Heterogeneity and Relative Concerns
Three Essays in Applied Economics

by

Roberta Distante

BSc, University of Bologna (2004)
MSc, University of Bologna (2006)
MSc, University College London (2008)

Submitted to the Department of Economics, University of Copenhagen
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

at the

UNIVERSITY OF COPENHAGEN

Evaluation Committee:
Professor Hans Christian Kongsted, University of Copenhagen (chairman)
Professor Nabanita Datta Gupta, University of Århus, Denmark
Professor Francesco Moscone, Brunel University, London, UK

Heterogeneity and Relative Concerns
Three Essays in Applied Economics

by

Roberta Distante

Submitted to the Department of Economics, University of Copenhagen
on August 24, 2012, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Abstract

The present Ph.D. thesis is made of three self-contained chapters in the form of scholarly articles. It is centered around the development and implementation of advanced micro-econometric techniques for the study of microeconomic outcomes in presence of heterogeneity and relative concerns.

The first chapter uses micro-data for the British Household Panel Survey to shed light on the role of absolute and relative income for self-reported individual well-being. We estimate a heteroskedastic pooled panel ordered probit model with unobserved individual-specific effects. The study shows that a reference-group mean income effect exerts an influence on subjective well-being. The effect is asymmetric, with poorer individuals being more affected. The econometric strategy allows us to realize that the effect exerted by absolute income on happiness is the same for happiest and least happy individuals, and that relative income exerts a homogeneous impact on all happiness categories, though with different intensity. These findings provide stronger evidence on the importance of relative concerns for subjective well-being.

In the second chapter we analyze the importance of social ties for eating behavior of the young in the US, devising a dynamic framework capable of overcoming the problem of identifying social endogenous effects. Specifically, we propose a dynamic linear-in-means model to estimate social effects and control for individual- and group-specific unobservable effects, by exploiting stationarity restrictions of a system GMM estimator. We show that the main drivers of eating behavior are habituation and social effects. Furthermore, we analyze eating behavioral patterns from adolescence to adulthood, showing that obese teenagers become obese adults enforcing their wrong habits with imitative behavior. For adults who were normal-weight and overweight during adolescence, instead, the role of peers at school has a crucial importance for their current Body Mass Index.

The third chapter develops an approach for making welfare comparisons between populations with multidimensional discrete well-being indicators observed at the micro-level. It introduces an efficient algorithm for multivariate first order dominance and employs a bootstrap approach that allows for cardinal rankings of populations. These

techniques are applied to household survey data from Vietnam and Mozambique with a focus on the distribution and evolution of child poverty through space and time. The analysis of child poverty relates to an increased interest in developing child-focused measures and definitions with a multidimensional and human rights-based view.

Resumé

Denne Ph.D. afhandling består af tre selvstændige kapitler i form af videnskabelige artikler. Afhandlingen er centreret omkring udvikling og implementering af avancerede mikro-økonometriske teknikker til undersøgelse af mikroøkonomiske modeller, når der er heterogenitet mellem individer.

Det første kapitel bruger mikrodata fra det "British Household Panel Survey" til at belyse betydningen af absolut og relativ indkomst for selvrapporteret tilfredshed/lykke. Vi estimerer en ordered probit panel model med uobserverede individ-specifikke effekter. Undersøgelsen viser, at en reference-gruppes gennemsnitlige indkomst har betydning for det subjektivt velbefindende. Virkningen er asymmetrisk, hvor individer med ringe tilfredshed er mere påvirket. Den anvendte økonometriske strategi gør det muligt at undersøge, at virkningen af absolut indkomst på lykke er den samme for de lykkeligste og mindst lykkelige individer, og at den relative indkomst har en homogen effekt på alle lykkekategorier, dog med en anden intensitet. Disse resultater giver stærkere indikationer på betydningen af relativ indkomst for subjektivt velbefindende.

I det andet kapitel, analyserer vi betydningen af sociale bånd for spiseadfærd for unge i USA. Vi opstiller en model, hvori vi udnytter en dynamisk model til at løse problemet med at identificere de sociale endogene effekter. Konkret foreslår vi en dynamisk linear-in-mean model til at vurdere de sociale virkninger og de individuelle- og gruppespecifikke observerbare effekter. Ved at udnytte stationaritetsantagelsen kan en system GMM estimator anvendes til at estimere modellens parameter. Vi viser, at de vigtigste drivkræfter bag spiseadfærd er vaner og sociale effekter. Desuden analyserer vi spiseadfærdsmønstre i overgangen fra ungdommen til voksenlivet, og viser at meget overvægtige teenagere bliver overvægtige som voksne idet de fortsætter deres forkerte vaner med efterlignende adfærd. For voksne, som var normalvægtige eller overvægtige i løbet af ungdomsårene, er det især peer-effekts som har afgørende betydning for deres nuværende "Body Mass Index".

Det tredje kapitel udvikler en strategi til at lave sociale sammenligninger mellem befolkninger med flerdimensionale diskrete trivselsindikatorer observeret på mikro-niveau. Der indføres en effektiv algoritme for multivariate førsteorden dominans og anvender en bootstrap metode, der giver mulighed for kardinal placeringer af befolkningen. Disse teknikker anvendes på data fra en husstandsundersøgelser i Vietnam og Mozambique med fokus på fordelingen og udviklingen i børnefattigdom over tid. Dette kapitel relaterer sig til en øget interesse i at udvikle børnefokuserede foranstaltninger og arbejdet med at udvikle definitioner på fattigdom baseret på et flerdimensional mål som også er baseret på menneskerettighedsaspekter.

Contents

Acknowledgements

I would like to thank all people who have helped and inspired me during my doctoral studies. This dissertation would not have been possible without the guidance and the help of several individuals who, in one way or another, contributed and extended their valuable assistance in the preparation and completion of this study.

First and foremost, I want to thank my advisors, Lars Peter Østerdal and Mette Ejrnæs, for their unflinching guidance during my research and study at the University of Copenhagen. Their energy and enthusiasm, as well as their genuine commitment in research have been my constant source of inspiration.

I am grateful to my coauthors, Lars Peter Østerdal, Channing Arndt, Azhar Hussain, and Luisa Corrado. It was a honour for me to work with them, as I learnt the essence of good research.

