The University Of Sheffield.

School of Health And Related Research

Health Economics & Decision Science (HEDS) Discussion Paper Series

Family Lifestyle Dynamics and Childhood Obesity: Evidence from the Millennium Cohort Study

Authors: Laura A Gray, Mónica Hernández Alava, Michael P Kelly, Michael J Campbell

Corresponding author: Laura A Gray

ScHARR, University of Sheffield,
Regent Court, 30 Regent Street, Sheffield, S1 4DA
Tel: (+44) (0)114 222 0885
Fax: (+44) (0)114 222 0710
Email: laura.gray@sheffield.ac.uk

No. 17.01

Family Lifestyle Dynamics and Childhood Obesity: Evidence from the Millennium Cohort Study

Laura A Gray[1]*, Mónica Hernández Alava[1], Michael P Kelly[2], Michael J Campbell[3]

1. Health Economics and Decision Science, School of Health and Related Research, University of Sheffield, Sheffield, S1 4DA
2. Institute of Public Health, University of Cambridge, Cambridge, CB2 0SR
3. Design, Trials and Statistics, School of Health and Related Research, University of Sheffield, Sheffield, S1 4DA

* corresponding author. Email: laura.gray@sheffield.ac.uk. Tel: 0114 222 0885. Fax: 0114 222 0749.

Acknowledgements

This work was supported by the Economic and Social Research Council and the National Institute of Health and Care Excellence [grant number ES/H037977/1].

We are grateful to the Centre of Longitudinal Studies, UCL Institute of Education for the use of these data and to the UK Data Archive and UK Date Service for making them available. However, they bear no responsibility for the analysis or interpretation of these data.

We are grateful to Sandy Tebeuf and participants in the 2016 Health Economics Study Group conference at the University of Manchester and to Sabine Grimm for helpful comments on earlier versions of this work.

Abstract

Using data from the Millennium Cohort Study, we investigate the dynamic relationship between underlying family lifestyle and childhood obesity during early childhood. We use a dynamic latent factor model, an approach that allows us to identify family lifestyle, its evolution over time and its influence on childhood obesity and other observable outcomes. We find that family lifestyle is persistent and has a significant influence on childhood weight status as well as other outcomes for all family members. Interventions should therefore be prolonged and persuasive and target the underlying lifestyle of a family as early as possible during childhood in order to have the greatest cumulative influence. Furthermore, the results indicate that to reduce inequalities in childhood obesity, policy makers should target disadvantaged families and design interventions specifically for these families.

1 Introduction

Childhood obesity has been a growing problem in recent years but the causes are still not fully understood. Recent public health interventions and guidance which aim to reduce childhood obesity are increasingly focusing on the family, for example Change4Life (Department of Health, 2009) and clinical and public health guidance from the National Institute for Health and Care Excellence (NICE, 2008, 2010, 2006), rather than only on the child. In doing so they acknowledged an association between family lifestyle and childhood obesity. However, there is a lack of empirical evidence on this relationship.

Previous studies have shown strong associations in BMI between family members (Abrevaya and Tang, 2011; Burke et al., 2001; Brown et al., 2013; Brown and Roberts, 2013) suggesting that childhood BMI could depend on the same influences as the BMI of other family members. In addition, childhood adiposity is associated with socioeconomic status (Shrewsbury & Wardle, 2008) and parental education (Lamerz et al., 2005; von Kries et al., 1999), hours spent watching television (Hancox and Poulton, 2006), diet and exercise (James et al., 2004; Taylor et al., 2015) and breastfeeding (Gibson et al., 2016). Consequently, fully understanding the complex relationship between childhood obesity and other outcomes of family lifestyle could improve future interventions by helping to determine the expected benefits of family lifestyle interventions, when to implement them and which children and families might benefit the most from them.

Given that childhood obesity and other outcomes of family lifestyle are expected to be dependent on the same underlying influences, it is important to model these outcomes jointly. Despite this, the majority of previous studies have estimated these variables independently (Mizutani et al., 2007; Francis et al., 2003; Dewey, 2003). This approach is less informative when considering policy implications because it is only possible to identify how potential lifestyle interventions might influence a single outcome. Other studies have jointly estimated a range of observable lifestyle outcomes, including diet, alcohol consumption and smoking habits (Balia and Jones, 2008; Contoyannis and Jones, 2004), allowing the benefits of potential interventions to a range of outcomes to be investigated. To our knowledge, no previous study has attempted to simultaneously model multiple lifestyle outcomes for multiple family members and how this underlying lifestyle impacts on childhood obesity. If lifestyle is a family-level concept which influences a range of outcomes for all family members, it is unlikely that interventions targeting individual children will be effective and family-level interventions which aim to improve the underlying lifestyle of a family will be more successful as well as having additional benefits. This could provide important evidence in support of public health interventions such as Change4Life (DH, 2009) as well as informing future interventions.

Existing studies show that early-life influences of obesity, particularly lifestyle during pregnancy and early infancy are important in predicting later obesity (Gibson et al., 2016; Reilly et al., 2005; Ajslev et al., 2011; Toschke et al., 2002). However, these studies are generally cross-sectional and do not allow the evolution of lifestyle behaviours over time to be investigated. These early-life influences might continue to have an effect

throughout childhood and new influences could emerge as children grow up and their immediate environment changes, for example starting school. Cunha & Heckman (2008) encouraged the use of more flexible dynamics when modelling development during childhood because children change so rapidly. If family lifestyle persists over time, then it is likely that it will be difficult to change and strong interventions will be required. However, any successful interventions could have long-lasting impacts on a range of outcomes for all family members and bring additional benefits throughout life.

This paper contributes to the exiting literature by using a dynamic latent factor model to investigate how family lifestyle evolves over time during early childhood and how family lifestyle dynamics influence childhood obesity. Underlying family lifestyle will be identified in each period of the model using a latent factor: a range of lifestyle outcomes identified using exploratory factor analysis (EFA) will be included in the construction of the factor at each time point. In addition, a structural model will be employed to investigate how this latent factor evolves over time. This approach has a number of advantages. First, structural models can explain much more than models which use a single equation and can be used to investigate multiple and more ambitious research questions than more modest models such as fixed effects or instrumental variable models (Heckman and Urzúa, 2010). Second, unlike more commonly used autoregressive models, structural models allow parameter estimates to differ over time. Third, different mean outcomes can be identified for children with different characteristics unlike existing studies into adiposity which are restricted to estimating a single average treatment effect for a sample (Conti et al., 2011). Identifying the full distribution of treatment effects allows those who will benefit most from potential interventions to be identified; this is vital evidence for policy makers in order for them to have the greatest possible impact.

Public health decisions, for example those made by the NICE, use a range of evidence. This includes cost-effectiveness analysis which requires the prediction of future outcomes as consequences of potential interventions or policies. These models are evidence-based wherever possible but inevitably rely on a variety of assumptions when evidence is lacking. Evidence is routinely taken from trials, policy evaluation or econometric studies but much of this evidence uses static models with short-term single outcomes. Dynamic models with multiple outcomes, like those employed in this paper, provide more long-term evidence without the need for further assumptions or extrapolation, leading to more robust public health interventions being recommended.

