Family Ties and Children Obesity in Italy

Crudu, Federico and Neri, Laura and Tiezzi, Silvia

Università di Siena, Università di Siena, Università di Siena

27 April 2018

Family Ties and Children Obesity in Italy[*]

Federico Crudu[†]   Laura Neri[‡]   Silvia Tiezzi[§]

University of Siena and CRENoS   University of Siena   University of Siena

October 2018

Abstract

This paper examines the impact of overweight family members on weight outcomes
of Italian children aged 6 to 14 years. We use an original dataset matching the
2012 cross sections of the Italian Multipurpose Household Survey and the Household
Budget Survey. Since identification of peer effects within the family is well known to
be a difficult challenge, we implement our analysis on a partially identified model by
means of valid inferential procedures recently introduced in the literature and based
on standard Bayesian computation methods. We find evidence of a strong, positive
effect of both overweight peer children in the family and overweight adults on children
weight outcomes. The impact of overweight peer children in the household is larger
than the impact of adults.

Key words: children obesity; peer effects within the family; partial identification; confidence sets.

JEL classification: I12, C15, C21, C35.

[†]Department of Economics and Statistics, University of Siena, Piazza San Francesco, 7/8 53100 Siena, federico.crudu@unisi.it.

[‡]Department of Economics and Statistics, University of Siena, Piazza San Francesco, 7/8 53100 Siena, laura.neri@unisi.it.

[§]Department of Economics and Statistics, University of Siena, Piazza San Francesco, 7/8 53100 Siena, silvia.tiezzi@unisi.it.

1 Introduction

In the past three decades children overweight and obesity prevalence have risen substantially in most high-income countries (Lobstein et al., 2015). Obesity rates are low in Italy compared to most OECD countries, but the picture is different for children. In fact, several stylized facts make the Italian case especially interesting. The rate of overweight children in Italy is one of the highest in OECD and non-OECD countries (OECD, 2015). Even though obesity prevalence has slightly decreased between 2008 and 2014, according to the fourth wave of the Italian Surveillance System OKkio alla Salute, in 2014 the rate of overweight primary school children was, respectively, 20.9% and 9.8%, with southern regions displaying higher prevalence compared to northern regions (Lauria et al., 2016).[1]

Paradoxically, the issue of childhood obesity is more prominent in Italy, where the Mediterranean diet is prevalent, than in other countries. Furthermore, even though maternal employment is usually associated with higher children weight outcomes, this is hardly the case in Italy where the labor market participation of mothers is much lower compared to other European countries (Brilli et al., 2016), while children obesity is higher. Family ties are culturally strong in Italy thus making social interaction within the family a particularly interesting issue to explore.

In addition to that, studying the determinants of children obesity is a compelling issue since it not only represents a direct threat for children's health and a cost to society, but it also has documented consequences for adult life, such as effects on self-esteem, body image and confidence, and lower wages as several habits take shape in early life and persist into adulthood (Schwartz et al., 2011).

The purpose of this paper is to examine the impact of overweight and obese members of the family, i.e. peer children and adults, on Italian children's weight outcomes. We consider children aged 6 to 14 years.

---

[1]The Surveillance System OKkio alla Salute (http://www.epicentro.iss.it/okkioallasalute/) monitors overweight and obesity of Italian children in primary schools (6-11 years of age). The System, promoted and financed by the Italian Ministry of Health, was started in 2007 and participates into the World Health Organization (WHO) European Childhood Obesity Surveillance Initiative (COSI).

No previous research that we are aware of has assessed the role of social interaction *within the family*, e.g. the role of the presence of other overweight and obese peer children or of overweight and obese adults in the family, as a determinant of childhood obesity despite the potential role for the presence of other overweight and obese family members in shaping children's weight outcomes.[2] If peer children in the family have important influences on children's weight outcomes, policies affecting one child in the family may have beneficial effects on the other children as well leading to a social multiplier effect. To our knowledge, only three papers, Nie et al. (2015), Asirvatham et al. (2014) and Gwozdz et al. (2015), deal with peer effects in children obesity and none of them explores social interaction *within the family*. Of these, only Gwozdz et al. (2015) deals with European data. This is surprising given the recognition that children consumption decisions are affected by those of their peers (Dishion & Tipsord, 2011) and the recent finding that peer effects are more pronounced in children compared to adolescents (Nie et al., 2015).

The impact of peer children in the family is a type of peer effect (Black et al., 2017). When the analysis focuses on a narrow peer group definition, like children in the same age class and family, the mechanism through which the peer effect plausibly operates is through imitation of good and bad behaviors such as eating habits (Nie et al., 2015), but also through shared limited parental resources (Black et al., 2017) since children grow up in the same household and are raised by the same parents.

Research in experimental psychology (Zmyj & Seehagen, 2013; Zmyj, Aschersleben et al., 2012; Zmyj, Daum et al., 2012) hypothesizes that prolonged individual experience with peers leads children to imitate peers more than adults. According to this literature, children imitate familiar behavior for social reasons, such as in order to identify with the model or to communicate likeness. Since age is an important indicator of the degree of being alike, children are more likely to imitate familiar behavior from peers than from adults.

---

[2]An exception is the famous study by Christakis & Fowler (2007) focusing on adults. One of their main finding was that, among pairs of adult siblings, if one sibling became obese, the chance that the other would become obese increased by 40% and that if one spouse became obese, the likelihood that the other spouse would also become obese increased by 37%.

When the contact with peers is prolonged, children may refer to their peers as valuable resources for learning also in unfamiliar situations. In this case imitation serves a cognitive function: prolonged contact makes peers a reliable model. Since children aged 6 to 14 plausibly spend extended periods of time with peer family members, prolonged contact is reflected in increased levels of peers imitation. If imitation behavior is the driving mechanism through which the peer effect operates, then the impact of peer children in the family should be larger than the impact of adults.

The impact of peer children in the family is particularly difficult to identify because of shared common traits and environments, and because of simultaneity effects (Black et al., 2017). To address this difficult identification problem we first ensure that a rich dataset at family level is available. We build a unique dataset resulting from Statistical Matching (SM henceforth) of the 2012 cross sections of two distinct surveys that share a set of variables and are representative of the same population. SM allows to integrate information on economic variables included in one of the two surveys into the other survey detailing family structure and composition as well as time use and weight outcomes of each family member.

Even though our data set provides a lot of information, the identification of peer effects remains a difficult task. Manski (1993) reports that in the context of the linear model peer effects are not identifiable due to the so called reflection effect. However, under certain conditions, when binary choice models are used, the reflection effect is not present (see, e.g., Blume et al., 2011; Brock & Durlauf, 2001, 2007). Despite the favourable features of binary choice models, due to the specific structure of the peer group, the identification problem is hard to solve. For such reasons, we use the partial identification results in Blume et al. (2011) for binary choice models. Clearly, providing valid inference for partially identified models is also a challenging task and a growing strand of literature has explored this issue. See for example Ciliberto & Tamer (2009), Chernozhukov et al. (2007), Chen et al. (in press) and references therein. Often, inferential procedures for partially identified models are rather complicated. However, the method we use in this paper, introduced in Chen et

al. (in press), is computationally rather simple and it boils down to calculating confidence sets for the parameters of interest by means of standard Bayesian computation methods.

We find evidence of a strong, positive and statistically significant effect of overweight and obese peer children and of a positive, smaller in magnitude and statistically significant effect of overweight and obese adults in the family on children's obesity. As a robustness check for the mechanism through which peer effects within the family might work, we split our sample of children between individuals aged 6 to 11 and those aged 12 to 14. If peers imitation increases with prolonged peer contact we should observe, as we do, a smaller peer effect for the second sub-group as pre-adolescents spend more time than their younger counterpart outside the family environment.

As stressed by Blume et al. (2011), the literature on partial identification for social interaction models has developed separately from that on the estimation of partially identified models via bounds initiated by Tamer (2003) and used in industrial organization. This paper is an attempt to integrate, in a very specific context, the two bodies of literature. Furthermore, this paper adds to a very scant literature on peer effects and children obesity. Finally, this is the only study on social interaction and childhood obesity in Italy.

The remainder of the paper unfolds as follows. Section 2 summarizes the literature. Section 3 describes the data and the statistical matching. Section 4 discusses our identification strategy. Section 5 presents the estimation methods and main results. Section 6 concludes.

2   Prior Research

The main recognized cause of the rise in children obesity is an imbalance between calories intake and calories expenditure. The factors driving this imbalance have been studied by a large literature. One strand of literature has addressed the relationship between maternal employment and children obesity in many developed countries (Cawley & Liu, 2012; Champion et al., 2012; Fertig et al., 2009; Gaina et al., 2009; García et al., 2006;

Greve, 2011; Gwozdz et al., 2013; Liu et al., 2009; Morrill, 2011, to cite only a few). Overall, these studies find empirical evidence of a positive relationship between maternal employment and childhood obesity. Maternal employment affects children weight outcomes through a number of channels. Cawley & Liu (2012), for example, find that employed women spend significantly less time eating and playing with their children and are more likely to purchase prepared foods. Fertig et al. (2009) find that maternal employment is related to children Body Mass Index (BMI) through the average number of meals consumed in a day, through reading/talking/listening to music and through TV watching. A related factor is the increasing use of non-parental child care (Hubbard, 2008; Herbst & Tekin, 2011). The growing use of non-parental care may play a crucial role in shaping children habits through quality of foods offered and the level of physical activity. Herbst & Tekin (2011) find that center-based care is associated with large and stable increases in BMI throughout its distribution, while the impact of other non-parental arrangements appears to be concentrated at the tails of the distribution. Hubbard (2008) also finds that using non-parental child-care (informal care from a relative, care from a baby-sitter, and center-based care) increases the likelihood of obesity.