Special thanks to my friends, Carla, Ina, Simona, Fiorenzo, Sara, Paolo, Paolino, Claudio, Annamaria, Steffan, Gosia, Peter, Elena, Luisa, Carmela, Jesper, Catia, Luigi, and to my lovely niece Cecilia, for their support, their smiles and their love. They have helped me to get through such a demanding path, especially in the most difficult and frustrating moments though partaking my joyful ones.

My family is my strength. My loving mother and father, and my grandparents have been always at my side, taking care of me and reinforcing my self-esteem. My brothers and my aunts have been source of peacefulness. I could never be grateful enough for what they have done for me.

My utmost gratitude goes to my husband, Emiliano Santoro, for his unflagging love and support throughout my life, and for giving me the strength to plod on despite wanting to give up. This dissertation is dedicated to him, who has always believed in me.

Roberta Distante
August, 2012

Chapter 1

Subjective Well-being, Income and Relative Concerns in the UK

Accepted for publication at *Social Indicators Research*. Available on line since June 9, 2012.

Abstract

We present an empirical model aimed at testing the relative income hypothesis and the effect of deprivation relative to mean income on subjective well-being. The main concern is to deal with subjective panel data in an ordered response model where error homoskedasticity is not assumed. A heteroskedastic pooled panel ordered probit model with unobserved individual-specific effects is applied to micro-data available in the British Household Panel Survey for 1996-2007. In this framework, absolute income impacts negatively on both completely satisfied and dissatisfied individuals, while relative income affects positively the most satisfied ones. Such an effect is asymmetric, impacting more severely on the relatively poor in the reference group. We argue that our results buttress the validity of the relative income hypothesis as an explanation of the happiness paradox.

1.1 Introduction

In recent years a new stream in the economic literature has boomed, which is mainly focused on explaining happiness determinants. Easterlin (1974) moved the first step towards a new conceptualization of happiness, overcoming the existing approaches built upon income-based measures of individual well-being. Easterlin's work sets out from a rather puzzling evidence, known as 'Easterlin Paradox' or 'Happiness Paradox': in developed countries, income is increasing while happiness levels are constant or decreasing. Economists believe that this puzzle should be unravelled by complementing income-based measures of welfare with alternative, more general measures of well-being (Graham, 2008). Among the factors supposed to play an important role for individual satisfaction with life are health, marital status, ethnicity, civic trust, and the so called 'relational goods,' which are referred to social aspects of life (see Dolan, Peasgood and White, 2008 for a review). Besides, economists are conscious that personal preferences for material goods are influenced by contextual effects that pertain to the social substrate and the environment individuals live in, hence by social comparison (e.g., *inter alia*, Firebaugh and Tach, 2000; Easterlin, 2001, 2003; McBride, 2001; Blanchflower and Oswald, 2004a, 2004b; Frijters, Haisken-DeNew and Shields, 2004a; Ferrer-i-Carbonell, 2005; Luttmer, 2005; Dynan and Ravina, 2007; Clark, Frijters and Shields, 2008; Boes, Staub and Winkelmann, 2010). However, the debate is still open on the validity of certain speculations, inasmuch other recent studies show that money matters for happiness by means of within- and between-country analyses (Frijters, Haisken-DeNew and Shields, 2004b; Headey, Muffels and Wooden, 2008; Stevenson and Wolfers, 2008; Sacks, Stevenson and Wolfers, 2010).

This paper builds upon the consideration that individual well-being is increasingly affected by social comparison, and that the relative income hypothesis – i.e., the fact that individuals derive utility from comparing their material achievements to income of others in a reference group – plausibly captures the complexity of the interplay between economic matters and individual well-being. Such a hypothesis has been tested in the happiness-related literature rather successfully (Ferrer-i-Carbonell, 2005; Luttmer, 2005; Clark, Frijters and Shields, 2008), as the usage of panel data allowed to control for unobservable conditions, situations or events, and personality traits (time-invariant individual effects). Indeed, accounting for individual heterogeneity is of paramount importance when dealing with unobservable outcomes inherently intertwined to inner life, as individual satisfaction (see Ferrer-i-Carbonell and Frijters, 2004 for a review on methodology in happiness-related studies).

Our aim is to convey robuster insights on the relationship between relative income and happiness, by pursuing a panel data analysis where we take heed of individual heterogeneity in two ways: by controlling for unobserved individual-specific effects, and by relaxing the homoskedasticity assumption. In particular, a heteroskedastic pooled panel ordered probit (HPPOP, henceforth) augmented to control for unobserved individual-specific effects is estimated using data from the British Household Panel Survey, and a large number of control variables (i.e., health, both at the subjective and objective levels, marital status, having children, age, gender, and employment status) are included.[1] We model the error variance to depend non-linearly on some time-varying and time-invariant observable factors of interest; this way, we are able to better account for heterogeneity in choices where individual-specific, time-invariant effects are not sufficient to capture unexplained error variation due to time-varying objective and subjective conditions spoiling individual perception. Also, in this framework we overcome ordinal probit's design rigidities in the analysis of marginal probability effects, and are able to show that absolute income impacts negatively on the probability of being generally unhappy as well as on the probability of being completely happy. Relative income, which is computed as the ratio between absolute and comparison income, appears to have a positive relationship with self-reported well-being, meaning that comparison income is negatively related to the level of self-reported satisfaction: in each reference group the

[1]The heteroskedastic ordered probit is also known as heterogeneous choice/ location-scale ordinal probit. We conied the term heteroskedastic pooled panel ordered probit for synthesizing the features of the model we use: an ordered probit, pooled, but still allowing more robustness than cross-sectional analyses (panel), and controlling for potential heteroskedasticity (heteroskedastic).

(relatively) rich and the (relatively) poor are both less satisfied if the comparison income increases. Such an effect is asymmetric: including a deprivation measure, we find that the mean income impact is severer for the poor, *ceteris paribus*.[2] In sum, our econometric analysis shows how the effect exerted by absolute income on happiness is the same for the extreme categories of response, and that relative income has a homogeneous impact on all response categories − though with a different intensity. These findings provide further and stronger evidence on the importance of relative concerns for subjective well-being: income matters for happiness if compared to a benchmark as individuals measure their own achievements in comparison to a general standard of living − the eponymous 'keeping up with the Joneses'.

The remainder of the present paper is laid out as follows: Section 2 reports technical details of the econometric framework used; Section 3 is devoted to data description, and overviews hypotheses and specification of the estimation models; Section 4 is dedicated to the estimation results; Section 5 concludes.

