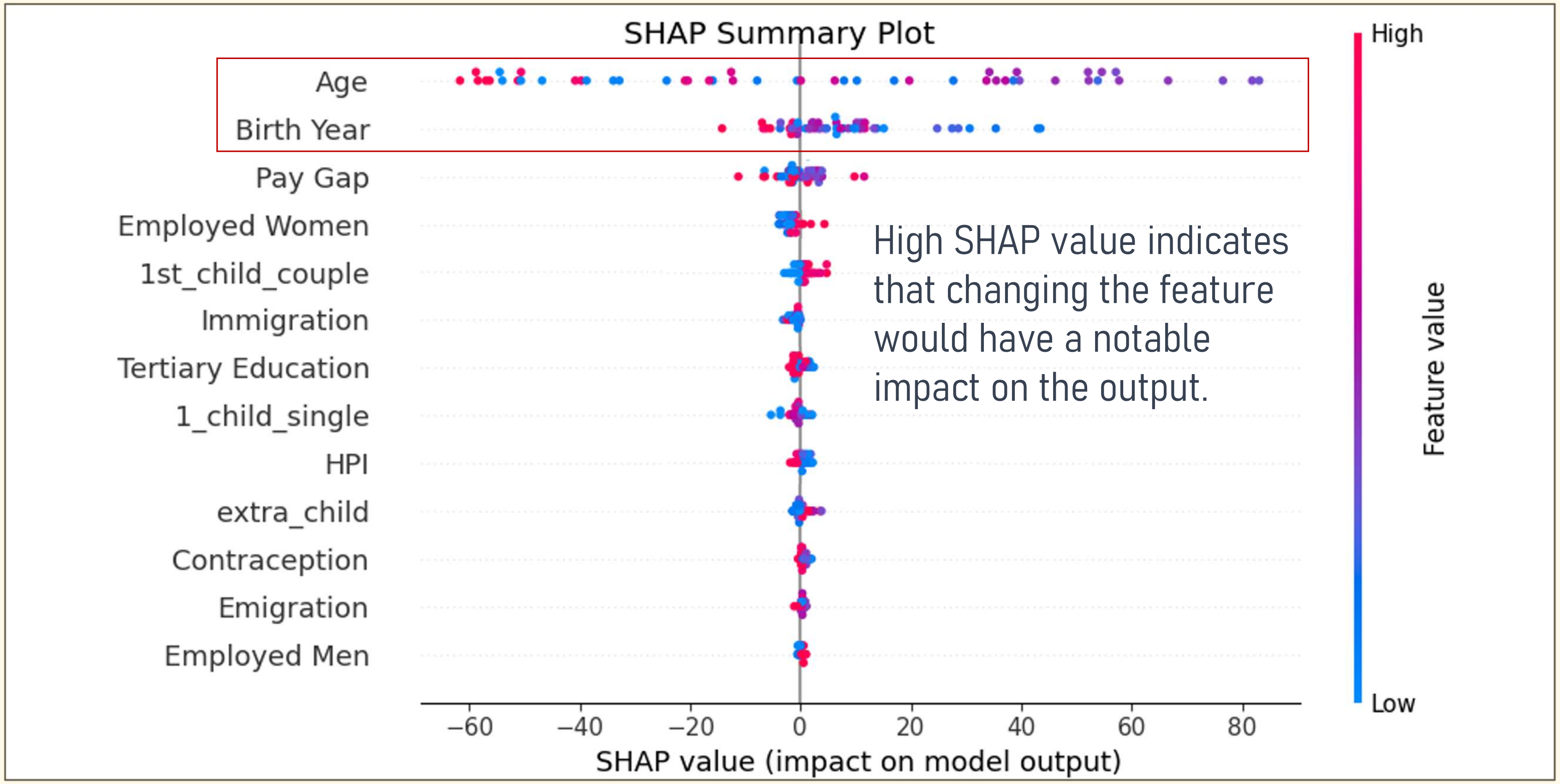
The impact of child benefits & immigration