The existing literature on children obesity in Italy mainly belongs to the medical sector. Binkin et al. (2010) estimated the prevalence of overweight and obese third-grade children by geographic area using the 2008 wave of the nationally represented nutrition survey OKkio alla Salute. In addition to explaining the high level of childhood obesity in the overall population - higher than that of most Western countries - they produced evidence of substantial geographic differences, with obesity prevalence twice as high in the South compared to the North. More recently, Bracale et al. (2013), using data from a special survey conducted in 2008, evaluated the prevalence of childhood overweight and obesity in a sample of school-age children (6-11 years of age) living in Milan and examined socio-cultural, parental and lifestyle factors associated with children's BMI that might affect the risk of obesity. While only moderate levels of above average weight and obesity were found among Milan children, this study confirmed the relevance of socio-cultural aspects

as factors affecting weight outcomes in a school-age children population.

A recent strand of literature, initiated by Christakis & Fowler (2007), has emerged in health economics that addresses the influence of social interaction, particularly of peers, on health status. Christakis & Fowler (2007) conducted a study to determine whether obesity might spread from person to person. Their starting point was that people embedded in social networks are influenced by the behaviors of those around them such that weight gain in one person might influence weight gain in others. A subsequent study by Fowler & Christakis (2008) produced evidence of person-to-person spread of obesity in adolescents. A recent review by Powell et al. (2015) has identified social contagion, i.e. the phenomenon whereby the network in which people are embedded influences their weight over time, as one of the social processes explaining the rise of adult overweight and obesity. The general finding is that weight-related behaviors of adolescents are affected by peer contacts (Fowler & Christakis, 2008; Halliday & Kwak, 2009; Mora & Gil, 2013; Renna et al., 2008; Trogdon et al., 2008). These studies take adolescents as the relevant age group and the classroom or friends as the relevant network. Much less is known about children as the relevant age group and the family as the relevant network.[3] We are aware of only three studies, besides ours, analyzing peer effects and children obesity. Asirvatham et al. (2014) study peer effects in elementary schools using measured obesity prevalence for cohorts within schools and using a panel dataset at grade level from Arkansas public schools. They found that changes in the obesity prevalence at the oldest grade are associated with changes in obesity prevalence at younger grades and the magnitude of the effect is greater in kindergarten to fourth-grade schools than in kindergarten to sixth-grade schools. Nie et al. (2015) analyze peer effects on obesity in a sample of 3 to 18 years old children and adolescents in China. Peer effects are found to be stronger in rural areas, among females and among individuals in the upper end of the BMI distribution. Finally, Gwozdz et al. (2015) analyze peer effects on childhood obesity using a panel of children aged 2 to 9 from eight European countries.

---

[3]Nie et al. (2015) report that most of the empirical literature on peer effects and obesity refers to adolescents or adults and uses US data.

They show that peer effects are larger in Spain, Italy, and Cyprus – compared to the other European countries in the sample. These studies adopt a fairly broad definition of peer effects, either peers at the same grade level within a school or children in a similar age group within a specific community. To the best of our knowledge, no study has yet used a very narrow peer group definition, such as children in the same age class in the same family. Even though the question may have crucial policy implications, the literature on the causal role of siblings on children's outcomes is small and very recent, because all the difficulties associated with peer effects estimation are magnified by the fact that it is difficult to disentangle the effect of one sibling on the other. This recent literature has focused on the effects of sibling health status on educational outcomes (Black et al., 2017; Fletcher et al., 2012), on the effect of early health shocks on child human capital formation (Yi et al., 2015) and on the effect of siblings on educational choices and early career earnings (Schrøter Joensen & Skyt Nielsen, 2018).

3   Data and Matching

The choice of the family as the relevant network to analyze peer effects complicates the problem of controlling for unobserved fixed effects. Thus, the amount of available information is a crucial issue in our case. Studies of peer effects and childhood obesity usually include information on economic characteristics of the household, such as income, in addition to personal and socio-demographic information as low-income individuals are more likely to be obese than high-income ones (Akee et al., 2013; Goisis et al., 2016). In addition, the relationship between income and weight is reported to vary by gender, race-ethnicity and age.[4] Due to the lack of a single Italian cross section containing individual weight outcomes, detailed family characteristics and socio-economic variables, we have used SM to match two datasets. The first is the 2012 cross section of the Multipurpose Survey on Households: Aspects of Daily Life (MSH) containing detailed information on family char-

---

[4]Food Research and Action Center: http://frac.org/obesity-health/relationship-poverty-obesity.

acteristics and the weight outcome of each member. The second is the 2012 cross-section of The Household Budget Survey (HBS) covering details of current and durable expenditures on more than 200 items. Both surveys are conducted by the Italian National Statistical Institute (ISTAT).

The MSH for year 2012 is a large nationally representative sample survey covering 19,330 households and 46,463 family members, including children aged 6 to 14 years. The questionnaire, administered by paper and pencil, contains 3 blocks of questions: a general questionnaire on individual characteristics of the first 6 members of the household; a family questionnaire collecting information about the household habits and lifestyles; a diary on health and nutritional information of each member of the household. For children and adolescents aged 6-17 a binary indicator for whether the child is overweight or obese is also included. The identification of a child as being overweight or obese is based on BMI threshold values for children aged 6 to 17 developed by Cole et al. (2000) and adopted by the International Obesity Task Force (IOTF). The MSH does not contain information on expenditures that could potentially be important covariates in our empirical model. We take this information from the 2012 cross-section of the HBS containing monthly consumption expenditures of 22,933 Italian households. ISTAT uses a weekly diary to collect expenditure data on frequently purchased items and a face-to-face interview to collect data on large and durable expenditures. Current expenditures are classified in about 200 elementary goods and services.

The survey also includes detailed information on the household structure and socio-demographic characteristics (such as regional location, household size, gender, age, education and employment condition of each household member). For both surveys, annual samples are independently drawn according to a two-stage design.[5] In addition to having a large set of variables in common, the two surveys share many characteristics such as the

---

[5]Details on the sampling procedure used to collect data in both surveys can be found in: ISTAT (2012) Indagine Multiscopo sulle Famiglie, aspetti della Vita quotidiana, Anno 2012, for the MSH survey; and in ISTAT (2012) File Standard-Indagine sui Consumi delle Famiglie-Manuale d'uso, anno 2012, for the HBS survey. Downloadable at $http://www.istat.it/it/archivio/4021$.

target population, sampling method, geographic frame, data collection procedure. These common characteristics allow us to use SM as an ideal method for combining information on households' quality of life and children weight outcomes with information on households' consumption expenditures.[6]

The unit of analysis is defined as children aged between 6 and 14 years. Our sample includes 3906 observations. For each individual a rich set of covariates is available.

Table 1 shows summary statistics of the relevant variables in the final dataset resulting from SM. In Table 1 we distinguish five sets of variables: individual characteristics of children (panel A), household characteristics (panel B), in some cases related to the household's reference person (RP), behavioral variables (panel C), proxies for genetic characteristics (panel D), regional variables (panel E). More specifically, the individual characteristics are the children overweight/obesity indicator (our dependent variable), gender and age of each child in the household. On average, overweight and obese children are 29% of Italian children aged 6-14 years. Male children are 49.5% and their mean age is 10 years. The mean size of their household is 4 persons. The share of other (overweight and obese) children in the family (excluding the considered child) is 7.2%, while the share of overweight and obese adult family members is 42%. Other included family controls are whether the parents consume soda drinks or smoke, whether the household lives in a central or northern Italian region, the employment status of the household RP, the logarithm of total monthly consumption expenditure of the household in Euros, whether the household is a single parent household, the level of education of the mother, the proportion of other children in the household watching TV every day, the share of siblings (excluding the child under consideration) aged between 6 and 18 having lunch at home, the share of siblings aged between 6 and 18 practicing physical activities, the share of siblings aged between 6 and 18 walking to go to school.[7] As proxies for the genetic variables we use the mean height and weight of the adult members of the family. Finally, additional variables are the 2012

---

[6]Statistical Matching of the two data sets is detailed in Appendix A.

[7]We use three dummies for whether the RP is employed, for whether she is a student or housewife, for whether she is retired or has another employment status.