In summary, we identify a comprehensive measure of family lifestyle which incorporates a range of family lifestyle outcomes, including childhood adiposity. We investigate the influence family lifestyle has on childhood weight status, and how family lifestyle evolves over time during early childhood. This is information which will be fundamental in determining the lifestyle interventions at a family level which are likely to be most effective.

2 Econometric Methods

In order to investigate the dynamic relationship between underlying family lifestyle and childhood obesity, we use a dynamic latent factor model, similar to that used by Cunha & Heckman (2008). They use this approach to identify the formation of skills during early childhood, whilst we use it to explore the evolution of family lifestyle and its relationship with obesity. The model is made up a set of latent factors (sometime known as measurement models) which identify the underlying lifestyle of a family using a range of outcomes and a structural model which estimates the relationship between these latent factors, in this case, the dynamic process of how family lifestyle evolves over time. Both parts of the model are outlined below and are jointly estimated using maximum likelihood. More details about structural models can be found in the literature (Bollen, 1989; Byrne, 2012).

2.1 Latent Factor for Family Lifestyle

Previous studies, such as Balia and Jones (2008) have jointly estimated multiple lifestyle outcomes using a multivariate probit model. Their model took into account the correlation of the error terms in each of the outcome equations and accounts for possible endogeneity of unobservables in the correlation of error terms. However, using this model it is not possible to directly estimate the underlying factor which is influencing each of these outcomes and therefore it is also not possible to estimate the effect that this underlying factor has on each of the outcomes. Due to the unobservable nature of this underlying factor, the only way to directly estimate it is using a common factor model, allowing this underlying concept to be identified without measurement error (Skrondal and Rabe-Hesketh, 2004).

We are interested in the relationship between childhood adiposity and underlying family lifestyle, so that

$$\boldsymbol{Y}_{it} = \boldsymbol{\lambda}_t \boldsymbol{\theta}_{it} + \boldsymbol{\delta}_t \boldsymbol{W}_{it} + \boldsymbol{\xi}_{it} \qquad (1)$$

where $\boldsymbol{Y}_{it}$ is childhood adiposity outcome at time t of child i, $\boldsymbol{\theta}_{it}$ is underlying family lifestyle with corresponding factor loading $\boldsymbol{\lambda}_t$ at time t, $\boldsymbol{W}_{it}$ is a vector of independent variables influencing the adiposity outcome at time t with vector of coefficients $\boldsymbol{\delta}_t$ and $\boldsymbol{\xi}_{it}$ is normally distributed error term.

Lifestyle is unobservable and cannot be identified using this single equation. The only way in which to identify underlying family lifestyle and estimate its influence on childhood adiposity is to use a latent factor. This uses additional observable outcomes which depend on family lifestyle as proxies for it. Because these observable outcomes each depend on underlying family lifestyle there is likely to be multicollinearity between them but by using a latent factor, this multicollinearity can be accounted for.

Similar to Equation 1, each additional outcome used to proxy family lifestyle is related to the underlying latent factor so that, for continuous outcome k

$$\boldsymbol{Y}_{kit} = \boldsymbol{\lambda}_{kt}\boldsymbol{\theta}_{it} + \boldsymbol{\xi}_{kit}; \tag{2}$$

the error terms are equivalent to that in Equation 1 and are independently and identically distributed. Other parameters are also equivalent to those in Equation 1, meaning that the effects of underlying family lifestyle on each of these outcomes can be investigated using this model. In both Equations 1 and 2, continuous outcomes are estimated using a linear regression and discrete outcomes are estimated using probit or ordered probit models, respectively. Threshold parameters for these discrete variables are jointly estimated and are strictly increasing.

The outcomes of family lifestyle can differ between periods. We assume here for simplicity that there is a single latent factor but this will be tested using the exploratory factor analysis (EFA). Outcomes in each period will be chosen using EFA and informed by existing literature. It is important that Equations 1 and 2 are jointly estimated in order to create the latent factor. If childhood adiposity were not included in the estimation of the same latent factor then underlying family lifestyle would be endogenous in Equation 1. It is assumed that there is no remaining correlation between outcomes at time t once the underlying factor for family lifestyle has been accounted for.

2.2 Structural Model

A structural model estimates the relationships between the latent factors; in this case, it creates the dynamic structure of underlying family lifestyle over time. This structure allows more long-term outcomes to be investigated and can show the extent to which influences can accumulate over time.

The initial underlying family lifestyle $\boldsymbol{\theta}_{i0}$, at time $t = 0$ around the time child i is born is

$$\boldsymbol{\theta}_{i0} = \boldsymbol{X}'_{i0}\boldsymbol{\beta}_0 + \boldsymbol{u}_{i0} \tag{3}$$

and depends on family characteristics $\boldsymbol{X}_{i0}$, with vector of corresponding coefficients $\boldsymbol{\beta}_0$. Error vector $\boldsymbol{u}_{i0}$ is made up of two parts; the family random effect $\boldsymbol{\eta}_i \sim N(0, \sigma_\eta)$ and independent error term $\boldsymbol{\varepsilon}_{i0} \sim N(0, \sigma_{\varepsilon_0})$ which is normally, independently and identically distributed.

Family lifestyle in the subsequent periods of the model depends on the underlying family lifestyle of the previous period, so that

$$\boldsymbol{\theta}_{it} = \alpha\boldsymbol{\theta}_{it-1} + \boldsymbol{X}'_{it}\boldsymbol{\beta}_t + \boldsymbol{u}_{it} \tag{4}$$

allowing underlying family lifestyle to evolve over time following a first order autoregressive process. Independent variables $\boldsymbol{X}_{it}$, as well as parameters α and $\boldsymbol{\beta}_t$, can differ over time. Again, the error terms $\boldsymbol{u}_{it}$ can be decomposed into a time-varying error term, $\boldsymbol{\varepsilon}_{it} \sim N(0, \sigma_{\varepsilon_t})$ and the time-invariant unobserved family random effect, $\boldsymbol{\eta}_i \sim N(0, \sigma_\eta)$. The inclusion of the family random effect allows us to account for any

unobservable influence on underlying family lifestyle over time. This allows us to ensure that the majority of variation in the observable lifestyle outcomes are accounted for within the model.

2.3 Model Identification

One cannot identify both the means and the intercepts in Equations 3 and 4 because both the latent factors $\boldsymbol{\theta}_t$ and the error terms are unobservable. As outlined in Cunha and Heckman (2008), in order to identify the model, we fix the variance of some of the error terms. The variance of the error term, $\boldsymbol{u}_0$ in Equation 3 (σ_{u_0}) is fixed at 0.05 and the variance of error terms, $\boldsymbol{u}_t$ in Equation 4 (σ_{u_t}) is fixed at 0.01. This identifies the structural part of this model and is equivalent to restricting the variance to one (normalisation) as is done in a probit model. In this case, model convergence was more easily achieved using values smaller than one but the magnitude of these values is arbitrary. A more detailed description and proof for the identification of this model is given by Cunha and Heckman (2008).