The impact of policy change in demographic segments



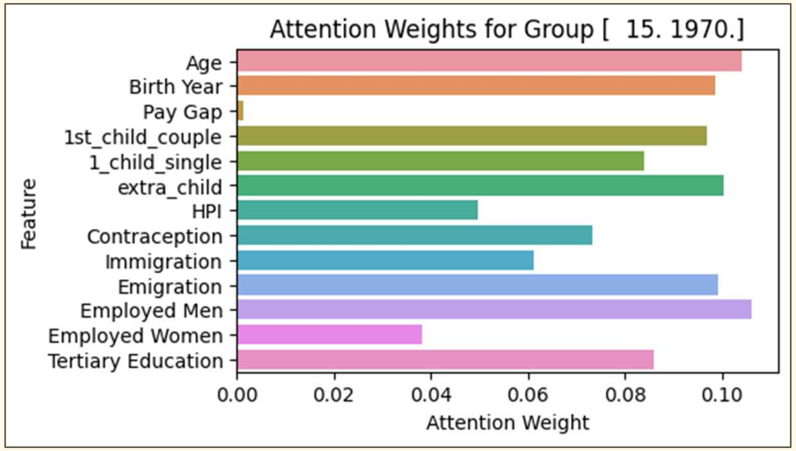


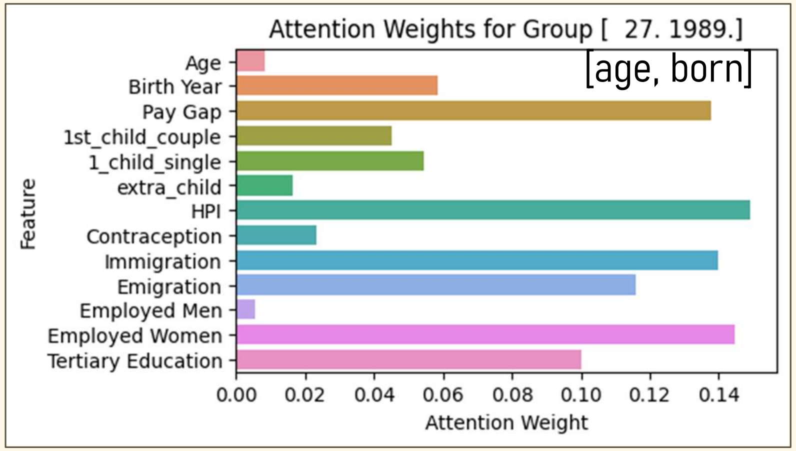


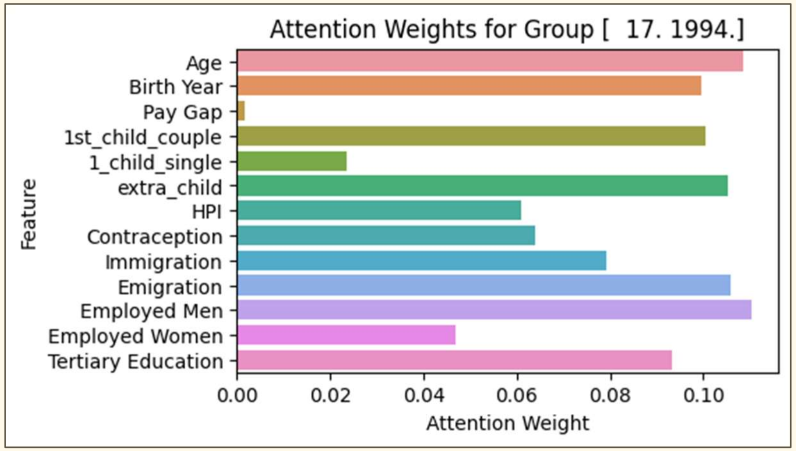
SHAP measures the impact of each feature in isolation.

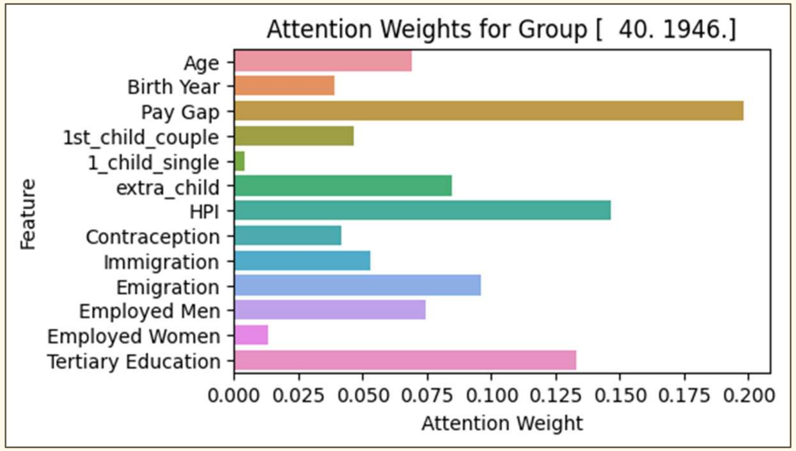
What about their interactions?