1.2 The Econometric Framework

As extensively discussed by Ferrer-i-Carbonell and Frijters (2004), the econometric models used in happiness economics generally present an ordinal latent-variable specification. The error term may be assumed to be either Normal or Logistic, this leading respectively to an ordered probit or logit. This framework is the most popular (for example, ordered probit analyses are pursued, among others, by Clark and Oswald,1994; Blanchflower and Oswald; 2001; Frey and Stutzer, 1999; 2000; while Winkelmann and Winkelmann, 1998; Blanchflower and Oswald; 2004b; Alesina, Di Tella and MacCulloch, 2004; rely on ordered logit models). Usually fixed effects are not directly included in the regression, provided the estimates obtained are inconsistent (Maddala, 1983). Other noteworthy studies are the ones by Winkelmann and Winkelmann (1998) where a conditional max-

[2]The deprivation measure consists of a multiplicative term which includes a dummy and relative income. The dummy takes on the unity when personal income is below the reference one.

imum likelihood estimator for a fixed effects logit model is implemented dichotomizing the dependent variable, and by Ferrer-i-Carbonell and Frijters (2004), who augment the Winkelmann and Winkelmann (1998) estimator with individual-specific thresholds. In alternative to the frameworks presented so far, other contributions assume a structural relationship existing between time-invariant variables and time-varying ones, including individual random time-invariant effects in ordered response models (Ferrer-i-Carbonell, 2005).

Relying on the achievements of the literature surveyed, and in consideration of the ordinal nature of subjective well-being data, we argue that analyses based on ordered discrete choice models should provide a better fit. Furthermore, we think that individual fixed effects as well as heteroskedasticity in choices need to be controlled for. First of all, we presume that it is appropriate to keep the ordered structure of the dependent, self-reported variable proxying well-being (generally life satisfaction), rather than conforming to other panel data analyses where the same variable is dichotomized;[3] this is because ordinal variables embed more information than binary ones. Secondly, given the strong heterogeneity of people surveyed, exacerbated by the psychological nature of such matter, the econometric framework needs to account for unobservable individual effects. Lastly, we want to avoid the assumption that error variances are the same for all cases, which might entail biased parameter estimates.

Therefore, in the remainder we specify a HPPOP model, which is augmented to

[3] Life satisfaction is thought of as being a good proxy for welfare, a more general concept the researchers actually focus on.

Also, life satisfaction is presumed to be ordinally comparable between individuals. Loosely speaking, we can recognize if any two individuals are better off, worse off or equally well off in terms of welfare. This implies that happiness is a concept perceived much the same way. Being life satisfaction a monotonic transformation of welfare, we are able to discern happier individuals from less happy ones.

Lastly, a cardinal comparability of life satisfaction (preferences) between individuals is assumed to be possible. This means assuming that the difference between any two consecutive scores in the satisfaction scale is the same regardless of the rank. Such a hypothesis is not very widespread for its perversity to the standard microeconomic theory. Indeed, a controversy on happiness (or utility) cardinal measurability exists in this literature. In these regards, Ferrer-i-Carbonell and Frijters (2004) produce evidence that the assumption of cardinality of life satisfaction scores has a negligible impact on empirical results. Indeed, we argue that such an assumption is closely related to the econometric method used for the empirical analysis, and that when ordinal discrete models are used, cardinality is not a major concern.

account for unobserved time-invariant individual effects. We control for unobserved effects which are neither considered as parameters to estimate nor as having a certain distribution and being independent from all covariates, accommodating the model by Mundlak (1978) to our case. In this way we do control for fixed effects, as Mundlak (1978) shows in his original article, where a modified random coefficients model leads to a 'within' estimator identical to the fixed effect estimator of the basic specification when unobserved effects are assumed to be normally distributed conditional on the covariates. In addition, we explicitly specify the determinants of heteroskedasticity in an attempt to correct for it. This leads to joint estimation of the explanators of heterogeneity and the explanators associated with choices.

1.2.1 Baseline Setting

Hereinafter, we explain the basic pooled panel ordered probit (PPOP, henceforth) in its standard form. Formally, the ordered categorical outcome for the variable life satisfaction S_{nt} is coded in a rank preserving manner:

$$S_{nt} \in \{1, 2, ..., j, ..., J\}$$

where we implicitly assumed repeated measurements ($t = 1, ..., T$) for a sample of N individuals ($n = 1, ..., N$). The vector of covariates $\mathbf{x}$ is, say, of dimension $(1 \times k)$. The cumulative probabilities of the outcome are linked to a single index of independent variables as follows:

$$\Pr(S_{nt} \leq j|\mathbf{x}_{nt}) = \Phi(\alpha_j - \mathbf{x}_{nt}\boldsymbol{\beta}),$$

where α_j and β are unknown parameters and Φ is the standard normal cumulative density function.

Well-defined probabilities are ensured if $\alpha_j > \alpha_{j-1}$, $\alpha_J = \infty$ such that $\Phi(\infty) = 1$ and $\alpha_0 = -\infty$ such that $\Phi(-\infty) = 0$. Ordered response models are expressed by means of an underlying continuous latent process S^*_{nt} and a response scheme:

$$
\begin{aligned}
S^*_{nt} &= \mathbf{x}_{nt}\beta + \epsilon_{nt} \qquad (1.1)\\
S_{nt} &= j \text{ iff } \alpha_{j-1} < S^*_{nt} = \mathbf{x}_{nt}\beta + \epsilon_{nt} < \alpha_j,\ j = 1, 2, ..., J,
\end{aligned}
$$

where S^*_{nt} represents the real line that is discretized in J categories by the threshold parameters α_j and it is in linear relation with observables and unobservables, the latter assumed to be distributed as a standard normal, $\Phi(\epsilon_{it})$. The estimated parameters are to be interpreted as indicative of the sign but not the magnitude of the effect. Indeed, conditional probabilities are crucial in this kind of analyses; they read as follows[4]:

$$
\Pr(S = j|X = \mathbf{x}) = \Phi(\alpha_j - \mathbf{x}\boldsymbol{\beta}) - \Phi(\alpha_{j-1} - \mathbf{x}\boldsymbol{\beta}).
$$

For identifying the parameters we need to assume that $\mathbf{x}$ does not contain a constant, this aimed at fixing the location of the arguments in Φ (Boes and Winkelmann, 2006a).