The dynamic latent factor model (Equations 1 to 3) is estimated by simulated maximum likelihood using Monte Carlo integration with 3,000 integration points. Robust standard errors are computed using a sandwich estimator. This requires the computation of a four-dimensional integration. The model is estimated using Mplus 6.1 (Muthen & Muthen, 2011) and data manipulation is carried out in Stata 13.

2.4 Data

We use data from the Millennium Cohort Study (MCS) which contains a rich set of information for a sample of 19,517 children born in and around the year 2000. Cohort members were recruited using child benefit records, at the time a universal benefit. The cohort members' carers were interviewed when the child was nine months old and subsequently when they were three, five and seven years old (see Hansen, 2012 for a guide to these datasets). During each of these subsequent interviews, data on height and weight of the children were collected, amongst other adiposity measures, allowing BMI and weight status to be calculated.

Childhood adiposity is an outcome of underlying family lifestyle in each period of the model. In the first period, we use child weight in kilograms because weight categories are not available at nine months of age. In subsequent periods, child weight status is included using the age and sex specific International Obesity Task Force (IOTF) definitions designed by Cole et al. (2000), which classify children as normal weight, overweight or obese. The median and interquartile ranges of BMI by age and sex are displayed in Figure 1 along with the IOTF cut-offs. The outcomes for the latent factors were chosen in accordance with the existing literature and using exploratory factor analysis (EFA) (Muthen and Muthen, 2010) for each period and are displayed in Table 1. These variables include maternal and paternal weight status (normal, overweight or obese), maternal smoking status (smoker, non-smoker), whether a pregnancy was planned (planned, not planned), exclusive breastfeeding duration (never breastfed, between four and thirteen weeks, between fourteen and seventeen

weeks and over seventeen weeks), screen time (3 hours or more each day), regular meal times, participation in sport (never, once, twice, three times, four or more times), visits to the park (at least once a week), unhealthy snacking between meals and daily breakfasting.

Socioeconomic and family background variables directly influence underlying lifestyle; these include variables which are found in the literature to influence the observable lifestyle outcomes outlined above. These include socioeconomic status (SES) using the five point National Statistics Socioeconomic Classification (NS-SEC) scale. The highest SES level of each of the cohort members' parents is used to measure the cohort members' family SES at birth. Maternal education at birth is also included. Both family SES and maternal education influence lifestyle only in the initial period of the model. This is because they do not differ a great deal over time and so any influence they have on later periods' lifestyle is assumed to be captured through the autoregressive process. Family structure, i.e. whether the family is a two-parent or single-parent family, is included in every period of the model because it has more variation throughout early childhood. Cunha & Heckman (2009) found that single-parent households were less able to invest in their children and suggested that this could be due to differences in time constraints between single-parent and two-parent households, or the lack of resources available to single-parent households. These variables all have an influence on childhood weight status through their influence on underlying family lifestyle. Ethnicity, age and sex are included as independent variables directly influencing child weight. We allow ethnicity to influence weight status in each subsequent period but because weight status is age and sex specific, age and sex are only included in the initial period.

Any observations which are not present in all four periods are removed from the analysis leaving a balanced sample of 11,484. In line with previous literature, children are also removed from the sample for a number of other reasons. These include children from multiple births, those weighing less than 2.5kg at birth, those taken to a special care unit straight after birth and those who's main carer is not their natural mother. Observations are also removed from the sample when independent variables are missing. This leaves a balanced panel sample of 8,462 observations. Missing data are assumed to be missing at random.

One benefit of latent factor models is that item-non-response in the outcomes does not necessarily result in observations being removed from the analysis. A latent factor can still be estimated using the remaining outcomes, provided that there are at least two non-missing outcomes for each observation. In accordance with the World Health Organisation recommendations for biologically implausible values, childhood and parental weight statuses are recorded as missing if the height, weight or BMI values used to calculate them were implausible. Although this means that childhood adiposity outcomes were recorded as missing for some observations, this does not result in the removal of any observations.

3 Results

Two different specifications of the dynamic latent factor model outlined above were implemented. Initially, a model was estimated with constant parameters across all periods. In this model, all lifestyle outcomes which appear in more than one period of the model had constant parameters, including factor loadings and threshold parameters. Independent variables influencing underlying family lifestyle or childhood adiposity and which appear in more than one period also had fixed parameters. In the second less restricted model, factor loadings, threshold parameters and independent variable coefficients were allowed to vary over time. All parameters were freed over time apart from the autoregressive component (α_t) and the factor loadings $\boldsymbol{\lambda}_{kt}$ for maternal and paternal weight categories along with their corresponding threshold parameters. These parameter estimates are restricted over time in order to achieve convergence in the model which was not possible when they were freed.

In both the restricted and unrestricted models, the family random effect $\boldsymbol{\eta}_i$ was found to be insignificant. This suggests that the majority of variation in the observable lifestyle outcomes is accounted for by the underlying latent factor. For this reason, and to enable the final model to converge more readily, this random effect was removed from the final models. This did not significantly affect our results.

Model fit of the unrestricted model showed an improvement on the restricted model using a likelihood ratio (LR) test as well as Akaike and Bayesian Information Criteria (AIC and BIC) supporting the claim by Cunha & Heckman (2008) that time-invariant parameters are not always best practice when analysing data on young children because they are constantly developing and changing. The remainder of this paper therefore focuses on results from the unrestricted model.

3.1 Parameter Estimates

Table 1 shows the factor loadings for each lifestyle outcome in each period of the model whilst Tables 2 and 3 present the parameter estimates relating to determinants of childhood adiposity and family lifestyle, respectively. As indicated in Table 1, all factor loadings are significant and have the expected sign; an improvement in underlying family lifestyle is associated with improved lifestyle outcomes, including but not limited to childhood adiposity. Childhood adiposity has a consistently positive and significant response to changes in the latent family lifestyle. Maternal weight status provides the largest informational content for the underlying lifestyle factor, particularly in comparison with paternal weight status, suggesting that the mother is largely responsible for the lifestyle of a family. The proportion of variance in childhood weight status explained by underlying family lifestyle is 11.3%, suggesting that improvements to family lifestyle could significantly reduce the likelihood of obesity in a child. Comparatively, the proportion of variance in maternal weight status explained by family lifestyle is 93.5%. This suggests that maternal weight status will be highly

influenced by family lifestyle and that maternal obesity could prove useful in identifying families that need more help improving their lifestyle.

Table 2 shows the parameter estimates for variables influencing childhood adiposity in each period. It shows that boys weigh more at nine months than girls do, ceteris paribus. At nine months of age, Asian children weigh significantly less than their white counterparts do. These associations are as expected. Asian children are significantly less likely to be obese or overweight at the age of three years, but this association is insignificant by the age of five. Conversely, black children are, on average, significantly more likely than white children to be obese or overweight at the age of five and seven years.