	Mean	S.d.	Min.	Max.	Obs.
	A. Individual characteristics				
Child obesity	0.292	0.455	0	1	3,906
Age	9.987	2.595	6	14	3,906
Gender	0.495	0.500	0	1	3,906
	B. Household characteristics				
Share of other overweight/obese children	0.072	0.174	0	0.667	3,906
Share of overweight/obese adults	0.420	0.351	0	1	3,906
Household size	4.120	1.018	2	11	3,906
log expenditure (Euro)	7.478	0.630	5.469	9.741	3,906
Employed RP	0.813	0.390	0	1	3,906
Student or housewife RP	0.053	0.225	0	1	3,906
Retired or other emp. status RP	0.023	0.150	0	1	3,906
Single parent	0.119	0.326	0	1	3,906
Mother's education	3.724	1.439	1	6	3,376
	C. Behavioral characteristics				
Central or northern region	0.578	0.494	0	1	3,906
Siblings sport habit	0.176	0.244	0	0.75	3,906
Siblings lunch location	0.186	0.246	0	0.75	3,906
Siblings means of transport	0.089	0.193	0	0.75	3,906
Others TV watching habit	0.810	0.317	0	1	3,906
Parents soda drinks	0.155	0.362	0	1	3,906
Parents smoking	0.372	0.483	0	1	3,906
	D. Genetic characteristics				
Mean adult weight (kg)	70.676	9.159	35.5	117.5	3,906
Mean adult height (cm)	168.959	5.694	110	193	3,906
	E. Other characteristics				
CPI (2010=100)	106.022	0.614	104.6	108.1	3,906
% obese adults by region	24.954	5.476	17.7	36.1	3,906

Table 1: This table includes summary statistics on individual characteristics of children (panel A), household characteristics (panel B), in some cases related to the household's reference person (RP), behavioral variables (panel C), proxies for genetic characteristics (panel D), regional variables (panel E).

consumer price index (CPI) at the regional level (2010=100) and the percentage of obese adults by region in 2012.

4 Identification

We aim at assessing whether the presence of other overweight/obese family members, i.e. children in the same age group and adults, has a positive and significant effect on the probability of a child being overweight/obese. If imitation behavior is the driving mechanism we also expect that the impact of overweight/obese peer children in the family is larger than the impact of overweight/obese adults. The impact of children's weight outcomes on the weight outcome of other children in the same family is a type of peer effect. We use a narrow peer-group definition that includes all children aged 6 to 14 years belonging to the same family (whether siblings or not). Narrow definitions of peer groups have been found in the literature to be more endogenous than broad ones. In particular Trogdon et al. (2008) report that broader measures of social networks (e.g. grade-level peer groups) are more exogenous than narrow ones (e.g. children in the same family) as they are likely to be determined by different causal mechanisms. While grade-level peer effects may be driven by BMI related social norms and body image concerns, family-level peer effects may also operate through additional channels such as influences of diets, habits and physical activities. Christakis & Fowler (2007) showed that the influence of the weight of friends, family members and neighbors decreases as the degree of separation from the person under investigation increases. Despite the large empirical literature on social interaction in a variety of contexts, identifying such effects remains a formidable challenge.

Let us consider the notation and the definitions in Brock & Durlauf (2007). We assume that individual binary weight outcomes are determined by five factors:

1. observable individual-specific characteristics known also as the exogenous effects, measured by an r-vector X_i;

2. unobservable individual characteristics summarized by a scalar ε_i;

3. observable group characteristics, measured by an s-vector Y_g; these are known as contextual effects and may directly influence individual decisions: for example, peers'

characteristics such as parents' income, education or occupation may influence children's weight;

4. unobservable (to the econometrician) group characteristics, measured by a scalar α_g that may affect individual outcomes; these are known as correlated effects: for example, genetic characteristics may affect the weight of all children in the same family;

5. the average outcome in the peer group excluding the child under consideration, $m_{g_{-i}}$, also known as the endogenous effect as it describes how the behaviors of peers affects each individual outcome.

Thus, our model of social interaction can be described as

$$\omega_i = k + c'X_i + d'Y_g + Jm_{g_{-i}} + \varepsilon_i + \alpha_g \tag{1}$$

where ω_i is a binary indicator that takes value one if, according to a BMI score, individual i is overweight/obese and zero otherwise. Brock & Durlauf (2007) provide sufficient conditions under which identification of the parameters of interest is achieved. In some cases, if some of such conditions are not met, it is possible to achieve partial identification. This means that we can at least identify the parameter associated to the endogenous effect J. The conditions are detailed in Brock & Durlauf (2007) (see also Chapter 4 in Horowitz, 2009). Here, though, we focus on three assumptions that in our setting are particularly relevant, namely the assumption that Y_g contain variables that are continuous and take values in the real line, the assumption that no unobserved group variables that are determinants of ω_i are left in the error term ($\alpha_g = 0$) and the assumption of random assignment. The latter assumption implies that the distribution of the individual characteristics is independent of the group characteristics. Since in our context the members of a group are consanguineous with high probability, the random assignment assumption is likely to fail. Hence, if we want to achieve at least partial identification, we must ensure that $\alpha_g = 0$.

To do that we include in our model a large number of group characteristics that may reasonably determine obesity and that are either related to genetic factors or to behavioral factors.[8] So, for ease of exposition, we define

$$\alpha_g = \alpha_g^G + \alpha_g^B \tag{2}$$

where α_g^G is the component of the fixed effects that includes the genetic characteristics, while α_g^B is the component of the fixed effects that includes the behavioral characteristics.[9] Our data set includes a large number of variables that may be used to extract information from α_g^B.[10] We introduce a large number of variables controlling for these contextual effects such as the logarithm of total monthly consumption expenditure at the family level in Euros, the employment status of the household head, the structure of the household, i.e. household size and whether the household is a single parent household, other controls capturing family behavior in terms of physical activity and eating habits. We do not have any explicit information on the genetic factors that determine obesity in the group. However, we may proxy it by means of the mean adults weight and height in the household. Finally, the assumption on the support of Y_g in conjunction with the fact that $m_{g_{-i}}$ is bounded between zero and one implies that the binary choice model of social interactions does not suffer from the reflection effect (see Manski, 1993; Brock & Durlauf, 2007, page 59).

5   Estimation

In this section we describe the method, due to Chen et al. (in press), to estimate valid confidence sets and the associated numerical results collected in Table 2 and Table 3. The confidence sets are estimated by means of both a logit loglikelihood and a probit

---

[8]A recent strand of literature stresses that any similarity in weight due to shared household environments is undetectable and ignorable (Cawley & Meyerhoefer, 2012; Kinge, 2016; Wardle et al., 2008).

[9]In Tables 2 to 9 we also introduce α_g^O to indicate the variables in panel E of Table 1.

[10]Such variables would correspond to the characteristics described in Table 1, panels C and D.

loglikelihood. Moreover, they are compared to the corresponding standard confidence intervals.

5.1 Confidence Sets

The methods proposed in Chen et al. (in press) exploit some classical ideas of Bayesian computation. The estimation of the confidence sets is in fact based on sampling from the quasi-posterior distribution of the parameters. Here we provide a brief description of the method.[11] Let us consider a parametric loglikelihood function that depends on a parameter vector θ that takes values in a set Θ and the data Z_i

$$L_N(\theta) = \frac{1}{N} \sum_{i=1}^{N} \log f(\theta, Z_i).$$

Let us denote the identified set as $\Theta_I = \{\theta \in \Theta : F_0 = F_\theta\}$, where F_θ is our parametric model and F_0 is the true distribution of the data. The quasi-posterior distribution, say Π_N, of θ given the data Z is defined as

$$d\Pi_N(\theta, Z) = \frac{\exp\left(N L_N(\theta)\right) d\Pi(\theta)}{\int_\Theta \exp\left(N L_N(\theta)\right) d\Pi(\theta)}$$

where $\Pi(\theta)$ is a prior distribution. The $100\alpha\%$ confidence set, say $\widehat{\Theta}_\alpha$, for Θ_I is computed in a three step procedure:

1. draw B samples $\{\theta^{(1)}, \dots, \theta^{(B)}\}$ from the quasi posterior distribution Π_N via a Monte Carlo Markov chain (MCMC) sampler;[12]

2. calculate the $(1 - \alpha)$ quantile of $\{L_N(\theta^{(1)}), \dots, L_N(\theta^{(B)})\}$, say $\zeta_{N,\alpha}$;

3. define the confidence set as $\widehat{\Theta}_\alpha = \{\theta \in \Theta : L_N(\theta) \geq \zeta_{N,\alpha}\}$.

In our case, we use the loglikelihood function of both the probit model and the logit model.

---

[11]Chen et al. (in press) introduce three procedures to compute confidence sets. We use procedure 1.

[12]Chen et al. (in press) suggest using a sequential Monte Carlo sampler as MCMC may be numerically unstable. We do not experience such problems in our application.

Moreover, the resulting confidence sets have exact asymptotic coverage for the identified set Θ_I in potentially partially identified regular models (Chen et al., in press). The confidence sets are compared to the confidence intervals provided by the standard probit and logit models.

5.2   Results

Tables 2 to 9 contain 95% confidence intervals obtained using the standard logit and probit models as if identification were possible and 95% confidence sets obtained using the approach described in Section 5 and denoted as CCT.

We estimate the models for three subsets of the data. The first three columns in each Table refer to different specifications of the binary choice model including families with one child only and no peer effects. Columns 4 to 9 refer to families with at least two children and to three different specifications of the binary choice and CCT models. The estimation of these models serves the purpose of checking the stability of the model parameters' confidence sets. In columns 10 to 15 we combine the two previous sub-samples of families. We here comment on the last subset of estimates only.