We are interested in understanding how a marginal variation in one covariate produces a change in the cumulative distribution of the dependent, thus a variation in all the outcome probabilities. For a continuous regressor x_h the marginal effects are computed as follows:

$$
M_{jh}(\mathbf{x}) = \frac{\partial \Pr(S = j|X = \mathbf{x})}{\partial x_h} = \left[\phi(\alpha_{j-1} - \mathbf{x}\boldsymbol{\beta}) - \phi(\alpha_j - \mathbf{x}\boldsymbol{\beta})\right]\beta_h,
$$

where $\phi(\cdot)$ is the standard normal probability density function. If the regressor is discrete, we compute the variation in probability before and after the discrete change:

$$
\Delta \Pr(S = j|X = \mathbf{x}) = \Pr(S = j|X = \mathbf{x} + \Delta x_h) - \Pr(S = j|X = \mathbf{x}).
$$

The size of the effects on the outcome probabilities depends on the values that the n^{th} observation takes on.

[4]Henceforth in this subsection we disregard subscripts for expositional neatness; the specification refers to individual n at period t.

The values at which the partial effects are to be evaluated are the means of the independent variables. The way to consistently estimate the average partial effects is to replace the population parameters with the estimates obtained by maximum likelihood and compute the average over the whole sample of observations.

A note is due on the limits of the ordered response models, because the ratio between the marginal probability effects of two different continuous regressors on the same response choice remains constant across individuals. Moreover, due to the shape of the normal distribution, we observe that the sign of marginal probability effects changes only once from the lowest to the highest category, being first negative and then positive or *vice versa*. Indeed, it is difficult to understand the effects for the categories included between the first and the last.

1.2.2 Extensions to the Baseline Setting

We operate two main adjustments to our baseline setting by introducing unobserved individual effects and controlling for potential heteroskedasticity of the errors.

Specifically, when unobserved individual specific effects are assumed to exist, the specification of the PPOP model changes as follows:

$$\begin{aligned} S^*_{nt} &= \xi_n + \mathbf{x}_{nt}\boldsymbol{\beta} + \epsilon_{nt} \ , \qquad (1.2) \\ n &= 1, ..., N; \quad t = 1, ..., T. \end{aligned}$$

In a linear model ξ_n would be eliminated by a first difference estimation or by a within-transformation. The ordered probit, instead, given its non-linear form, does not permit similar methods. Applying a dummy variable approach is not advisable either, mainly for two reasons: too many degrees of freedom are lost in this case and the incidental parameters problem[5] would lead to inconsistent estimators.

[5] In fixed effects models, the number of parameters increases with the number of individuals, because we estimate them as unknown parameters. When n becomes large, but T is finite, the maximum likelihood estimator is inconsistent.

What we do for taking into account unobserved individual effects is modeling the conditional distribution of such a term with respect to the covariates:

$\xi_n|\mathbf{x}_n \backsim N(\overline{\mathbf{x}}_n\boldsymbol{\gamma}, \sigma^2_{\varpi})$, where $\overline{\mathbf{x}}_n$ is the average over time of $\mathbf{x}_{nt}$, and σ^2_{ϖ} is an unknown parameter. In other terms, $\xi_n = \overline{\mathbf{x}}_n\boldsymbol{\gamma} + \varpi_n$, where ϖ_n is an orthogonal error with $\varpi_n \mid \mathbf{x}_n \backsim N(0, \sigma^2_{\varpi})$.

In practice, we extend the approach *à la* Mundlak (1978) to an ordered setting. Mundlak originally proposes a modified random coefficients model in which unobserved effects are assumed to be normally distributed conditional on the mean of the covariates, thus obtaining a 'within' estimator in the random effects framework. In Mundlak's specification the error distribution is symmetrical, thus the resulting GLS estimator is identical to the fixed effect estimator of the basic specification. Therefore it is unbiased (Hsiao, 1986).

The other adjustment regards the error term. We model the error variance structure, as suggested in the literature on heterogeneous choice models, assuming that $\epsilon_{nt}|\mathbf{x}_{nt} \backsim iiN(0, \sigma^2_{\epsilon})$, where $\sigma^2_{\epsilon} = \exp(\mathbf{z}_{nt}\boldsymbol{\vartheta})^2$. The vector $\mathbf{z}_{nt}$ can contain all the variables that the researcher considers as possible sources of heteroskedasticity, even variables already included in the set of regressors. Such a method should avert potential heteroskedasticity to bias our results. Heteroskedastic models like this one have been frequently used to explore heterogenous behaviors (Alvarez and Brehm, 1997, 1998, 2000; Busch and Reinhardt, 1999; Gabel, 1998; Lee, 2002; Krutz, 2005). So far, heteroskedastic probit and heteroskedastic ordered probit models are the most used tools in investigating discrete heterogenous choices. The advantage of these models is the ability to cure probit with non-homogeneous error variances or to test hypotheses about heterogenous choices that immediately relate to σ^2_{ϵ} (Keele and Park, 2006).

Back to our model, all parameters are now scaled by

$$(\sigma^2_{\epsilon} + \sigma^2_{\varpi})^{-1/2} = (\exp(\mathbf{z}_{nt}\vartheta)^2 + \sigma^2_{\varpi})^{-1/2}$$

that will be denoted with $\Omega_{nt(z)}$. By assuming that the individual-specific effects are normally distributed conditional on the individual means of time-varying covariates, we end up with a sum of normal variables; the response probabilities for individual n at period t, $p_j(\mathbf{x},\mathbf{z}) = \Pr(S = j \mid X = \mathbf{x}, Z = \mathbf{z})$, look like:

$$
\begin{aligned}
&p_1(\mathbf{x},\mathbf{z}) = \Phi\left[(\alpha_1 - \mathbf{x}\boldsymbol{\beta} - \overline{\mathbf{x}}\gamma)\cdot\Omega_{(z)}\right] \\
&p_2(\mathbf{x},\mathbf{z}) = \Phi\left[(\alpha_2 - \mathbf{x}\boldsymbol{\beta} - \overline{\mathbf{x}}\gamma)\cdot\Omega_{(z)}\right] - \Phi\left[(\alpha_1 - \mathbf{x}\boldsymbol{\beta} - \overline{\mathbf{x}}\gamma)\cdot\Omega_{(z)}\right] \\
&\ldots \\
&p_{J-1}(\mathbf{x},\mathbf{z}) = \Phi\left[(\alpha_{J-1} - \mathbf{x}\boldsymbol{\beta} - \overline{\mathbf{x}}\gamma)\cdot\Omega_{(z)}\right] - \Phi\left[(\alpha_{J-2}\mathbf{x}\boldsymbol{\beta} - \overline{\mathbf{x}}\gamma)\cdot\Omega_{(z)}\right] \\
&p_J(\mathbf{x},\mathbf{z}) = 1 - \Phi\left[(\alpha_J - \mathbf{x}\boldsymbol{\beta} - \overline{\mathbf{x}}\gamma)\cdot\Omega_{(z)}\right].
\end{aligned}
$$

