

## Study Limitations

The main study limitations are due to the retrospective design, the reliance on self-reported data, and the potential for bias in the selection of households for the surveys. Both household surveys (surveys ran between November 2020 and January 2021) ask respondents to provide estimates of costs incurred pre-Covid, meaning the data is subject to recall bias. This was a pragmatic approach to mitigate against the potential influence of Covid on spending patterns during the study period. The FHS household survey was also disseminated via the partner charities using membership email lists, social media channels and websites in addition to promotion of the survey on both FSA and RSM social media channels. This could have resulted in sampling bias as participants will likely have been members of the partner charities, thus potentially excluding those that are not aware or affiliated with the partner charities. However, this was done for practical purposes as utilising the charities' outreach channels provided better access to FHS respondents, which helped maximise sample size.

Additionally, prospective power analysis to determine the minimum sample size needed to detect an effect of a given size, was not performed. However, it must be noted that despite power analysis not being performed, this is the largest survey of its kind in the UK, and it may be difficult to obtain larger sample sizes in similar timeframes.

The FHS household survey was largely completed by females (79%) whereas the non-FHS household survey was more representative of the population (52% female). This means there is a risk of respondent bias when analysing differences within FHS types and between the non-FHS group. The study addressed this issue by adjusting for the confounding effect of gender in the multivariate regression models. Further sensitivity analysis was conducted to assess whether the relationship between FHS types and cost differentials varied between females and males. The results from this analysis found no difference between females and males.

A further limitation is that kitchen equipment costs were monetised by deriving a UK high street average price for all the pieces of equipment named by respondents. As respondents were not asked to detail when they purchased the additional kitchen equipment, the assumption is that these were bought fairly recently and that the market price in 2021 has not significantly changed in the last few years. The alternative was to ask respondents to recall how much they had spent on each item, but this would have been subject to unacceptable recall bias. Therefore, the approach selected provided the most pragmatic estimation for kitchen equipment costs.

For the propensity score matching model, we were unable to achieve 'perfect' matching (for example, not all sub-characteristics had a post-matched standardised mean difference less than 0.1) which means that the FHS survey respondents and non-FHS group respondents are not completely balanced on demographic and household characteristics. This was due to there being a trade-off between achieving 'perfect' matching and losing observations which would have affected sample size. Although there were other matching parameters – described in Appendix 6 – which would have resulted in 'perfect' matching, the sample size would have been reduced by more than 50%. We addressed this remaining imbalance through adjustment for the demographic and household characteristics in the multivariate regression analysis as part of a 'doubly robust' approach to addressing confounding. However, we note that our approach to addressing covariate imbalance may not have removed residual confounding by unmeasured factors.