Factor scores provide numerical values for underlying family lifestyle. Although these factor scores have no cardinal meaning and cannot be directly compared across time periods, they can be used to rank families in terms of their lifestyle to determine where each family lies on a lifestyle distribution. Families with higher factor scores have 'healthier' lifestyle than families with lower factor scores. Our results show that the variation in previous family lifestyle accounts for 98.7% of variation in current family lifestyle when the child is four years old. Table 3 shows the parameter estimates for the variables influencing these factor scores. Previous family lifestyle has a positive and statistically significant influence on current family lifestyle. Determinants of family lifestyle in this model are consistent with the literature. Family SES, maternal education and being from a single-parent family each have a statistically significant effect on initial latent family lifestyle. Families with high SES are at the higher end of the lifestyle distribution in the initial period and those with a low SES are towards the lower end of the distribution, ceteris paribus. Single-parent families are on average higher up the lifestyle distribution across all periods. However, this effect is only significant in the initial period and when the child is seven years old.

3.2 Factor Scores

From the model, we can determine the factor scores for the underlying family lifestyle factors for each individual. Factor scores are the numerical values of the underlying factors and are estimated using the observable characteristics of each observation (Skrondal and Laake, 2001). In this case, they have no cardinal meaning but factor rankings and percentiles can be used to compare observations with respect to family lifestyle. From this, the persistence and mobility of underlying family lifestyle can be investigated. In addition, we can use the variance-covariance matrix from the model to calculate the proportion of variance in variables of interest explained by the latent factor.

Correlations between the factors scores in each period are consistently above 0.982, demonstrating an immobility in the family lifestyle distribution. Table 4 shows what proportion of families remain in the same part of the lifestyle distribution over time. For example, 87.43% of families which were above the ninety-fifth percentile on the lifestyle distribution in the initial period remain above the ninety-fifth percentile when a child is seven years of age, showing some movement at the upper end of the distribution. Families that are

initially in the bottom five percentiles almost never improve their lifestyle. The increased movement at the upper end of the family lifestyle distribution suggests that general interventions at a population level are likely to help only those at the upper end of the distribution and are unlikely to benefit those that are in need of the most help to improve their lifestyles. This suggests that in order to improve the lifestyles of families at the lower end of this distribution, interventions should be specifically designed to help these families and targeted directly at them. Interventions designed to help everyone, might inadvertently, and disproportionately, help those at the top, rather than those at the lower end of the distribution.

Table 5 shows the difference in characteristics between families in the top and bottom five percentiles of the lifestyle distribution. Children in families above the 95th percentile have a lower BMI and are less likely to be obese during childhood than those from families below the 5th percentile. This differences increases as children get older as those in the lowest 5 percentiles become more likely to be obese. The most overwhelming difference between those at the upper and lower ends of this distribution is that between SES; families with low SES are almost always at the lower end of the lifestyle distribution. The information displayed in Table 5 can help to identify families at the lower end of the lifestyle distribution in order to help policy makers design and target interventions more effectively and reduce inequalities in childhood obesity.

4 Simulations

In order to investigate the influences of underlying family lifestyle on childhood obesity, the posterior distributions of observable childhood weight status can be calculated, that is the predicted outcome of childhood weight status, conditional on other observable lifestyle outcomes and independent variables. This equation requires the computation of several integrals and for this reason we approximate these predictions with simulations using the estimated parameters from the dynamic latent factor model. This prevents the need for the complex calculations and allows us to estimate the likelihood of obesity in children with given sets of observable characteristics and at different ages using a single model. We use 10,000 simulated repetitions in order to stabilise the expected means. All simulations are estimated using Stata 13.

Using simulations along with the parameter estimates discussed above, it is possible to investigate a range of policy relevant relationships within this model. Here, we outline just a few which we feel are of particular policy interest. In order to investigate the inequalities in obesity prevalence between advantaged and disadvantaged children, we predict the likelihood of obesity, and the expected percentile of the lifestyle distribution, for two hypothetical children using a multidimensional measure of disadvantage. The first is an ‘advantaged’ child who is from a family with high SES, has a highly educated mother and is from a two-parent family. The second ‘disadvantaged’ child is from a family with low SES, has a poorly educated mother and is from a single-parent family. Both children are white girls and are 42.21 weeks old, the mean age of the cohort at the time of the initial MCS interviews.

Table 6 shows that the advantaged child has a lower risk of obesity than the disadvantaged child, an observation which is consistent over time. The difference is noticeable as early as the age of three years, when children from disadvantaged backgrounds are around 50% more likely to be obese than those from the most advantaged backgrounds. This difference increases with age and by the age of five years, the disadvantaged child is more than twice as likely to be obese than the advantaged child.

Table 7 shows the expected percentile of underlying family lifestyle for the advantaged and the disadvantaged child. There is a substantial difference in the relative underlying family lifestyle between these hypothetical children from these different backgrounds. The simulated kernel density distributions of expected lifestyle for each of these hypothetical children at the age of seven years are displayed in Figure 2 and show very little overlap in the distributions of family lifestyle between the two children. These shows how the family background characteristics, SES, maternal education and family structure, account for significant differences in underlying family lifestyle and in doing so create inequalities in childhood obesity. This difference in family lifestyle mediates the relationship between family background and childhood adiposity. In addition to the parameter estimates fromthe dynamic latent factor model outlined above, these simulations emphasise the importance of targeting children from disadvantaged backgrounds when aiming to reduce inequalities in obesity prevalence through the use of lifestyle interventions.

5 Discussion

This study investigates the persistence of family lifestyle and the influence that this has on childhood adiposity. The extent to which this underlying family lifestyle mediates the impacts of socioeconomic and family background on childhood adiposity is also explored. Our results show that family lifestyle is persistent; the largest influence on family lifestyle was previous family lifestyle. Families rarely move up or down the lifestyle distribution over time, particularly those at the lower end of this distribution. We also find an increasing positive association between disadvantage and childhood obesity throughout early childhood.

This study adds to the existing literature in a number of ways. First, the latent factors used in each period allow a range of outcomes to be used to estimate an underlying family lifestyle. These latent factors provide a more comprehensive measure of lifestyle compared to single-item lifestyle proxies, such as those used by many studies within the existing literature, see Reilly et al. (2005), Bauer et al. (2011), Haug et al. (2009) and Janssen et al. (2005). Second, the use of latent factors also builds on work by Balia & Jones (2008) who use a multivariate probit model to simultaneously estimate a range of lifestyle behaviours but who do not directly estimate the underlying influence effecting these outcomes. Third, this study uses a dynamic model of lifestyle. Previous studies such as Janssen et al. (2005), Haug et al. (2009) and Giles-Corti et al. (2003), among others, investigated lifestyle variables using static or cross-sectional models. The dynamic nature of the latent factor model allows the evolution and persistence of family lifestyle to be explored during early

childhood making it possible to investigate the effects of early-life and family background influences on childhood adiposity over time. The dynamic nature of the model is also important for providing economic models with more long-term evidence that can be used to identify the most cost-effective interventions using fewer extrapolations. Finally, this study uses a large dataset which is representative of children and families in the UK. To our knowledge there is no other study which investigates the effects of underlying family lifestyle on a range of childhood outcomes using such a large number of children. By estimating the same outcomes over a period of time using longitudinal data, this study provides more long-term evidence than many other studies in the area and could lead to stronger public health guidance.