The joint distribution of $(S_{n1}, ..., S_{nT})$ conditional on the explanatory variables is obtained by integrating ϖ_n out in the response probabilities:

$$
f\left(S_{n1}, ..., S_{nT}\right) = \int_{-\infty}^{+\infty} \prod_{t=1}^{T} \prod_{j=1}^{J} p_j(\mathbf{x},\mathbf{z})^{\mathbf{1}(S_{nt}=j)} \frac{1}{\sigma_{\varpi}^2} \phi\left(\frac{\varpi_n}{\sigma_{\varpi}^2}\right) d\varpi_n.
$$

The parameters $\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{\vartheta}$ and σ_{ϖ}^2 are estimated by maximum likelihood, the total partial log-likelihood function reading as:

$$
\ell(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{\vartheta}, \sigma_{\varpi}^2 \mid \mathbf{x}, \mathbf{z}) = \sum_{n=1}^{N} f\left(S_{n1}, ..., S_{nT}\right).
$$

Without further assumptions, a robust variance matrix estimator is needed to account for serial correlation in the scores across the time periods. Indeed, we adjust robust standard errors for clustering at the individual level, i.e. correct for correlation between responses of the same individual across time periods.

As to the marginal partial effects, it is straightforward to see how their magnitude and sign are dependent on the inclusion of a function for modeling the error variance. The first case to be considered is that of continuous variables included in $\mathbf{z}$ when such vector is a subset of $\mathbf{x}$. Consider the marginal effect of $x_h \in \mathbf{z} \subseteq \mathbf{x}$:

$$
\begin{aligned}
M_{jh}(\mathbf{x}) &= \frac{\partial \Pr(S=j|X=\mathbf{x}, Z=\mathbf{z})}{\partial x_h} = \qquad (1.3)\\
&= \phi\left[(\alpha_{j-1} - \mathbf{x}\boldsymbol{\beta} - \overline{\mathbf{x}}\gamma)\cdot\Omega_{(z)}\right]\left\{\beta_h\cdot\Omega_{(z)} + \theta_h\cdot(\alpha_{j-1} - \mathbf{x}\boldsymbol{\beta} - \overline{\mathbf{x}}\gamma)\exp(\mathbf{z}\boldsymbol{\vartheta})^2\mathbf{z}\boldsymbol{\vartheta}\right\}\\
&\quad -\phi\left[(\alpha_{j} - \mathbf{x}\boldsymbol{\beta} - \overline{\mathbf{x}}\gamma)\cdot\Omega_{(z)}\right]\left\{\beta_h\cdot\Omega_{(z)} + \theta_h\cdot(\alpha_{j} - \mathbf{x}\boldsymbol{\beta} - \overline{\mathbf{x}}\gamma)\exp(\mathbf{z}\boldsymbol{\vartheta})^2\mathbf{z}\boldsymbol{\vartheta}\right\},
\end{aligned}
$$

where the mean component for x_h is considered to be negligible. This way it is easy to understand how the structure imposed to the model allows the marginal effects to be non-trivial. Different from the basic model, the ratio of marginal probability effects of two distinct continuous covariates on the same outcome is not constant across individuals and the outcome distribution. Moreover, marginal probability effects may change their sign more than once when moving from the smallest to the largest outcome. Therefore, while the standard model precludes a flexible analysis of marginal probability effects by design, when turning our attention to the effects on the full distribution of outcomes this extension appears to be more appropriate.

For a continuous variable x_h in $\mathbf{x}$ but not in $\mathbf{z}$, the marginal partial effect is much simpler:

$$
\begin{aligned}
M_{jh}(\mathbf{x}) &= \frac{\partial \Pr(S=j|X=\mathbf{x}, Z=\mathbf{z})}{\partial x_h} = \qquad (1.4)\\
&= \left[\phi(\alpha_{j-1} - \mathbf{x}\boldsymbol{\beta} - \overline{\mathbf{x}}\gamma) - \phi(\alpha_{j} - \mathbf{x}\boldsymbol{\beta} - \overline{\mathbf{x}}\gamma)\right]\beta_h\cdot\Omega_{(z)}.
\end{aligned}
$$

Finally, for discrete variables in $\mathbf{z}$ the partial effect is easy to compute and similar to the baseline case:

$$
\Delta p_j(\mathbf{x},\mathbf{z}) = \Pr(S=j|X=\mathbf{x}+\Delta x_h, Z=\mathbf{z}+\Delta x_h) - \Pr(S=j|X=\mathbf{x}, Z=\mathbf{z}),
$$

while for discrete variables in $\mathbf{x}$ but not in $\mathbf{z}$ the partial effect is exactly the same as in the baseline setting.

1.3 Data

1.3.1 The Life Satisfaction Variable

The BHPS is a longitudinal panel survey of households in Great Britain. The first wave of data was collected in 1991,[6], originally including 5,500 households. Members of these households who were aged 16 years and over in 1991 have been interviewed every year, and their children included as respondents when older than 16, as well as any new member of the household. About 10,300 individuals are interviewed every year from 1996 to 2007 on the general question "How satisfied are you with life overall?".[7] They can choose based on an ordinal scale from 1 to 7, where 1 means 'not satisfied at all' and 7 'completely satisfied'. The dependent variable is therefore a 1 to 7 ordered response variable denoted as 'Satisfaction with Life Overall' and is meant to measure subjective well-being. By means of a single question it is possible to register individuals' self-reported level of happiness. The person surveyed makes a cognitive assessment on her own perceived quality of life, and we are driven by the belief that these data are significantly reliable for disclosing individuals' state. Studies on subjective well-being generally take two main perspectives referred to the concept they want to capture by means of the satisfaction variable: *hedonism* and *eudaimonia* (Kahneman et al., 2003). Hedonism can be expressed as the pursuit of satisfaction by self-gratification or pleasure, thus well-being is merely related to the material goods and the immediate enjoyment of such goods. Eudaimonia refers to the human desire for overall fulfillment - originally eudaimonia ($\varepsilon\upsilon\delta\alpha\iota\mu o\iota\nu\iota\alpha$, happiness etymologically) was a concept belonging to greek philosophy[8] which considered happiness as the final goal, the moral perfection of the

[6] The number of waves an individual is surveyed may change due to several reasons, such as death, immigration and attrition or because new individuals become part of the household.