Moreover, while Table 2 and Table 3 include confidence sets for children between 6 and 14 years of age, Tables 6 and 7, and Tables 8 and 9 consider children between 6 and 11 years of age and between 12 and 14 years of age respectively (Tables 6 to 9 are found in Appendix B).

Our dependent variable is a binary variable for a child being overweight/obese. The explanatory variables of interest are the share of other overweight and obese children in the household (the peer effect) and the share of overweight and obese adults in the household. Other included covariates are described in Section 3.

In most specifications in Tables 2 and 3, the confidence sets associated to the peer effect (*Share of other obese children*) do not include zero, have positive bounds and often-times display upper bounds around one. Thus we find a strong and positive effect of peer overweight and obese children in the household on the weight outcome of the child un-

der consideration. Interestingly, the confidence set in column 15 associated with the logit model is considerably shorter than the one associated with the probit model. In line with our predictions, we also find upper bounds for the share of overweight and obese adults in the family to be smaller than that of the peer effect across all models.

We conjecture the peer effect to be driven by imitation behavior. Recent experimental research (Zmyj & Seehagen, 2013; Zmyj, Ascherslebel et al., 2012; Zmyj, Daum et al., 2012) hypothesizes that prolonged individual experience with peers leads children to imitate peers more than adults. Our results support this hypothesis. Indeed we find confidence sets associated with the share of overweight and obese adults in the household on the weight outcome of children to be either zero or positive and smaller than the impact of the share of other overweight and obese children in the household across all specifications and models. To further corroborate this hypothesis, as a robustness check of the mechanism through which peer effects within the family might work, we split our sample between children aged 6 to 11 and those aged 12 to 14. If peers imitation increases with prolonged peer contact we should observe a smaller peer effect for the second sub-group as pre-adolescents spend more time outside the family environment than their younger counterparts. Confidence sets are shown in Tables 6 to 7 (for children aged 6 to 11) and Tables 8 to 9 (for children aged 12 to 14) in Appendix B. Again, we focus on the last 6 columns. Confidence sets for the share of peer children in the household are consistently positive and larger for the sample of children aged 6 to 11, while confidence sets for the share of overweight and obese adults in the household are consistently smaller across all models and specifications.

16

	One child families			Families with at least two children						All families					
	Logit	Logit	Logit	Logit	CCT	Logit	CCT	Logit	CCT	Logit	CCT	Logit	CCT	Logit	CCT
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
Share of other obese children				[1.085, 1.938]	[0.458,1.372]	[0.688, 1.576]	[0.406, 1.588]	[0.732, 1.668]	[0.515, 0.704]	[0.912, 1.699]	[0.369, 1.499]	[0.688, 1.576]	[0.406, 1.588]	[0.737, 1.656]	[0.469, 0.501]
gender	[0.047,0.532]	[0.020, 0.517]	[0.037, 0.537]	[0.084, 0.454]	[-0.081, 0.517]	[0.069, 0.448]	[-0.076, 0.516]	[0.066, 0.448]	[0.139, 0.374]	[0.133, 0.426]	[0.048, 0.418]	[0.069, 0.448]	[-0.076, 0.516]	[0.121, 0.422]	[0.166, 0.230]
age	[-0.090, 0.758]	[-0.136, 0.733]	[-0.128, 0.749]	[-0.266, 0.379]	[-0.379, 0.431]	[-0.242, 0.421]	[-0.420, 0.186]	[-0.221, 0.447]	[0.079, 0.130]	[-0.103, 0.407]	[-0.145, 0.361]	[-0.242, 0.421]	[-0.420, 0.186]	[-0.077, 0.449]	[0.262, 0.341]
age^2	[-0.043, -0.001]	[-0.043, -0.000]	[-0.043, -0.000]	[-0.025, 0.007]	[-0.028, 0.014]	[-0.028, 0.005]	[-0.016, 0.015]	[-0.029, 0.005]	[-0.012, -0.010]	[-0.026, -0.000]	[-0.024, 0.001]	[-0.028, 0.005]	[-0.016, 0.015]	[-0.029, -0.003]	[-0.023, -0.019]
Share of obese adults	[0.483, 1.178]	[0.359, 1.078]	[-0.657, 0.549]	[0.593, 1.127]	[0.338, 1.160]	[0.475, 1.027]	[0.289, 0.961]	[0.002, 0.865]	[0.131, 0.285]	[0.631, 1.053]	[0.465, 1.092]	[0.475, 1.027]	[0.289, 0.961]	[-0.066, 0.629]	[0.226, 0.334]
log expenditure (Euro)	[-0.285, 0.097]	[-0.165, 0.258]	[-0.156, 0.274]	[-0.323, -0.026]	[-0.362, 0.047]	[-0.152, 0.176]	[-0.227, 0.090]	[-0.141, 0.193]	[-0.184, 0.097]	[-0.278, -0.045]	[-0.340, -0.067]	[-0.152, 0.176]	[-0.227, 0.090]	[-0.092, 0.169]	[0.010, 0.071]
Y_g	NO	YES	YES	NO		YES		YES		NO		YES		YES	
α_g^O	NO	YES	YES	NO		YES		YES		NO		YES		YES	
α_g^G	NO	NO	YES	NO		NO		YES		NO		NO		YES	
α_g^B	NO	NO	YES	NO		NO		YES		NO		NO		YES	
N	1374	1374	1374	2297		2297		2297		3671		3671		3671	

Table 2: This table provides 95% confidence intervals using the standard logit model as if it were identified and 95% confidence sets using the approach described in Chen et al. (in press) and denoted as CCT. We consider three specifications of the model and three subsets of the initial data set. The first specification does not include group variables Y_g apart from those explicitly included in the table. Moreover, we exclude α_g^O (these variables are CPI and the percentage of obese adults by region), genetic fixed effects (α_g^G) and behavioral fixed effects (α_g^B). The second specification introduces Y_g and α_g^O, while in the third specification we also have α_g^G and α_g^B. The first data subset only includes families with only one child. This implicitly implies that there are no peer effects. In the second subset we use families that include at least two children, while the third sample combines the former two subsamples.

	One child families			Families with at least two children							All families			
												Dependent variable: child obesity			
	Probit (1)	Probit (2)	Probit (3)	Probit (4)	CCT (5)	Probit (6)	CCT (7)	Probit (8)	CCT (9)	Probit (10)	CCT (11)	Probit (12)	CCT (13)	Probit (14)	CCT (15)
Share of other obese children				[0.665, 1.183]	[0.295, 1.201]	[0.426, 0.961]	[0.165, 1.045]	[0.457, 1.019]	[0.159, 1.071]	[0.560, 1.041]	[0.338, 1.064]	[0.427, 0.962]	[0.165, 1.045]	[0.460, 1.012]	[0.105, 1.105]
gender	[0.038, 0.327]	[0.021, 0.315]	[0.029, 0.325]	[0.056, 0.277]	[-0.040, 0.378]	[0.048, 0.274]	[-0.073, 0.348]	[0.044, 0.271]	[-0.072, 0.332]	[0.086, 0.261]	[0.028, 0.310]	[0.048, 0.274]	[-0.073, 0.348]	[0.077, 0.256]	[-0.038, 0.336]
age	[-0.048, 0.459]	[-0.069, 0.449]	[-0.063, 0.458]	[-0.167, 0.217]	[-0.210, 0.290]	[-0.152, 0.240]	[-0.366, 0.212]	[-0.136, 0.259]	[-0.277, 0.340]	[-0.062, 0.241]	[-0.122, 0.279]	[-0.152, 0.240]	[-0.366, 0.212]	[-0.044, 0.268]	[-0.143, 0.335]
age^2	[-0.026, -0.001]	[-0.026, -0.000]	[-0.026, -0.001]	[-0.015, 0.005]	[-0.018, 0.007]	[-0.016, 0.004]	[-0.015, 0.016]	[-0.017, 0.003]	[-0.021, 0.011]	[-0.015, -0.000]	[-0.018, 0.003]	[-0.016, -0.004]	[-0.015, 0.016]	[-0.017, -0.002]	[-0.021, 0.004]
Share of obese adults	[0.293, 0.708]	[0.227, 0.654]	[-0.393, 0.324]	[0.364, 0.683]	[0.227, 0.795]	[0.287, 0.616]	[0.143, 0.688]	[-0.003, 0.510]	[-0.315, 0.632]	[0.383, 0.635]	[0.331, 0.680]	[0.288, 0.617]	[0.143, 0.688]	[-0.044, 0.370]	[-0.303, 0.463]
log expenditure (Euro)	[-0.171, 0.056]	[-0.099, 0.150]	[-0.092, 0.161]	[-0.196, -0.018]	[-0.247, 0.032]	[-0.093, 0.102]	[-0.202, 0.102]	[-0.083, 0.115]	[-0.213, 0.172]	[-0.168, -0.029]	[-0.201, -0.005]	[-0.092, 0.103]	[-0.202, 0.102]	[-0.054, 0.101]	[-0.111, 0.172]
Y_g	NO	YES	YES	NO		YES		YES		NO		YES		YES	
α_g^O	NO	YES	YES	NO		YES		YES		NO		YES		YES	
α_g^G	NO	NO	YES	NO		NO		YES		NO		NO		YES	
α_g^B	NO	NO	YES	NO		NO		YES		NO		NO		YES	
N	1374	1374	1374	2297		2297		2297		3671		3671		3671	

Table 3: This table provides 95% confidence intervals using the standard probit model as if it were identified and 95% confidence sets using the approach described in Chen et al. (in press) and denoted as CCT. We consider three specifications of the model and three subsets of the initial data set. The first specification does not include group variables Y_g apart from those explicitly included in the table. Moreover, we exclude α_g^O (these variables are CPI and the percentage of obese adults by region), genetic fixed effects (α_g^G) and behavioral fixed effects (α_g^B). The second specification introduces Y_g and α_g^O, while in the third specification we also have α_g^G and α_g^B. The first data subset only includes families with only one child. This implicitly implies that there are no peer effects. In the second subset we use families that include at least two children, while the third sample combines the former two subsamples.