5.1 Policy Implications

Maternal weight status provides the greatest informational content to the latent factor for family lifestyle in each period suggesting that maternal influences are more important when investigating family lifestyle than paternal influences. This could be due to the role that mothers play in the lifestyles of young families. Mothers are most often responsible for family diet, exercise and other lifestyle behaviours and this could mean that underlying family lifestyle is most highly driven by maternal lifestyle outcomes. This would suggest that any family-based lifestyle policies could be easiest implemented through maternal education and providing mothers with additional help to improving the lifestyle of their family.

The largest influence on family lifestyle is previous family lifestyle. This persistence of underlying family lifestyle suggests that an exogenous shock to family lifestyle, caused by an intervention or otherwise, which successfully improves underlying lifestyle, could have long-lasting influences on childhood adiposity as well as the other observable outcomes for all family members. Family lifestyle interventions should be carried out as early in childhood as possible in order to have the greatest cumulative impact on the outcomes for the child. Targeting the families of very young children or expectant parents could have effects that last throughout childhood. This is consistent with other studies in the economics literature that find that other childhood outcomes are most improved when interventions focus on the very early years (Heckman, 2006; Cunha et al., 2010; Heckman, 2012; Conti and Heckman, 2013) in order to have the greatest cumulative influences. Although interventions carried out before birth might be the most effective in reducing childhood obesity, targeting families this early is not always possible. Despite earlier interventions being most effective, there is also evidence that successful lifestyle interventions later in childhood could still significantly reduce childhood adiposity. The persistence of family lifestyle means that any interventions which aim to improve family lifestyle will need to be substantial, in order to cause a significant improvement family lifestyle. The need for large and effective policies suggests that interventions will need to be prolonged in order to achieve the desired effects.

Socioeconomic variables play a large part in shaping underlying family lifestyle throughout early childhood. Our results showed a large difference in the prevalence of childhood obesity expected in advantaged and

disadvantaged children. This large difference in obesity prevalence was largely explained by the differences in underlying family lifestyle in advantaged and disadvantaged families. This, coupled with the fact that there is very little mobility of family lifestyle particularly at the lower end of the family lifestyle distribution, suggests that interventions should be specifically targeted at disadvantaged families. This is similar to findings by Heckman (2006) who encouraged early interventions targeted at disadvantaged children to improve skill formation. Furthermore, interventions designed to improve the lifestyle of the general population are unlikely to be effective in improving the lifestyle of those at the immobile lower end of the family lifestyle distribution. Therefore, interventions should be designed specifically with these disadvantaged families in mind. The differences in childhood obesity prevalence between advantaged and disadvantaged children increases as children get older. This reinforces the suggestions that interventions should target disadvantaged families at the earliest possible stage of childhood in order to reduce these growing inequalities and to have the greatest cumulative effect over time.

Results from this study support the use of interventions such as those that aim to tackle attitudes to lifestyle and educate families about how they can improve their lifestyles and what the benefits of doing so might be. Encouraging change in specific lifestyle behaviours cannot singlehandedly address the obesity epidemic, nor can tackling social determinants of underlying lifestyle. However, simultaneously targeting a range of different lifestyle outcomes through improved education and attitudes towards lifestyle could be one way of effectively reducing the prevalence of childhood obesity with further benefits to other observable outcomes for the family. Current UK policies such as Change4Life have identified the need to target families rather than individuals when aiming to improve childhood outcomes. In addition, NICE (2013) have also recognised that disadvantage and obesity are closely related. However, there is a lack of specific focus, in both design and implementation, on families from a disadvantaged background. NICE (2013) recommended the targeting of specific neighbourhoods alongside more widespread interventions, but they do not go as far as suggesting that interventions in these neighbourhoods should specifically designed for disadvantaged families. Future interventions that aim to reduce inequalities in family lifestyle, childhood obesity and a range of other outcomes for all family members should consider the design of interventions specifically for disadvantaged families.

The multiple outcomes estimated in this study, using a single dynamic model, mean that policy implications go beyond childhood obesity. The results emphasise the need for policy makers to consider the long-term influences and effects on multiple outcomes that their family lifestyle interventions could have. It is important to consider how these interventions might improve a number of observable outcomes for multiple family members over time as well as reducing inequalities in these outcomes. Not only is this important for policy makers but also for cost-effectiveness modellers wishing to provide robust evidence to decisions makers such as NICE on public health interventions. The joint estimation of the system of equations included in these latent factors can also provide economic or cost-effectiveness models with estimates of correlations between these equations which means that fewer assumptions are required.

5.2 Limitations

Although we find that the underlying factor for family lifestyle accounts for the vast majority of variation in the observable outcomes, it is possible that genetics could be playing a role here. In our sample, the mother is always the biological mother of the child but the father figure is not always a biologic father and sometimes no information on a father is collected at all. This could suggest that genetics could, to some extent, be responsible for some of the association between weight statuses of family members; child weight might be more correlated with maternal than paternal weight status because they are more often genetically related. This could increase the correlation between maternal and childhood adiposity relative to the correlation between paternal and childhood adiposity, meaning that maternal weight status provides the higher informational content. We can be confident however, that any part genetics does play in this underlying factor is minimal because many of the other outcomes used to create the latent factors are clearly related to lifestyle and not to genetics. There is a growing literature on obesity and epigenetics and future research could investigate how epigenetics might influence the relationship between childhood obesity and family lifestyle.

The MCS contains data from when a cohort child is born. However, data from before birth might have proven useful in identifying family lifestyle before the birth of a child. This would have allowed the effects of having a child on family lifestyle to be investigated. More detailed data on siblings might also have been useful and future research from later waves which contain such data could focus on the differences between individual and family effects. There could also be a cohort effect here. All children in the sample were born around the turn of the millennium; results might be slightly different for children born today. That said, given the rise in both childhood obesity and inequalities faced by disadvantaged families, any associations between the two could be even stronger.

Family lifestyle is already well established by the time a child reaches seven years old. However, as children become adolescents and increasingly interact with people outside the family home, they might be less influenced by the lifestyle of their family and could develop a more individual lifestyle as they become more independent. Further research could investigate how the dynamic path of lifestyle changes throughout childhood and into adolescence when they begin to have increasing individual influences. Likewise, further research into the intergenerational transmission of lifestyle could be useful for policy makers aiming to target families before the birth of a child.

5.3 Conclusion

In order to better tackle childhood obesity and childhood obesity inequalities, interventions should be focussed on the family. In particular, mothers should be provided with additional support and information on the benefits of a healthy lifestyle for her family. This should be done as early as possible in childhood in order to have the greatest cumulative influence. Interventions should also be designed specifically to help families

from disadvantaged backgrounds as well as being targeted specifically at these families. Successful interventions will need to be prolonged and substantial in order to overcome the persistent nature of family lifestyle. These types of intervention have the potential to benefit a wide range of outcomes, including childhood obesity.