[7] We drop all the non-full interviews. From Wave 7 (1997) there is oversampling of low income people for comparability with the European Community Household Panel. Moreover, many more observations have been sampled for Scotland and Wales. In order to maintain comparability with previous waves and random sampling, we keep only observations belonging to the original sample.

[8] Socrates was the first philosopher using this term; Aristoteles and Plato contributed to develop the concept in relationship with the moral and political disciplines.

human-being achieved by means of the *Virtus*, and for this reason material circumstances were conceived to be only corollary to pure happiness. By interpreting the meaning of eudaimonia for the present society, we might consider it as the multidimensional actualization of the *self* and a commitment to socially-shared goals. Despite the fact that both are considered separately as inputs into subjective well-being, for the purposes of this work we focus on the concept of 'eudaimonia', given the use of variables other than income in our analysis of well-being determinants.

1.3.2 Income, Relative Income and Deprivation

Our main interest is to assess the importance exerted by material circumstances on individual well-being. For this reason, such regressors play a crucial role in the analysis and deserve a special mention.

The variable income is meant to capture the consumption capacity of the person surveyed. It is intended as the compound of annual nominal household labour income and household non-labour income both deflated at the UK CPI[9] (basis year: 2005). We opt for household rather than individual income for the simple reason that life tenor depends on the familiar monetary wealth more than on the individual one.

Relative income, instead, is computed as the ratio between the real household income and the average income in the neighborhood.

In the following digression we will explain in which way relative income is thought of proxying a measure of social comparison and what is the definition of neighborhood used.

In line with the economic literature on subjective well-being, we assume that happiness responses give us a perception of individuals' preferences. In practice, we hypothesize that individuals make a cognitive assessment of their overall situation and express their self-measured level of satisfaction deriving from the utility function maximization.

Let us consider a function of the form:

[9] Source UK National Statistics (http://www.statistics.gov.uk/hub/index.html)

$$
\begin{aligned}
U_{nt} &= S_{nt}^{*}\left[\left(y_{jt}\right),\left(y_{jt}/y^{*}\right),\mathbf{x}_{nt}\right], &(1.5)\\
n &= \text{individual, } j = \text{household, } t = \text{time.}
\end{aligned}
$$

where U stands for utility, y_{jt} is real household income and y^* is a specific benchmark income, also called comparison income. Finally, $\mathbf{x}_{nt}$ is a vector of covariates- in our case, demographic and socioeconomic variables.

The term that includes relative income expresses social comparison. Since Duesenberry (1949), the relative income hypothesis- i.e. that people care about what their income is compared to other people in the same country more than their absolute one- has been used in many speculations on individual preferences and reciprocity. Nevertheless, it is only recently that the happiness economics literature focuses on the importance of material comparison for individual well-being. In particular, neighborhood more than country effects are thought of playing a role in these regards. Neighborhood effects are in general defined as "social interactions that influence the behavior or socioeconomic outcome of an individual" (Dietz, 2002) They include influences on individual behavior or outcomes due to the characteristics of an individual's neighbors and neighborhood, and spatial aspects of the neighborhood (the spatial relationship is defined with respect to location of residence). However, a measure of social distance may also be appropriate. Therefore, how choosing the reference (or comparison) group is of crucial importance for measuring social and economic interdependencies correctly. The main question here is whether the size of the neighborhood, as *a priori* determined by the researcher, influences the conclusions of the study. At present, there are no convincing answers to such a question. In our specific case the neighborhood delineation is driven by limitations of the data set. Specifically, we select reference groups based on sub-region and age-cohort, lacking of theoretically motivated definitions of neighborhood. If this presents an estimation bias is not known with certainty, given that no studies in the neighborhood effects literature exist which empirically test the effect of different neighborhood definitions. The common sense suggests that individuals are likely to compare with people they are in contact with in everyday life, and who share similar characteristics, e.g., are same-aged and live in the same area. As regards the geographical area, two options were available using the BHPS: either considering the so called ‘Primary Sampling Units’ (PSU’s) or UK sub-regions. The former contain, at minimum, 500 households and are stratified into an ordered listing by region and three socio-demographic variables. The latter refer to 18 sub-regions. Considering PSU’s defined neighborhoods would mean having very small groups in most of the cases, as well as too much variability in the size of the different groups. That is the reason why we opt for grouping by 18 sub-region, and 6 age-cohorts, singling out 108 neighborhoods. In this last case, in fact, we increase the size of each neighborhood and minimize its within-region variability. Furthermore, we assume within neighborhood effects only, i.e. that the neighborhood has no spillover characteristics. Thus, neighborhoods with identical characteristics but dissimilar neighboring neighborhoods are considered equivalent. In attempting to embed the educational dimension into the neighborhood choice we encountered a problem of collinearity with the income variable, which is present in the estimation as well. Finally, we imagine that income comparisons are not symmetric, affecting the poor more than the rich (Ferrer-i-Carbonell, 2005). For this reason a deprivation relative to mean income measure is introduced, leading the empirical function to be conceived as follows[10]

$$S^*_{nt} = \xi_n + \ln(y_{jt})\beta_1 + \ln(y_{jt}/y^*)\beta_2 + D \cdot \ln(y_{jt}/y^*)\beta_3 + \underset{1\times k-3}{\mathbf{x}_{nt}}\boldsymbol{\beta}_k + \epsilon_{nt} \tag{1.6}$$

where

$$D = \begin{cases} 1 \text{ if } y_{jt} \leq y^* \\ 0 \text{ otherwise} \end{cases},$$

and S^*_{nt} is the conditional expected value of individual well-being. If $\beta_2 > 0$, an increase

[10] Notice that equation (1.6) represents the latent random utility model, as in equation (1.2).

in the comparison income reduces the well-being of those with an income above the mean. An increase in the reference income produces a worsening in well-being for individuals with a given income below the mean if $\beta_2 + \beta_3 > 0$. Finally, if $\beta_3 > 0$ the comparison income has a greater effect on the poor.[11] Gravelle and Sutton (2009) introduce the same measure for studying the relationship between perceived health and income in the UK. We find its design appropriate to our purpose as well, because we want to test for asymmetries in the impact that relative income might have on the relatively poor and the relatively rich in the comparison group.