6   Conclusions

This paper contributes to the literature on children obesity by assessing the effect of peers on children's weight outcomes in the context of a narrow peer group. Hence, we assess whether the presence of overweight and obese family members – other children and adults – affects children's weight outcomes. To our knowledge no study has yet analyzed the impact on children obesity of the obesity status of other members of their own family. Due to the fact that point identification is often not possible, we choose to carry out our analysis by means of a partially identified model. With respect to that aspect we contribute to the integration, even though in a rather specific context, of the literature on partial identification for social interaction models and that on partially identified models in industrial organization (Blume et al., 2011; Tamer, 2003).

We use a data set on Italian children resulting from statistical matching of the 2012 cross sections of two surveys, the Multipurpose Household Survey and the Household Budget Survey, both supplied by ISTAT. To provide valid inference for our partially identified models we use the method proposed by Chen et al. (in press). We find evidence of a strong, positive impact of both overweight and obese peer children in the family and of overweight and obese adults on children weight outcomes. Interestingly, in all empirical models we find that the impact of overweight and obese peer children in the household is larger than the impact of adults. Our results are consistent with studies on children imitation behavior and the age of the role model (Zmyj, Ascherslebel et al., 2012; Zmyj, Daum et al., 2012; Zmyj & Seehagen, 2013), where prolonged contact with peers leads children to imitate peers more than adults.

Despite the need to develop targeted approaches for obesity prevention in Italy, empirical evidence on the factors affecting Italian children weight outcomes remains poor. Further exploration of causal pathways linking social interaction within the family and children obesity is thus desirable. We show that social interaction within the family is one important causal factor explaining the prevalence of Italian children weight outcomes.

Appendix A Statistical Matching

In the basic framework SM integrates two data sources $\mathcal{A}$ and $\mathcal{B}$ drawn from the same target population (Cohen, 1991; Radner et al., 1980; Rodgers, 1984). $\mathcal{A}$ contains vector-valued variables (X, Y), whereas $\mathcal{B}$ contains vector-valued variables (X, Z) such that X is shared by both sources. SM uses the X variables common to both surveys as a bridge to create records containing (X, Y, Z) which can then be used to investigate the relationship between Y and Z (D'Orazio et al., 2006). In practice, matching procedures impute the target variables from a donor to a recipient survey. Our purpose was to integrate households' total consumption expenditure (*totexp*) from the HBS (denoted survey $\mathcal{A}$) into the MSH dataset (survey $\mathcal{B}$).

The first step was to identify the vector of matching variables X. Since $\mathcal{A}$ and $\mathcal{B}$ are representative samples of the same population, the common variables are expected to share the same marginal/joint distribution. This check was performed using the Cramer's V association measures. Potentially, all the variables identified and chosen according to this check could be used in the SM. In fact, just the most relevant ones have been identified and selected according to a linear model for predicting the logarithm of total consumption expenditure (*ltotexp*), our target variable. Table 4 shows the set of matching variables used to predict/impute the target variable. The listed common variables explain 70% of the total variability of the target variable. In addition to these, we include a number of interaction terms in the specification.

The next step was imputation from the donor to the recipient. The chosen imputation method was the Sequential Regression Multiple Imputation (Raghunathan et al., 2001) implemented by the software IVE-ware that allows imputation of missing data both in the recipient variables and in the set of matching variables. IVE-ware has a number of desirable properties that make it particularly well suited for imputing missing data in large datasets. For example, each imputation model can be specified according to the nature of the variable to be imputed. The software easily handles arbitrary missing data patterns

NO	Household living in the North West
NE	Household living in the North East
$Central$	Household living in the Centre
$Typfam$	Family type: single parent
$Typfam1$	Family type: both parents
$Typfam2$	Family type: Single
$n\_members$	Household size
$n\_members0\_5$	# persons 0-5 years old
$n\_members6\_17$	# persons 6-17 years old
$n\_members18\_34$	# persons 18-34 years old
$n\_members35\_65$	# persons 35-65 years old
$Gender\_RP$	Gender of the reference person
$Mstatus\_RP$	Marital status of the reference person
$Prof\_pos\_RP$	Professional position of the reference person
$Home$	Home ownership
$Rooms$	# of rooms
$Ec\_resource$	Adequacy of economic resources
PI	Gereral Price Index at regional level
PI_food	Gereral Price Index of Food at regional level

Table 4: Final matching variables.

with categorical and continuous variables. Finally, a sequential method is used to impute missing values: the variable with the least amount of missing data is imputed first and then used in subsequent imputations; the next variable with the second least amount of missing data is then imputed and used in subsequent imputations. The resulting quality of the matching can be assessed by comparing the marginal distribution of the target variable ($ltotexp$) in the observed data (i.e. in the donor dataset) and in the dataset obtained after the matching.

The distributions of the observed and imputed data are very close, as expected given the high explanatory power of the predictors included in the model of the target variable. This closeness increases the reliability of the statistical matching. In addition, it reduces the problem of conditional independence (CI), a required hypothesis for the validity of SM.

Given three random variables (X, Y, Z), and the model defined by (XY, XZ), the CI condition implies that any relationship between Y and Z can be explained by the set of matching variables X.
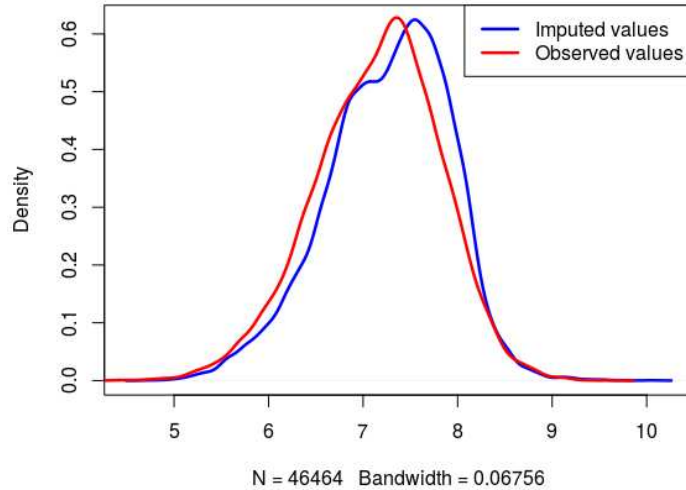
Figure 1: Probability density function of the observed target variable (*ltotexp*) collected in the HBS survey (2012) and probability density function of the imputed target variable in the final dataset.

In our study, the CI condition implies that any existing relationship between Y (household consumption expenditure) and Z (the binary indicator of a child obesity) can be explained by the set of matching variables X. In order to meet this assumption, a third data set is needed including complete information on (X, Y, Z). Since this dataset is normally not available, we test the CI assumption using a variable that expresses the subjective evaluation regarding the household's economic condition as a proxy for the household's total consumption expenditure. This is a binary indicator (*cond_econ_yes*), equal to 1 if the household's economic conditions are considered to be "very good" or "adequate" and equal to 0 if "scarce" or "absolutely insufficient".

The odds ratio of the logit model for obesity given *cond_econ_yes* indicates a statistically significant relationship (see column 1 in Table 5), implying that good or adequate economic resources at the household level reduce the probability of a child obesity. In Table 5 column 2 we consider a second model including *cond_econ_yes* and a subset of our matching variables as explanatory variables. All the covariates are significant, but the variable *cond_econ_yes* has odds ratio close to one implying a statistically insignificant

effect on obesity. A similar check has been conducted considering *cond_econ_yes* as the dependent variable and the obesity binary indicator as independent variable, obtaining similar results [13]. Thus, we conclude that results are consistent with the CI assumption.

	Dependent variable:	
	obechild	
	Logit	*Logit*
	(1)	(2)
constant	0.442***	$1.61e+07$**
	(0.022)	$(1.06e+08)$
cond_econ_yes	0.876*	0.934
	(0.062)	(0.070)
cn		0.526***
		(0.040)
age		1.256*
		(0.154)
age2		0.983***
		(0.006)
pgener		0.847***
		(0.052)
Observations	3,906	3,906
Pseudo R^2	0.001	0.030
Log Likelihood	-2,357.584	-2,288.024
LR Statistic	3.550* (df = 1)	126.220*** (df = 5)

*Note: standard errors in parentheses; *$p<0.1$; **$p<0.05$; ***$p<0.01$*

Table 5: Test of the CI assumption

---

[13]Available from the authors upon request.