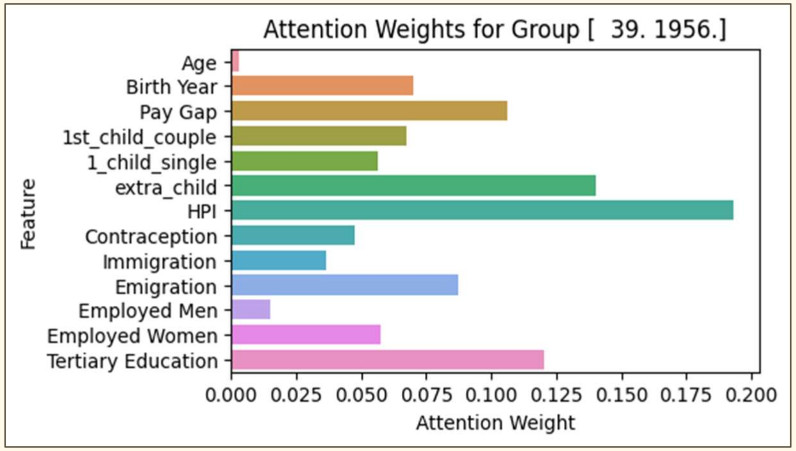
Attention Mechanism: contextual importance of features in combination