1.3.3 Control Variables

A large number of control variables is included among the regressors for rendering the analysis more robust.

First of all, we think that health status could strongly drive the happiness response. In order to capture the impact of the health status on life satisfaction, we first make use of a self-reported measure of subjective health. Data were collected by registering answers to the question "How would you define your health status over the last 12 months" on a 1-5 scale (from excellent to very poor). We dichotomize the variable by assigning it value 1 if the original were 1 and 2, and value 0 otherwise, by relying upon the median point to group responses into good or bad health status. Criticism may arise on the endogeneity of such variable: an individual saying she is happy can subjectively consider herself in a good health status and the other way around. This is why we repeat the analysis by replacing this measure for health with the variable 'Limits in Activities of Daily Life (ADL)'. This is a dichotomous variable that takes on value 1 if individuals say that a list of health problems limit their daily activities (doing the housework, climbing the stairs, getting dressed, walking more than 10 minutes, limits in type or amount of work) and 0 otherwise. We argue that in this way it is possible to synthesize individuals'

[11] A person whose income is 20,000 GBP, and confronts herself with a reference income of 30,000 GBP, experiences the same relative deprivation of an individual having 90,000 GBP per year and a comparison income of 100,000 GBP.

health objectively, by taking into account the possible consequences of several factors, such as illness, obesity and injuries. Our aim is to check that the results obtained under a subjective measure of health status are not too dissimilar from those obtained by including a more objective proxy, which rules out possible psychological interferences.

Marital status is indicated by the binary variable 'Married'. We include both legally married and living-as-a-couple individuals, given that we are interested in the effect of sharing everyday life with someone rather than the importance of the mere relationship type. 'Children' is a dummy indicating the presence of own children in the household, while 'Employed' is a binary variable that indicates being in-paid employed.

Age is calculated from the date of birth, and is included in the regression squared and cubed, in order to control for potential non-linearities in the relation with happiness.

Finally, we include gender, ethnicity, year and geographical dummies. In this case, compared to what we have done for computing the relative income, we group geographical regions into macro-areas: Southern England, Northern England, London, Scotland, and Wales.

Although the BHPS offers a good range of educational variables, only one suited our purposes, specifically a qualitative variable on educational attainment. Nevertheless, even when properly modified, we faced the problem of collinearity between this variable and the income one, which makes good sense if we consider income as a proxy for education. Therefore, we could not explicitly include any educational variable.

1.3.4 Potential Sources of Heteroskedasticity

A last note is due on the choice of the variables to be included in the set of potential heteroskedasticity sources, i.e. the vector $\mathbf{z}$ in $\sigma_\epsilon^2 = \exp(\mathbf{z}_{nt}\vartheta)^2$. We mentioned that the vector $\mathbf{z}$ can contain either some or all regressors, or variables which are not included among the explanatories, or a mixture of both. In our case, we have selected income, sex, age and ethnicity to appear in the variance structure, this leading $\mathbf{z}$ to be a subset of $\mathbf{x}$.

Income has been chosen for taking into consideration the possibility that an increase in income has a greater impact for poor people than for rich people. Therefore, given the high correlation between poverty and low self-reported well-being, we are driven to think that the variation in income might cause the perceived satisfaction to vary more for the poor than for the rich. Loosely speaking, a very poor person who rated herself as completely unsatisfied and experiences ameliorations in her income might change her response by one unit, for example. The same variation might not cause a similar reaction for a rich individual who rated herself as satisfied 'six' on a one-to-seven scale, simply because more income does not matter for being one score happier. A similar behavior, which is likely to bias our results, is not controllable otherwise, nor the inclusion of unobservable individual effects can assure that we properly account for it . Heterogeneity can arise due to several factors. For example, it may be the by-product of different levels of perception about a choice: certainty about if and how much satisfied one is with her life might depend on mental sophistication,[12] cultural heirloom, personal ambition. In fact, age, gender and ethnicity dummies are added for capturing some more variation in choices, even though we have included them also in the main regression. Again, the point is to relate heterogeneity in choices, therefore potential error heteroskedasticity, with its plausible causes, and we are persuaded that those variables are indeed good factors for explaining human complexity and heterogeneity.

1.3.5 Descriptive Statistics

Figure 1 displays the percentage of the responses to the subjective well-being question. In accordance with the literature exploring individual well-being in western countries, about 75% of the people surveyed assert to be very satisfied.[13]

Figure 1 about here

The transition matrix reported in Table 1 gives us a rather clear perception of how responses change over time. Probabilities located on the main diagonal are quite high,

[12] For instance, men and women have different sensibility and ambitions, as it is well-known.
[13] Percentage computed from the sum of densities relative to responses between 5 and 7

meaning that choosing the same response is frequent, especially for 'very satisfied' people; higher volatility is observed for responses from 1 to 3. A reasonable interpretation for this is that individuals who consider themselves very unsatisfied could find an improvement in their lives more significant than already 'happy' individuals, as already discussed in the previous subsection.

Table 1 about here

As a preliminary clue on the nature of the relationship between life satisfaction and real income, let us notice that, according to Figure 2, real average household income has significantly increased while life satisfaction has been on average fairly constant. Not surprisingly, what we find in our data is adherent to what other studies on western economies have already found (e.g., *inter alia*, Easterlin, 1974, 1995, 2001; Clark and Oswald, 1994; Hagerty and Veenhoven, 2003; Ferrer-i-Carbonell, 2005; Clark, Frijters and Shields, 2008).

Figure 2 about here

Table 2 shows some descriptive statistics. Given that the regressors are mainly binary variables, we have computed the mean level of life satisfaction and how it varies when individuals surveyed are women or men, married or not, in good or bad health status, employed or unemployed, have babies or not, have an income above/below the average in their neighborhood or in the whole sample. Individuals with a good perceived health have an average satisfaction 0.61 units higher than the average of the whole sample; such a difference in the mean may be quite important. Who lives as a couple has a higher level of average satisfaction, while women and men in the sample have almost the same average level of life satisfaction. Moreover, people older than the average are happier than younger respondents.