Appendix B   Robustness Checks

	One child families			Families with at least two children						All families					
							Dependent variable: child obesity (6 ≤ age ≤ 11)								
	Logit (1)	Logit (2)	Logit (3)	Logit (4)	CCT (5)	Logit (6)	CCT (7)	Logit (8)	CCT (9)	Logit (10)	CCT (11)	Logit (12)	CCT (13)	Logit (14)	CCT (15)
Share of other obese children				[1.055, 2.071]	[0.291, 1.201]	[0.666, 1.730]	[0.004, 0.827]	[0.718, 1.847]	[0.271, 0.852]	[0.927, 1.888]	[0.278, 1.287]	[0.666, 1.730]	[0.004, 0.827]	[0.737, 1.849]	[0.402, 0.979]
gender	[-0.332, 0.272]	[-0.385, 0.239]	[-0.373, 0.255]	[-0.134, 0.288]	[-0.259, 0.292]	[-0.171, 0.265]	[-0.163, 0.184]	[-0.186, 0.253]	[0.009, 0.272]	[-0.129, 0.215]	[-0.225, 0.297]	[-0.171, 0.265]	[-0.163, 0.184]	[-0.159, 0.197]	[-0.057, 0.103]
age	[-0.787, 1.265]	[-0.737, 1.395]	[-0.794, 1.360]	[-1.126, 0.321]	[-0.318, 0.531]	[-1.248, 0.249]	[-0.507, -0.040]	[-1.269, 0.242]	[-0.443, 0.316]	[-0.795, 0.382]	[-0.350, 0.364]	[-1.248, 0.249]	[-0.507, -0.040]	[-0.822, 0.400]	[-0.128, 0.338]
age^2	[-0.077, 0.042]	[-0.086, 0.038]	[-0.084, 0.041]	[-0.023, 0.062]	[-0.038, 0.020]	[-0.019, 0.069]	[-0.003, 0.029]	[-0.019, 0.070]	[-0.026, 0.023]	[-0.026, 0.043]	[-0.028, 0.018]	[-0.019, 0.069]	[-0.003, 0.029]	[-0.028, 0.044]	[-0.022, 0.008]
Share of obese adults	[0.253, 1.126]	[0.136, 1.055]	[-0.607, 0.931]	[0.517, 1.123]	[0.262, 1.027]	[0.395, 1.029]	[0.390, 0.697]	[0.085, 1.082]	[-0.086, 0.674]	[0.532, 1.028]	[0.337, 1.004]	[0.395, 1.029]	[0.390, 0.697]	[0.045, 0.870]	[-0.212, 0.622]
log expenditure (Euro)	[-0.254, 0.225]	[-0.142, 0.403]	[-0.134, 0.416]	[-0.371, -0.033]	[-0.357, 0.029]	[-0.155, 0.220]	[-0.166, 0.041]	[-0.123, 0.259]	[-0.242, 0.094]	[-0.291, -0.019]	[-0.328, 0.050]	[-0.155, 0.220]	[-0.166, 0.041]	[-0.066, 0.243]	[-0.050, 0.100]
Y_g	NO	YES	YES	NO		YES		YES		NO		YES		YES	
α_g^O	NO	YES	YES	NO		YES		YES		NO		YES		YES	
α_g^G	NO	NO	YES	NO		NO		YES		NO		NO		YES	
α_g^B	NO	NO	YES	NO		NO		YES		NO		NO		YES	
N	789	789	789	1678		1678		1678		2467		2467		2467	

Table 6: This table provides 95% confidence intervals using the standard logit model as if it were identified and 95% confidence sets using the approach described in Chen et al. (in press) and denoted as CCT. We consider the subset of children aged 6 to 11 years. Moreover, three specifications of the model and three subsets of the initial data set are considered. The first specification does not include group variables Y_g apart from those explicitly included in the table. Moreover, we exclude α_g^O (these variables are CPI and the percentage of obese adults by region), genetic fixed effects (α_g^G) and behavioral fixed effects (α_g^B). The second specification introduces Y_g and α_g^O, while in the third specification we also have α_g^G and α_g^B. The first data subset only includes families with only one child. This implicitly implies that there are no peer effects. In the second subset we use families that include at least two children, while the third sample combines the former two subsamples.

| | One child families | | | Families with at least two children | | | | | | | | All families | | | |
	Probit (1)	Probit (2)	Probit (3)	Probit (4)	CCT (5)	Probit (6)	CCT (7)	Probit (8)	CCT (9)	Probit (10)	CCT (11)	Probit (12)	CCT (13)	Probit (14)	CCT (15)
Share of other obese children				[0.653, 1.279]	[0.239, 1.251]	[0.413, 1.061]	[0.133, 1.101]	[0.449, 1.134]	[0.111, 1.115]	[0.573, 1.168]	[0.346, 1.057]	[0.413, 1.060]	[0.133, 1.101]	[0.465, 1.414]	[0.004, 1.157]
gender	[-0.200, 0.168]	[-0.228, 0.149]	[-0.221, 0.157]	[-0.079, 0.176]	[-0.214, 0.277]	[-0.098, 0.162]	[-0.221, 0.270]	[-0.109, 0.153]	[-0.188, 0.247]	[-0.077, 0.131]	[-0.144, 0.184]	[-0.098, 0.162]	[-0.221, 0.270]	[-0.093, 0.120]	[-0.162, 0.199]
age	[-0.482, 0.772]	[-0.452, 0.838]	[-0.493, 0.812]	[-0.686, 0.188]	[-0.382, 0.432]	[-0.756, 0.139]	[-0.641, 0.175]	[-0.763, 0.139]	[-0.977, 0.424]	[-0.486, 0.228]	[-0.280, 0.318]	[-0.756, 0.139]	[-0.641, 0.175]	[-0.502, 0.232]	[-0.734, 0.503]
age^2	[-0.047, 0.026]	[-0.052, 0.023]	[-0.050, 0.026]	[-0.014, 0.038]	[-0.027, 0.021]	[-0.011, 0.042]	[-0.012, 0.036]	[-0.011, 0.042]	[-0.029, 0.056]	[-0.016, 0.026]	[-0.020, 0.015]	[-0.011, 0.042]	[-0.012, 0.036]	[-0.016, 0.027]	[-0.032, 0.041]
Share of obese adults	[0.151, 0.681]	[0.092, 0.647]	[-0.361, 0.570]	[0.318, 0.684]	[0.163, 0.871]	[0.240, 0.619]	[0.006, 0.757]	[0.041, 0.638]	[-0.154, 0.674]	[0.324, 0.624]	[0.248, 0.715]	[0.240, 0.619]	[0.006, 0.757]	[0.020, 0.515]	[-0.212, 0.573]
log expenditure (Euro)	[-0.153, 0.137]	[-0.085, 0.242]	[-0.080, 0.250]	[-0.226, -0.022]	[-0.262, 0.087]	[-0.096, 0.128]	[-0.229, 0.139]	[-0.074, 0.154]	[-0.160, 0.208]	[-0.177, -0.012]	[-0.181, 0.046]	[-0.096, 0.129]	[-0.229, 0.139]	[-0.040, 0.145]	[-0.143, 0.183]
Y_g	NO	YES	YES	NO		YES		YES		NO		YES		YES	
α_g^O	NO	YES	YES	NO		YES		YES		NO		YES		YES	
α_g^G	NO	NO	YES	NO		NO		YES		NO		NO		YES	
α_g^B	NO	NO	YES	NO		NO		YES		NO		NO		YES	
N	789	789	789	1678		1678		1678		2467		2467		2467	

Table 7: This table provides 95% confidence intervals using the standard probit model as if it were identified and 95% confidence sets using the approach described in Chen et al. (in press) and denoted as CCT. We consider the subset of children aged 6 to 11 years. Moreover, three specifications of the model and three subsets of the initial data set are considered. The first specification does not include group variables Y_g apart from those explicitly included in the table. Moreover, we exclude α_g^O (these variables are CPI and the percentage of obese adults by region), genetic fixed effects (α_g^G) and behavioral fixed effects (α_g^B). The second specification introduces Y_g and α_g^O, while in the third specification we also have α_g^G and α_g^B. The first data subset only includes families with only one child. This implicitly implies that there are no peer effects. In the second subset we use families that include at least two children, while the third sample combines the former two subsamples.