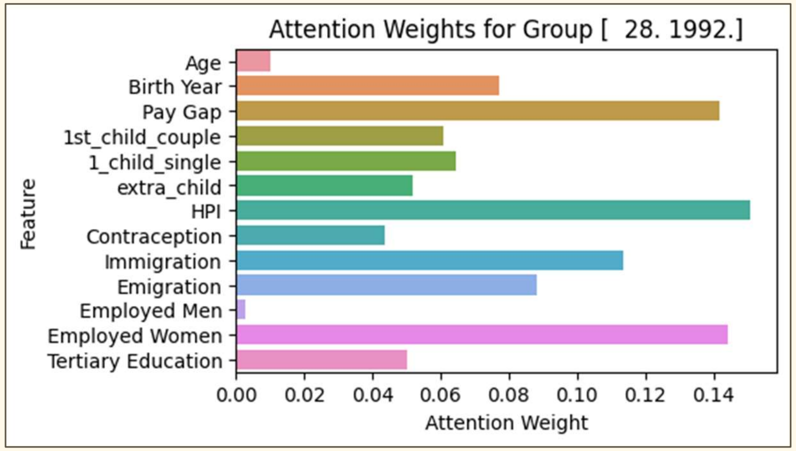




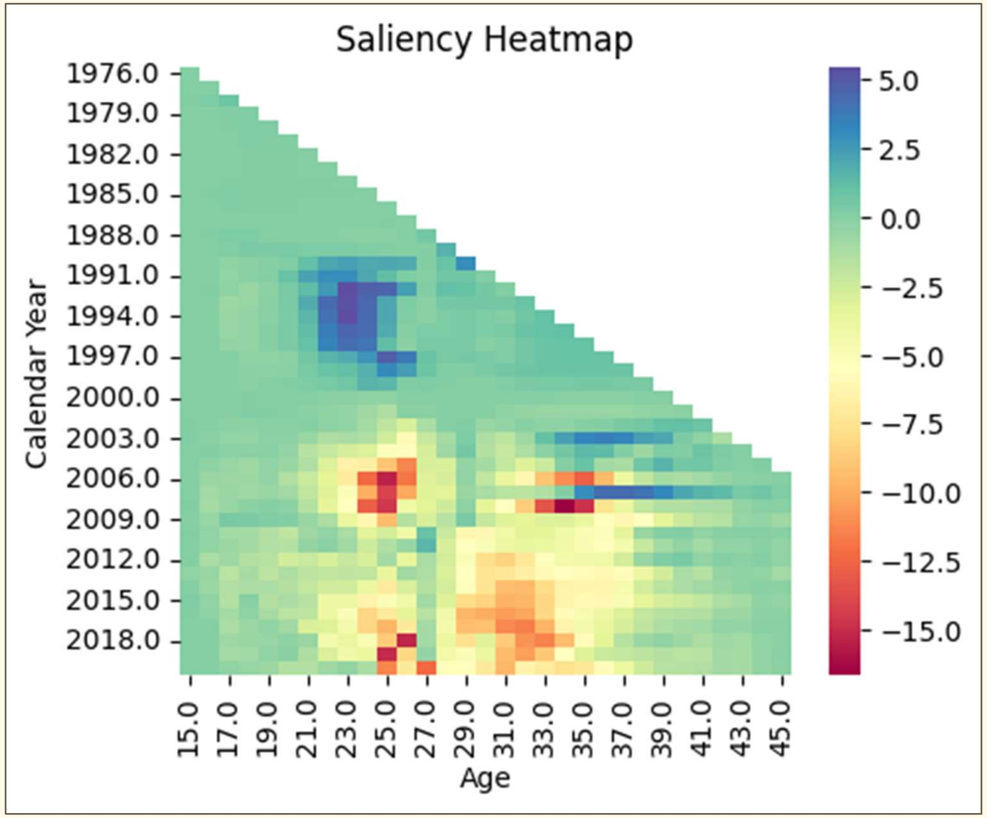






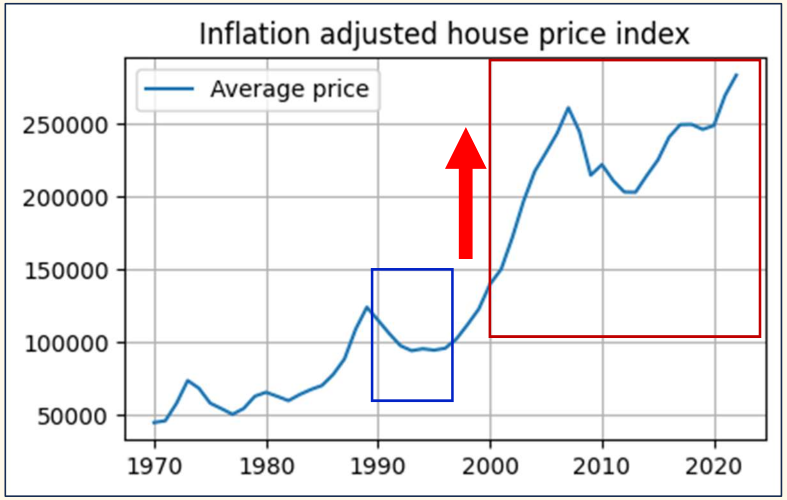


Saliency maps (sensitivity)



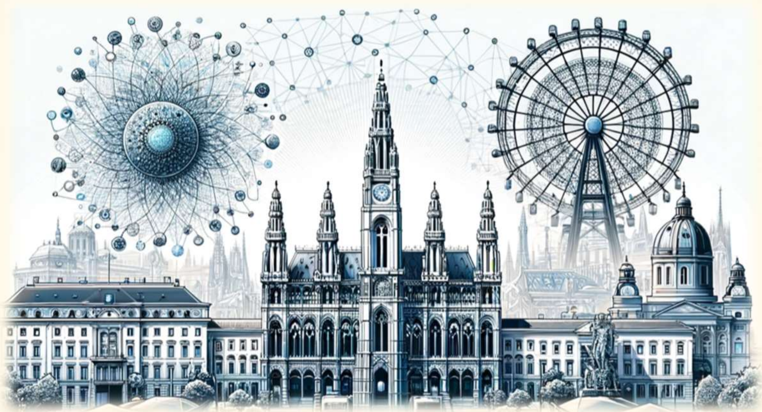
Rising house prices had significant impact on the model's predictions.

Did they influence fertility decisions (directly or as a part of broader economic conditions)?