References

Abrevaya, J. & Tang, H. (2011) Body mass index in families: Spousal correlation, endogeneity, and intergenerational transmission. Empirical Economics. 41 (3), 841–864.

Ajslev, T. a et al. (2011) Childhood overweight after establishment of the gut microbiota: the role of delivery mode, pre-pregnancy weight and early administration of antibiotics. International journal of obesity. 35(4) 522–529.

Balia, S. & Jones, A. M. (2008) Mortality, lifestyle and socio-economic status. Journal of Health Economics. 27 (1), 1–26.

Bauer, K. W. et al. (2011) Familial correlates of adolescent girls' physical activity, television use, dietary intake, weight, and body composition. The International Journal of Behavioral Nutrition and Physical Activity. 8 (1), 25–34.

Bollen, K. A. (1989) Structural Equations with Latent Factors. New York: Wiley.

Brown, H. et al. (2013) Going the Same 'Weigh': Spousal Correlations in Obestiy in the UK. Applied Economics. 46 (2), 153–166.

Brown, H. & Roberts, J. (2013) Born to be wide? Exploring correlations in mother and adolescent body mass index. Economics Letters. 120 (3), 413–415.

Burke, V. et al. (2001) Family lifestyle and parental body mass index as predictors of body mass index in Australian children: a longitudinal study. International journal of obesity related metab disord 25 (2), 147–157.

Byrne, B. M. (2012) 'The Full Latent Variable Model', in Structural Equation Modeling with Mplus: Basic Concepts, Applications and Programming. Hove, East Sussex: Taylor and Francis Group LLC. pp. 6–7.

Cole, T. J. et al. (2000) Establishing a Standard Definition for Child Overweight and Obesity Worldwide: International Survey. British Medical Journal. 320, 1240–1245.

Conti, G. et al. (2011) Early Endowment, Education and Health. Human capital and Economic Opportunity: A Global Working Grou, Working Paper Series 2011-001, 1-54.

Conti, G. & Heckman, J. J. (2013) The Developmental Approach to Child and Adult Health. Pediatrics. 131 (supplement), s133–s141.

Contoyannis, P. & Jones, A. M. (2004) Socio-economic Status, Health and Lifestyle. Journal of Health Economics. 23 (5), 965–995.

Cunha, F. et al. (2010) Estimating the Technology of Cognitive and Noncognitive Skill Formation. Econometrica : journal of the Econometric Society. 78 (3), 883–931.

Cunha, F. & Heckman, J. J. (2008) Formulating, Identifying and Estimating the Technology of Cognitive and Noncognitive Skill Formation. Journal of Human Resources. 43 (4), 738–782.

Cunha, F. & Heckman, J. J. (2009) The Economics and Psychology of Inequality and Human Development. Journal of the European Economic Association. 7 (2-3), 320–364.

Dewey, K. G. (2003) Is Breastfeeding Protective Against Child Obesity? Journal of Human Lactation. 19 (1), 9–18.

DH (2009) Change4Life Marketing Strategy. In Support of ealthy Weight, Healthy Lives.

Francis, L. A. et al. (2003) Parental weight status and girls' television viewing, snacking, and body mass indexes. Obesity Research. 11 (1), 143–151.

Gibson, L. A. et al. (2016) The effects of breastfeeding on childhood BMI: a propensity score matching approach. Journal of Public Health. 1–9. [online]. Available from: http://jpubhealth.oxfordjournals.org/lookup/doi/10.1093/pubmed/fdw093.

Giles-Corti, B. et al. (2003) Environmental and Lifestyle Factors Associated with Overweight and Obesity in Perth, Australia. American Journal of Health Promotion. 18 (1), 93–102.

Hancox, R. J. & Poulton, R. (2006) Watching television is associated with childhood obesity: but is it clinically important? International Journal of Obesity. 30 (1), 171–175.

Hansen, K. (2012) Millennium Cohort Study: First, Second, Thurd and Fourth Surveys: A Guide to the Datasets (Seventh Edition). (February).

Haug, E. et al. (2009) Overweight in school-aged children and its relationship with demographic and lifestyle factors: results from the WHO-Collaborative Health Behaviour in School-aged Children (HBSC) study. International Journal of Public Health. 54 (supplement 2), 167–179.

Heckman, J. J. (2006) Skill formation and the economics of investing in disadvantaged children. Science (New York, N.Y.). 312 (5782), 1900–1902. [online]. Available from: http://www.ncbi.nlm.nih.gov/pubmed/16809525 (Accessed 30 May 2013).

Heckman, J. J. (2012) The developmental origins of health. Health Economics. 21 (1), 24–29.

Heckman, J. J. & Urzúa, S. (2010) Comparing IV with structural models: What simple IV can and cannot identify. Journal of Econometrics. 156 (1), 27–37.

Janssen, I. et al. (2005) Comparison of overweight and obesity prevalence in school-aged youth from 34 countries and their relationships with physical activity and dietary patterns. Obesity Reviews. 6 (2), 123–132.

von Kries, R. et al. (1999) Breast Feeding and Obesity: Cross Sectional Study. British Medical Journal. 319, 147–150.

Lamerz, A. et al. (2005) Social class, parental education, and obesity prevalence in a study of six-year-old children in Germany. International Journal of Obesity. 29 (4), 373–380.

Mizutani, T. et al. (2007) Association of maternal lifestyles including smoking during pregnancy with childhood obesity. Obesity. 15 (12), 3133–3139.

Muthen, L. K. & Muthen, B. O. (2010) 'Exploratory Factor Analysis', in Mplus User’s Guide. Sixth Edit Los Angeles, CA: Muthen & Muthen. pp. 41–50.

Muthen, L. K. & Muthen, B. O. (2011) Mplus 6.1.

NICE (2006) NICE Clinical Guideline 43: Obesity guidance on the prevention, identification, assessment and management of overweight and obesity in adults and children. (December 2006).

NICE (2013) NICE Local Government Briefing 9: Preventing obesity and helping people to manage their weight. (May).

NICE (2008) NICE Public Health Guidance 11: Maternal and Child Nutrition. [online]. Available from: http://guidance.nice.org.uk/PH11 (Accessed 16 November 2011). [online].

NICE (2010) NICE Public Health Guidance 27: Dietary interventions and physical activity interventions for weight management before, during and after pregnancy. (July 2010).

Reilly, J. J. et al. (2005) Early life risk factors for obesity in childhood: cohort study. British Medical Journal. 330, 1357–1363.

Shrewsbury, V. & Wardle, J. (2008) Socioeconomic status and adiposity in childhood: a systematic review of cross-sectional studies 1990-2005. Obesity. 16 (2), 275–284.

Skrondal, A. & Laake, P. (2001) Regression among factor scores. Psychometrika. 66 (4), 563–575.