Table: Dependent variable: child obesity (12 ≤ age ≤ 14)

	One child families			Families with at least two children						All families					
	Logit (1)	Logit (2)	Logit (3)	Logit (4)	CCT (5)	Logit (6)	CCT (7)	Logit (8)	CCT (9)	Logit (10)	CCT (11)	Logit (12)	CCT (13)	Logit (14)	CCT (15)
Share of other obese children				[0.701, 2.333]	[-0.274, 1.162]	[0.381, 2.094]	[-0.179, 0.895]	[0.512, 2.354]	[-0.149, 0.228]	[0.477, 1.893]	[-0.073, 1.439]	[0.381, 2.094]	[-0.179, 0.895]	[0.482, 2.257]	[-0.410, 0.636]
gender	[0.466, 1.319]	[0.451, 1.335]	[0.445, 1.344]	[0.535, 1.345]	[0.070, 1.045]	[0.514, 1.346]	[0.090, 1.065]	[0.548, 1.410]	[0.029, 0.794]	[0.626, 1.210]	[0.346, 1.057]	[0.514, 1.346]	[0.090, 1.065]	[0.648, 1.257]	[0.492, 0.823]
age	[-14.494, 7.797]	[-12.915, 10.180]	[-14.470, 9.081]	[-14.401, 7.755]	[-0.595, 0.605]	[-14.057, 8.714]	[-0.524, 0.328]	[-12.784, 10.674]	[0.046, 0.170]	[-10.848, 4.812]	[-0.318, 0.625]	[-14.057, 8.714]	[-0.524, 0.328]	[-10.853, 5.418]	[-0.445, 0.394]
age²	[-0.311, 0.548]	[-0.403, 0.487]	[-0.361, 0.546]	[-0.305, 0.548]	[-0.042, 0.025]	[-0.344, 0.533]	[-0.026, 0.016]	[-0.420, 0.484]	[-0.020, -0.008]	[-0.194, 0.409]	[-0.039, 0.008]	[-0.344, 0.533]	[-0.026, 0.016]	[-0.219, 0.408]	[-0.034, 0.009]
Share of obese adults	[0.519, 1.695]	[0.367, 1.590]	[-1.208, 0.867]	[0.354, 1.506]	[-0.325, 1.189]	[0.196, 1.377]	[-0.217, 0.664]	[-0.984, 0.883]	[-0.245, 0.327]	[0.602, 1.421]	[0.305, 1.057]	[0.196, 1.377]	[-0.217, 0.664]	[-0.750, 0.615]	[-0.480, 0.234]
log expenditure (Euro)	[-0.573, 0.073]	[-0.490, 0.216]	[-0.517, 0.217]	[-0.401, 0.239]	[-0.580, 0.308]	[-0.394, 0.319]	[-0.410, 0.240]	[-0.439, 0.298]	[-0.036, 0.295]	[-0.410, 0.040]	[-0.530, 0.148]	[-0.394, 0.319]	[-0.410, 0.240]	[-0.339, 0.166]	[-0.370, 0.029]
Y_g	NO	YES	YES	NO	NO	YES	YES	YES	YES	NO	NO	YES	YES	YES	YES
α_g^O	NO	YES	YES	NO	NO	YES	YES	YES	YES	NO	NO	YES	YES	YES	YES
α_g^G	NO	NO	YES	NO	NO	NO	NO	YES	YES	NO	NO	NO	NO	YES	YES
α_g^B	NO	NO	YES	NO	NO	NO	NO	YES	YES	NO	NO	NO	NO	YES	YES
N	585	585	585	619	619	619	619	619	619	1204	1204	1204	1204	1204	1204

Table 8: This table provides 95% confidence intervals using the standard logit model as if it were identified and 95% confidence sets using the approach described in Chen et al. (in press) and denoted as CCT. We consider the subset of children aged 12 to 14 years. Moreover, three specifications of the model and three subsets of the initial data set are considered. The first specification does not include group variables Y_g apart from those explicitly included in the table. Moreover, we exclude α_g^O (these variables are CPI and the percentage of obese adults by region), genetic fixed effects (α_g^G) and behavioral fixed effects (α_g^B). The second specification introduces Y_g and α_g^O, while in the third specification we also have α_g^G and α_g^B. The first data subset only includes families with only one child. This implicitly implies that there are no peer effects. In the second subset we use families that include at least two children, while the third sample combines the former two subsamples.

	One child families			Families with at least two children						All families					
	Probit (1)	Probit (2)	Probit (3)	Probit (4)	CCT (5)	Probit (6)	CCT (7)	Probit (8)	CCT (9)	Probit (10)	CCT (11)	Probit (12)	CCT (13)	Probit (14)	CCT (15)
Share of other obese children				[1.085, 1.938]	[0.458, 1.372]	[0.688, 1.576]	[0.406, 1.588]	[0.732, 1.668]	[0.515, 0.704]	[0.912, 1.699]	[0.369, 1.499]	[0.688, 1.576]	[0.406, 1.588]	[0.737, 1.656]	[0.469, 0.501]
gender	[0.047, 0.532]	[0.020, 0.517]	[0.037, 0.537]	[0.084, 0.454]	[-0.081, 0.517]	[0.069, 0.448]	[-0.076, 0.516]	[0.066, 0.448]	[0.139, 0.374]	[0.133, 0.426]	[0.048, 0.418]	[0.069, 0.448]	[-0.076, 0.516]	[0.121, 0.422]	[0.166, 0.230]
age	[-0.090, 0.758]	[-0.136, 0.733]	[-0.128, 0.749]	[-0.266, 0.379]	[-0.379, 0.431]	[-0.242, 0.421]	[-0.420, 0.186]	[-0.221, 0.447]	[0.079, 0.130]	[-0.103, 0.407]	[-0.145, 0.361]	[-0.242, 0.421]	[-0.420, 0.186]	[-0.077, 0.449]	[0.262, 0.341]
age^2	[-0.043, -0.001]	[-0.043, 0.000]	[-0.043, 0.000]	[-0.025, 0.007]	[-0.028, 0.014]	[-0.028, 0.005]	[-0.016, 0.015]	[-0.029, 0.005]	[-0.012, -0.010]	[-0.026, 0.000]	[-0.024, 0.001]	[-0.028, 0.005]	[-0.016, 0.015]	[-0.029, -0.003]	[-0.023, -0.019]
Share of obese adults	[0.483, 1.178]	[0.359, 1.078]	[-0.657, 0.549]	[0.593, 1.127]	[0.338, 1.160]	[0.475, 1.027]	[0.289, 0.961]	[0.002, 0.865]	[0.131, 0.285]	[0.631, 1.053]	[0.465, 1.092]	[0.475, 1.027]	[0.289, 0.961]	[-0.066, 0.629]	[0.226, 0.334]
log expenditure (Euro)	[-0.285, 0.097]	[-0.165, 0.258]	[-0.156, 0.274]	[-0.323, -0.026]	[-0.362, 0.047]	[-0.152, 0.176]	[-0.227, 0.090]	[-0.141, 0.193]	[-0.184, 0.097]	[-0.278, -0.045]	[-0.340, -0.067]	[-0.152, 0.176]	[-0.227, 0.090]	[-0.092, 0.169]	[0.010, 0.071]
Y_g	NO	YES	YES	NO		YES		YES		NO		YES		YES	
α_g^O	NO	YES	YES	NO		YES		YES		NO		YES		YES	
α_g^G	NO	NO	YES	NO		NO		YES		NO		NO		YES	
α_g^B	NO	NO	YES	NO		NO		YES		NO		NO		YES	
N	585	585	585	619		619		619		1204		1204		1204	

Dependent variable: child obesity (12 ≤ age ≤ 14)

Table 9: This table provides 95% confidence intervals using the standard probit model as if it were identified and 95% confidence sets using the approach described in Chen et al. (in press) and denoted as CCT. We consider the subset of children aged 12 to 14 years. Moreover, three specifications of the model and three subsets of the initial data set are considered. The first specification does not include group variables Y_g apart from those explicitly included in the table. Moreover, we exclude α_g^O (these variables are CPI and the percentage of obese adults by region), genetic fixed effects (α_g^G) and behavioral fixed effects (α_g^B). The second specification introduces Y_g and α_g^O, while in the third specification we also have α_g^G and α_g^B. The first data subset only includes families with only one child. This implicitly implies that there are no peer effects. In the second subset we use families that include at least two children, while the third sample combines the former two subsamples.

References

Akee, R., Simeonova, E., Copeland, W., Angold, A. & Costello, E. J. (2013). Young adult obesity and household income: Effects of unconditional cash transfers. *American Economic Journal: Applied Economics*, *5*(2), 1–28.

Asirvatham, J., Nayga, R. M., Jr. & Thomsen, M. R. (2014). Peer-effects in obesity among public elementary school children: A grade-level analysis. *Applied Economic Perspectives and Policy*, *36*(3), 438-459.

Binkin, N., Fontana, G., Lamberti, A., Cattaneo, C., Baglio, G., Perra, A. & Spinelli, A. (2010). A national survey of the prevalence of childhood overweight and obesity in Italy. *Obesity Reviews*, *11*(1), 2–10.

Black, S., Breining, S., Figlio, D., Guryan, J., Karbownik, K., Skyt Nielsen, H., ... Simonsen, M. (2017). *Sibling spillovers* (Working Papers No. 23062). National Bureau of Economic Research.

Blume, L. E., Brock, W. A., Durlauf, S. N. & Ioannides, Y. M. (2011). Identification of social interactions. In *Handbook of social economics* (Vol. 1, pp. 853–964).

Bracale, R., Milani, L., Ferrara, E., Balzaretti, C., Valerio, A., Russo, V., ... Carruba, M. O. (2013). Childhood obesity, overweight and underweight: a study in primary schools in milan. *Eating and Weight Disorders - Studies on Anorexia, Bulimia and Obesity*, *18*(2), 183–191.

Brilli, Y., Del Boca, D. & Pronzato, C. (2016). Does child care availability play a role in maternal employment and children's development? Evidence from Italy. *Review of Economics of the Household*, *14*(1), 27-51.