Summary

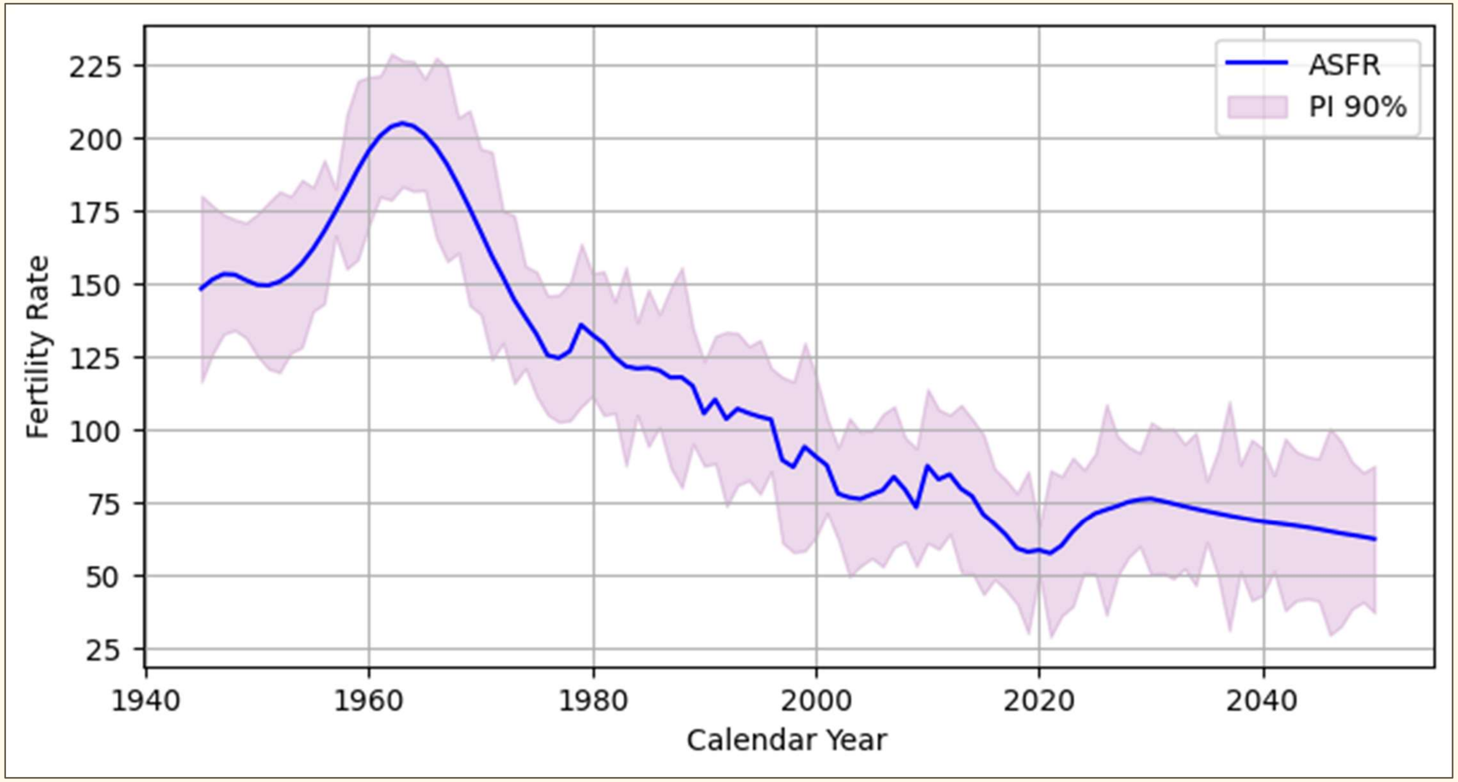
- Selection of ML techniques used to investigate the influence of socio-economic factors on age-specific fertility rates in England & Wales:
    - **SHAP** - quantified the contribution of each feature to the fertility model's predictions
    - **Saliency Maps** - visualized the importance of features like house price index in the model's decision-making process
    - **Attention Mechanism**: shown on which features the model focused when making predictions; indicates their strong interplay
- The model indicates that fertility decisions are multifaceted and influenced by a complex interplay of socio-economic factors.
- The analysis suggests that changes in child benefits relative to household income, significantly impact fertility trends. It also demonstrates a direct correlation between fertility rates and house prices.
- The use of ML interpretability & explainability methods can provide a deeper understanding of the model's predictions, which may support research and policy formulation.



Thank you!

Prediction interval for a NN model

Introducing a bias to the loss function to estimate the lower and upper bounds of the PI – intuitive, simple and efficient approach.



Prediction interval for neural network models using weighted asymmetric loss functions. M. Grillo, Y. Han, AW (2023)