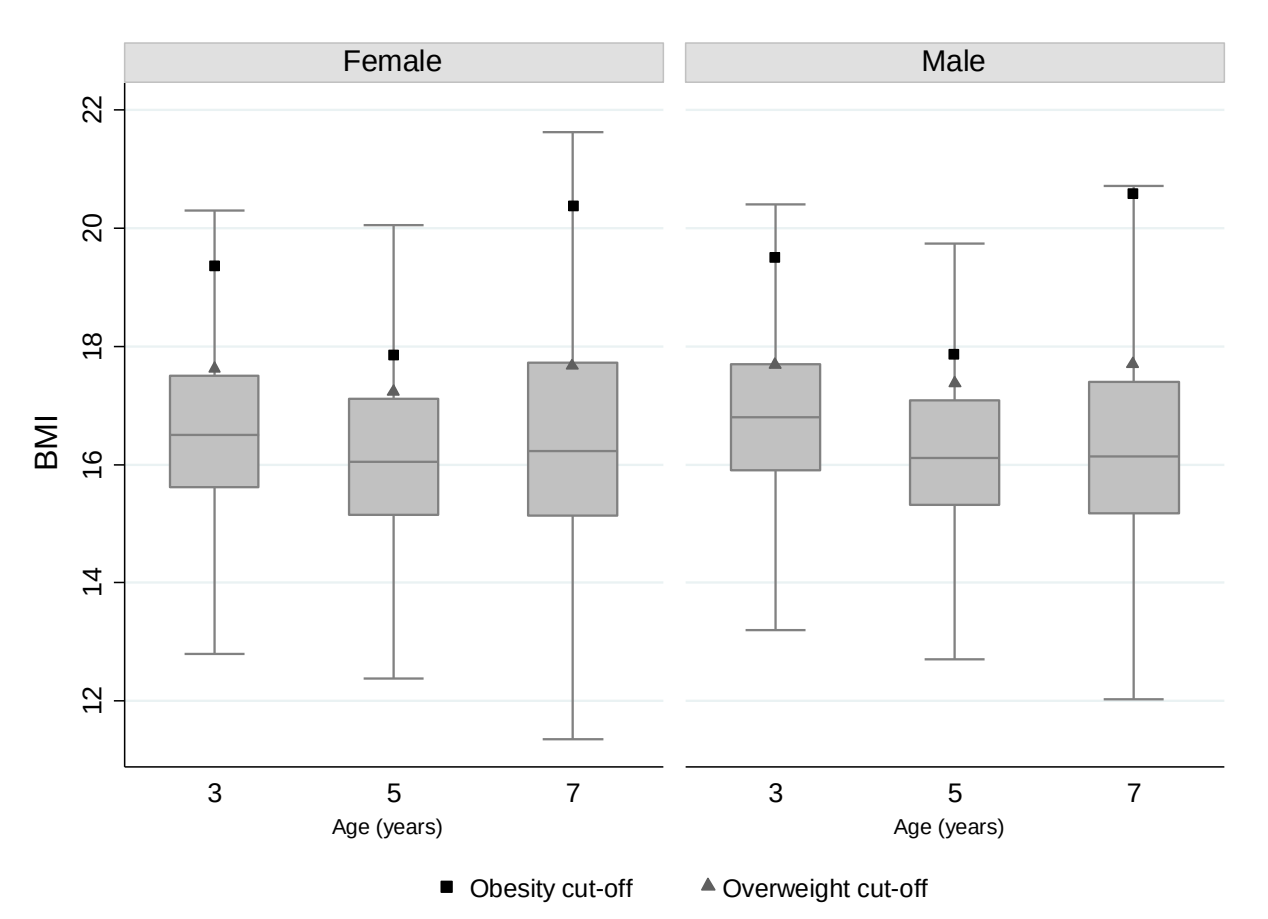
Skrondal, A. & Rabe-Hesketh, S. (2004) 'Factor Models and Item Response Models', in N Keiding et al. (eds.) Generalized Latent Variable Modeling: Multilevel, Longitudinal and Structural Equation Models. Florida: Chapman and Hall CRC. pp. 49–94.

Taylor, M. J. et al. (2015) Measuring perceived exercise capability and investigating its relationship with childhood obesity: a feasibility study. International journal of obesity (2005). 386S71. [online].

Toschke, A. M. et al. (2002) Childhood obesity is associated with maternal smoking in pregnancy. European Journal of Pediatrics. 161 (8), 445–448.