Brock, W. A. & Durlauf, S. N. (2001). Discrete choice with social interactions. *The Review of Economic Studies*, *68*(2), 235–260.

Brock, W. A. & Durlauf, S. N. (2007). Identification of binary choice models with social interactions. *Journal of Econometrics*, *140*(1), 52–75.

Cawley, J. & Liu, F. (2012). Maternal employment and childhood obesity: A search for mechanisms in time use data. *Economics and Human Biology*, *10*(4), 352-364.

Cawley, J. & Meyerhoefer, G. (2012). The medical care costs of obesity: an instrumental variables approach. *Journal of Health Economics*, *31*(4), 210-230.

Champion, S. L., Rumbold, A. R., Steele, E. J., Giles, L. C., Davies, M. J. & Moore, V. M. (2012). Parental work schedules and child overweight and obesity. *International Journal of Obesity*, *36*, 573–580.

Chen, X., Christensen, T. & Tamer, E. (in press). Monte Carlo confidence sets for identified sets. *Econometrica*.

Chernozhukov, V., Hong, H. & Tamer, E. (2007). Estimation and confidence regions for parameter sets in econometric models. *Econometrica*, *75*(5), 1243–1284.

Christakis, N. A. & Fowler, J. H. (2007). The spread of obesity in a large social network over 32 years. *The New England Journal of Medicine*, *357*, 370–379.

Ciliberto, F. & Tamer, E. (2009). Market structure and multiple equilibria in airline markets. *Econometrica*, *77*(6), 1791–1828.

Cohen, M. L. (1991). Statistical matching and microsimulation models. In C. F. Citro & E. A. Hanushek (Eds.), *Improving information for social policy decisions – the uses of microsimulation modeling: Volume II, technical papers* (pp. 62–88). National Academy Press.

Cole, T. J., Bellizzi, M. C., Flegal, K. M. & Dietz, W. H. (2000). Establishing a standard definition for child overweight and obesity worldwide: international survey. *British Medical Journal*, *320*(7244), 1240.

Dishion, T. J. & Tipsord, J. M. (2011). Peer contagion in child and adolescent social and emotional development. *Annual Review of Psychology*, *62*(1), 189-214.

D'Orazio, M., Zio, M. D. & Scanu, M. (2006). *Statistical Matching: Theory and Practice (Wiley Series in Survey Methodology)*. John Wiley & Sons.

Fertig, A., Glomm, G. & Tchernis, R. (2009). The connection between maternal employment and childhood obesity: inspecting the mechanisms. *Review of Economics of the Household*, *7*(3), 227-255.

Fletcher, J., Hair, N. & Wolfe, B. (2012). *Am I my brother's keeper? Sibling spillover effects: the case of developmental disabilities and externalizing behavior* (Working Papers No. 18279). National Bureau of Economic Research.

Fowler, J. & Christakis, N. (2008). Estimating peer effects on health in social networks: A response to Cohen-Cole and Fletcher; and Trogdon, Nonnemaker, and Pais. *Journal of Health Economics*, *27*(5), 1400-1405.

Gaina, A., Sekine, M., Chandola, T., Marmot, M. & Kagamimori, S. (2009). Mother employment status and nutritional patterns in Japanese junior highschool children. *International Journal of Obesity*, *33*(7), 753–757.

García, E., Labeaga, J. M. & Masagué, A. C. O. (2006, July). *Maternal Employment and Childhood Obesity in Spain* (Working Papers No. 2006-17). FEDEA.

Goisis, A., Sacker, A. & Kelly, Y. (2016). Why are poorer children at higher risk of obesity and overweight? a uk cohort study. *The European Journal of Public Health*, *26*(1), 7–13.

Greve, J. (2011). New results on the effect of maternal work hours on children's overweight status: Does the quality of child care matter? *Labour Economics*, *18*(5), 579-590.

Gwozdz, W., Sousa-Poza, A., Reisch, L. A., Ahrens, W., Eiben, G., Fernandéz-Alvira, J. M., . . . others (2013). Maternal employment and childhood obesity–A European perspective. *Journal of Health Economics*, *32*(4), 728–742.

Gwozdz, W., Sousa-Poza, A., Reisch, L. A., Bammann, K., Eiben, G., Kourides, Y., . . . others (2015). Peer effects on obesity in a sample of European children. *Economics & Human Biology*, *18*, 139–152.

Halliday, T. & Kwak, S. (2009). Weight gain in adolescents and their peers. *Economics and Human Biology*, *7*(2), 181-190.

Herbst, C. M. & Tekin, E. (2011). Child care subsidies and childhood obesity. *Review of Economics of the Household*, *9*(3), 349–378.

Horowitz, J. L. (2009). *Semiparametric and nonparametric methods in econometrics* (Vol. 12). Springer.

Hubbard, M. (2008). The effect of mothers' employment and child care decisions on the body mass status of young children. *The University of North Carolina at Chapel Hill Working Paper*.

Kinge, J. (2016). Body mass index and employment status: a new look. *Economics and Human Biology*, *22*, 117-125.

Lauria, L., Pizzi, E., Andreozzi, S. & Galeone, D. (2016). Il Sistema di sorveglianza OKkio alla SALUTE: risultati 2014. *Istituto Superiore della Sanità Technical Report*.

Liu, E., Hsiao, C., Matsumoto, T. & Chou, S. (2009). Maternal full-time employment and overweight children: Parametric, semi-parametric, and non-parametric assessment. *Journal of Econometrics*, *152*(1), 61–69.

Lobstein, T., Jackson-Leach, R., Moodie, M. L., Hall, K. D., Gortmaker, S. L., Swinburn, B. A., . . . McPherson, K. (2015). Child and adolescent obesity: part of a bigger picture. *The Lancet*, *385*(9986), 2510–2520.

Manski, C. F. (1993). Identification of endogenous social effects: The reflection problem. *The review of economic studies*, *60*(3), 531–542.

Mora, T. & Gil, J. (2013). Peer effects in adolescent BMI: evidence from Spain. *Health Economics*, *22*(5), 501–516.

Morrill, M. S. (2011). The effects of maternal employment on the health of school-age children. *Journal of Health Economics*, *30*(2), 240–257.

Nie, P., Sousa-Poza, A. & He, X. (2015). Peer effects on childhood and adolescent obesity in China. *China Economic Review*, *35*, 47–69.

OECD. (2015). Overweight and obesity among children.

Powell, K., Wilcox, J., Clonan, A., Bissell, P., Preston, L., Peacock, M. & Holdsworth, M. (2015). The role of social networks in the development of overweight and obesity among adults: a scoping review. *BMC Public Health*, *15*(1), 996.

Radner, D., Allen, R., Gonzalez, M., Jabine, T. & Muller, H. (1980). Report on exact and statistical matching techniques. *Statistical Policy Working Paper 5. U.S. Department of of Commerce, Office of Federal Statistical Policy and Standards*.

Raghunathan, T. E., Lepkowski, J. M., Van Hoewyk, J. & Solenberger, P. (2001). A multivariate technique for multiply imputing missing values using a sequence of regression models. *Survey Methodology*, *27*(1), 85–96.

Renna, F., Grafova, I. & Thakur, N. (2008). The effect of friends on adolescent body weight. *Economics and Human Biology*, *6*(3), 377-387.

Rodgers, W. L. (1984). An evaluation of statistical matching. *Journal of Business & Economic Statistics*, *2*(1), 91–102.

Schrøter Joensen, J. & Skyt Nielsen, H. (2018). Spillovers in education choice. *Journal of Public Economics*, *157*, 158-183.

Schwartz, D. L., Chase, C. C., Oppezzo, M. A. & Chin, D. B. (2011). Practicing versus inventing with contrasting cases: The effects of telling first on learning and transfer. *Journal of Educational Psychology*, *103*(4), 759.

Tamer, E. (2003). Incomplete simultaneous discrete response model with multiple equilibria. *The Review of Economic Studies*, *70*(1), 147–165.

Trogdon, J. G., Nonnemaker, J. & Pais, J. (2008). Peer effects in adolescent overweight. *Journal of Health Economics*, *27*(5), 1388–1399.

Wardle, J., Carnell, S., Hawhort, C. & Plomin, R. (2008). Evidence for a strong genetic influence on childhood adiposity despite the force of the obesogenic environment. *American Journal of Clinical Nutrition*, *87*, 398-404.

Yi, J., Heckman, J., Zhang, J. & Conti, G. (2015). Early health shocks, intra-household resource allocation and child outcomes. *The Economic Journal*, *125*, F347-F371.

Zmyj, N., Ascherslebel, G., Prinz, W. & Daum, M. (2012). The peer model advantage in infants' imitation of familiar gestures performed by differently aged models. *Frontiers in Psychology*, *3*, 1-7.

Zmyj, N., Daum, M., Prinz, W., Nielsen, M. & Ascherslebel, G. (2012). Fourteen-months-old's imitation of differently aged models. *Infant and Child Development*, *21*, 250-266.

Zmyj, N. & Seehagen, S. (2013). The role of a model's age for the young children's imitation: a research review. *Infant and Child Development*, *22*, 622–641.