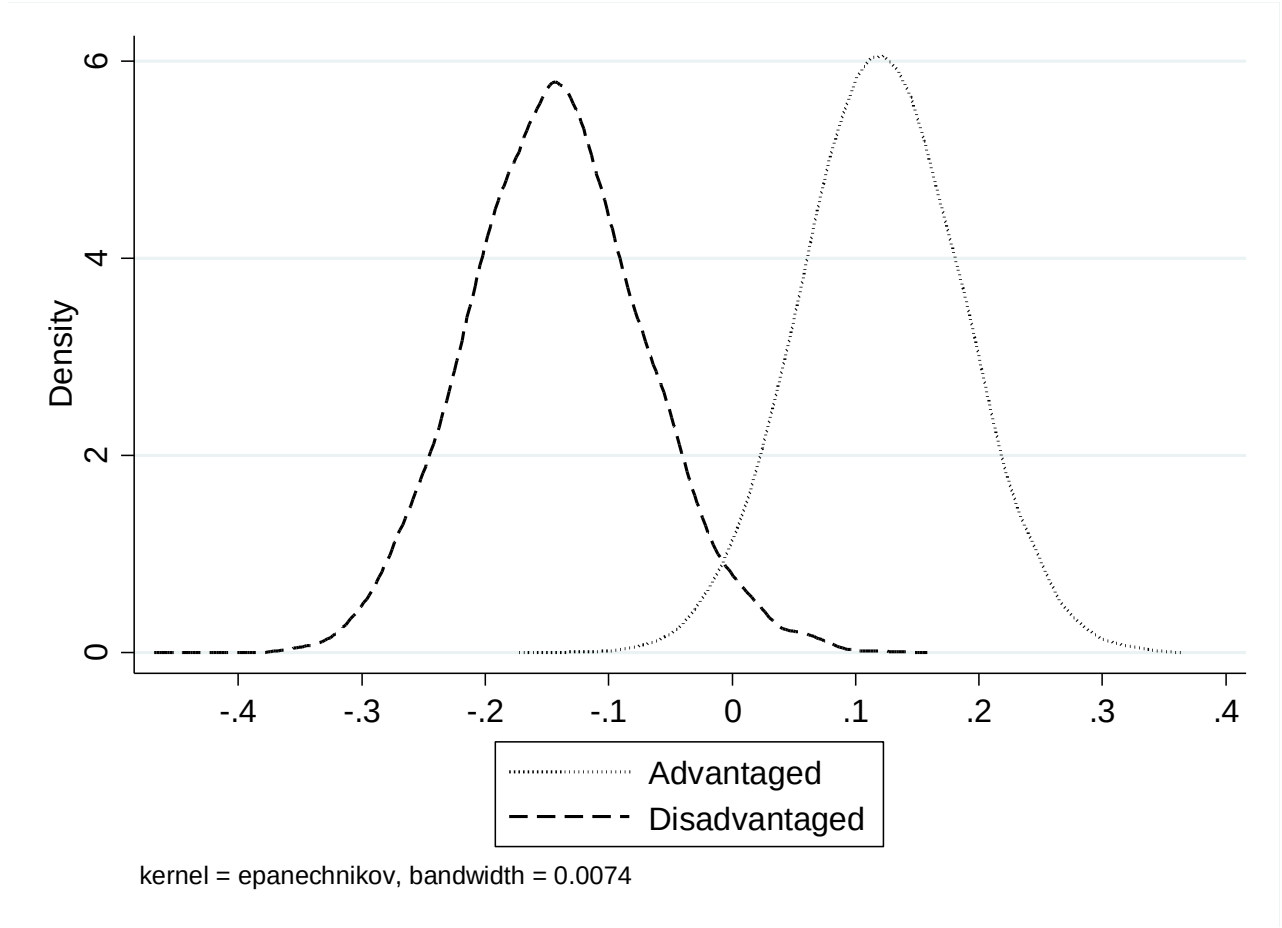
Tables and Figures

Figure 1: Median BMI and Interquartile Range by Age and Sex



Box plots showing median and interquartile range for BMI by age and sex using data from the Millennium Cohort Study. Outliers not included. International Obesity Taskforce (IOTF) age and sex specific cut-offs for obesity and overweight also shown.

Figure 2: Kernel Densities of Lifestyle Distributions in Advantaged and Disadvantaged Children aged 7.



Kernel density of posterior lifestyle distributions for advantaged and disadvantaged children aged seven years.

Table 1: Estimated Factor Loadings

	Factor Loading, **λ** (Equations 1 and 2) (Standard Error)			
Dependent Variable	Initial	Age Three Years	Age Five Years	Age Seven Years
Weight (kg)	-0.051*** (0.007)	-	-	-
Weight Category	-	-1.205*** (0.079)	-1.535*** (0.080)	-1.518*** (0.078)
Maternal Weight Category[a]	-8.527*** (0.321)	-12.574*** (0.418)	-12.574*** (0.418)	-12.574*** (0.418)
Fathers Weight Category	-1.393*** (0.102)	-1.215*** (0.088)	-1.215*** (0.088)	-1.215*** (0.088)
Mothers' Smoking Behaviour[b]	-0.739*** (0.105)	-0.757*** (0.101)	-0.697*** (0.092)	-0.643*** (0.083)
Planned Pregnancy	0.712*** (0.079)	-	-	-
Breastfeeding Behaviour	1.056*** (0.064)	-	-	-
Regular Meals	-	0.577*** (0.091)	0.648*** (0.090)	-
Over Three Hours TV per day	-	-0.867*** (0.076)	-0.545*** (0.070)	-0.431*** (0.062)
Sport	-	-	0.669*** (0.053)	0.561*** (0.047)
Playground/Park	-	-	0.154*** (0.057)	0.182*** (0.051)
Unhealthy Snacks	-	-	-	-0.290*** (0.056)
Regular Breakfast	-	-	-	0.553*** (0.082)
N	8,462			

This table shows factor loadings from the factor models. * $p < 0.01$, ** $p < 0.05$, *** $p < 0.001$, [a] for initial conditions this is pre-pregnancy weight category, [b] for initial conditions this is smoking behaviour during pregnancy.

Table 2: Parameter Estimates for Covariates influencing Childhood Adiposity

	Coefficient (Equation 2) (Standard Error)			
	Weight at first Interview (kg)	Weight Category Age 3	Weight Category Age 5	Weight Category Age 7
		λ		
Family Lifestyle[a]	-0.051*** (0.007)	-1.205*** (0.079)	-1.535*** (0.080)	-1.518*** (0.078)
		δ		
Male	0.066*** (0.003)	-	-	-
Age (weeks)	0.004*** (0.001)	-	-	-
Black	-0.010 (0.012)	0.186 (0.113)	0.352*** (0.103)	0.339*** (0.101)
Asian	-0.077*** (0.007)	-0.262*** (0.083)	-0.091 (0.080)	0.096 (0.076)
Other	-0.028*** (0.009)	-0.011 (0.092)	-0.041 (0.097)	0.058 (0.098)
N			8,462	

This table shows the parameter estimates for variables having a direct influence on childhood adiposity. [a] These are the factor loadings for childhood adiposity, also displayed in Table 1. * $p < 0.01$, ** $p < 0.05$, *** $p < 0.001$.

Table 3: Parameter Estimates for Covariates Influencing Family Lifestyle

	Coefficient (Standard Error)			
Independent Variable	Initial Family Lifestyle	Family Lifestyle Age 3	Family Lifestyle Age 5	Family Lifestyle Age 7
		α (Equation 4)		
Previous Latent Family Lifestyle, α	-	1.094*** (0.007)	1.094*** (0.007)	1.094*** (0.007)
		β (Equations 3 and 4)		
Currently High SES	0.028*** (0.008)	-	-	-
Currently Low SES	-0.072*** (0.008)	-	-	-
Maternal Education at Birth	0.013*** (0.003)	-	-	-
Single Parent	-0.044*** (0.010)	-0.002 (0.007)	-0.003 (0.005)	-0.012** (0.005)

This table shows the autoregressive parameter on lifestyle and the coefficients for independent variables directly influencing underlying family lifestyle. * $p < 0.01$, ** $p < 0.05$, *** $p < 0.001$.

Table 4: Proportion of Families Remaining in Initial Lifestyle Percentile Group

Initial percentile	3 Years	5 Years	7 Years
≥ 95th	95.48%	91.27%	87.43%
≥ 90th	95.94%	92.77%	88.96%
≥ 75th	95.70%	93.84%	91.52%
Inter-quartile range	97.57%	96.46%	94.98%
< 25th	>99.99%	>99.99%	>99.99%
< 10th	99.99%	99.99%	99.76%
< 5th	>99.99%	>99.99%	99.99%
N		8,462	

Table 5: Characteristics of those at Top and Bottom of Family Lifestyle Rankings

	Initial Family Lifestyle Ranking	
Variable	≥ 95th percentile	< 5th percentile
Percentage Male	49.58%	51.34%
Mean Weight (kg) (standard deviation)	8.784 (1.444)	8.935 (1.513)
Percentage Obese Age 3	4.05%	6.01%
Percentage Obese Age 5	2.06%	6.44%
Percentage Obese Age 7	2.37%	8.37%
High SES at birth	83.99%	0.42%
Low SES at birth	0.14%	98.73%
N	8,462	

Table 6: Obesity Prevalence in Advantaged and Disadvantaged Children

	Advantaged (%)	Disadvantaged (%)
3 Years	3.79	6.43
5 Years	2.81	6.17
7 Years	2.59	6.42
N	8,462	

Table7: Lifestyle Percentiles in Advantaged and Disadvantaged Children

	Advantaged	Disadvantaged
9 months	85.63	7.50
3 Years	84.96	6.86
5 Years	84.59	6.29
7 Years	84.39	4.97
N	8,462	