Average life satisfaction is higher for individuals with a household income greater than the average, both in the reference group and in the whole sample, and lower for those lagging behind the others. At this first attempt, we are inclined to think that our guess on relative concerns is correct and that other factors rather than income itself are at work to determine increases in happiness.

Table 2 about here

1.4 Estimation Results

Tables 3 and 4 display the HPPOP and PPOP estimates, both with and without individual effects, using respectively a subjective health measure (Health) and a more objective one (ADL, limits in Activities of Daily Life). The first question we address is whether one of the models presented uses the information inherent in the data optimally. For this purpose, we perform information criteria comparisons between each model: a smaller value indicates a better fit while penalizing for the escalation of parameters. Akaike, Hannan-Quinn and Schwarz Bayesian criteria are reported at the bottom of both Table 4 and Table 5. It can be observed that all these criteria suggest the HPPOP model with individual-specific effects should be favored to all the others. That is why such a model is considered as the benchmark. For completeness reasons, though, the other models estimates are included in our comments.

All our results show that income and relative income are both significant, but they exert an opposite effect on happiness: specifically, absolute income is in negative relationship with happiness, while relative income has a positive link with it. Therefore, the total effect of absolute income, obtained by summing the coefficients of absolute and relative income, is almost null. As to the deprivation measure, when subjective health is considered, it is significant and positive in sign only in the benchmark model, i.e. in the HPPOP with unobserved individual effects. In the objective health estimation, it is instead always significant and positive. Such results mirror what conjectured: when the temporal dimension is added, absolute income matters very little for happiness, because of adaptation and income shock absorption in the long run. The positive coefficient attached to the relative income variable, instead, signals that an increase in the comparison income reduces the well-being of those with a household income above the mean. Furthermore, the sum between the relative income and the deprivation measure coefficients is positive, meaning that an increase in the reference income produces a worsening in well-being for individuals with a given income below the mean. Finally, the deprivation coefficient is positive, thus the comparison income has a greater effect on the poor than on the rich, relatively to the neighborhood they belong to. We argue that this explanation could constitute a solution to the Easterlin Paradox in that the impact of absolute income is compensated from the one of reference income, leading happiness to depend more on material social comparison than on household wealth itself. While this idea is not new in the subjective well-being literature, yet our methodological analysis renders such findings more reliable.

Not surprisingly, the most relevant variables for subjective well-being are health and marital status. As to the role played by health status, a good perceived health positively and substantially affects happiness. Intuitively, limits in ADL have a negative effect on life satisfaction. Marital status is found to exert a positive effect on happiness as well, while having children has a negative effect on the whole sample of individuals. Finally, employment status is in positive relation with happiness, but shows a smaller impact than health and marital status.

The variable age is included squared and cubed in order to determine the nature of its relation with the dependent variable and to allow for potential non-linear patterns. Our estimates suggest that age can be related to life satisfaction through a convex decreasing relationship. It is interesting to mention that several cross-sectional or random-effects analyses highlight a U-shaped pattern (e.g., Oswald, 1997; Blanchflower and Oswald, 2004a; Lelkes, 2006). However, the marginal effect of an additional year in the age distribution is typically small.

Finally, the ethnicity dummies are significant only for white and black individuals, essentially because they are the most numerous groups. The magnitude of the impact on the response probabilities is approximately the same.

Notice how the individual-specific time-invariant effects standard errors are smaller when the variance is structured as described in the previous sections. This means that, although the coefficients relative to the **z** variables are meaningless *per se*, still we are able to capture some more error variation and, perhaps, to correct upward/downward biases. Besides, it is only in the HPPOP with individual effects that the deprivation measure shows significance in the main model specification (with subjective health).

In order to better understand the magnitude of the effect that such variables exert on life satisfaction, as well as to know how the impact changes across categories, we now turn our attention to average marginal probability effects of the income variables on happiness (Tables 5 and 6).

First of all, let us focus on the absolute household income variable in Table 5.[14] The interpretation of, for example, first column $MPE_5 = 0.0414$ is that a one-percent increase in log-income raises the probability of life satisfaction = 6 by approximately 0.0414 percentage points. A quite striking result of our benchmark model (column 1) is that of a negative marginal partial effect for both low happiness responses and the highest one, meaning that an increase in absolute income actually reduces the probability of being completely satisfied as well as of being generally dissatisfied. Looking at the magnitude of the effects, we can observe that the negative impact on the individuals who rated themselves as the happiest is about -4%, while for the low categories the percentage is on average -0.45%. This would signal that absolute income is not the key variable driving happiness. Only the individuals who perceive themselves as moderately happy (4 and 5 responses) show a positive income impact. Performing the same estimation with no fixed effects (column 2) simply leads to an underestimation of the magnitude for the dissatisfied individuals and an overestimation of the impact on the highest categories.

The same behavior cannot be inferred from the PPOP (column 3 and 4), where the effects' sign is allowed to change only once by design. Indeed, the somewhat perverse result in this model is that the high responses are associated to a negative sign and the others to a positive one.

[14] Marginal partial effects computed as in (1.3)

On the contrary, the results are unambiguous regarding relative income:[15] a positive variation in this variable due to either an increase in absolute income, or a decrease in reference income, or both, increases the probability of rating oneself very happy or completely happy of about 0.9% and decreases the probability of being dissatisfied or moderately happy of approximately 2% on average. Such a result is confirmed for all the models, where accounting for fixed effects allows to avoid, again, overestimation.

Finally, variations in the deprivation variable follow, intuitively, those in relative income, and have to be interpreted as "getting less deprived" increases the probability of being very/completely happy , while decreasing the one of being less happy. Fixed effects are crucial to have significant results, for both HPPOP and PPOP.

The results displayed in Table 6 mimic those pertaining Table 5 just commented, confirming that using a subjective measure of health instead of an objective one does not spoil the basic variable relationships.

1.5 Conclusions

In the last 30 years research in economics has experienced a booming in the exciting field of happiness and well-being studies. Many are the unsolved questions about what determines life satisfaction, and economists started focusing on the role of money in people's happiness. The well-known Easterlin Paradox, the economics of happiness milestone, finds that increasing trends in income are associated with flat average levels of life satisfaction in western countries. In a first instance this signals that in developed societies money does not necessarily bring the contentment we might think, thus other factors might be at work. When accounting for other determinants such as a good health and family status, cultural and civic trust as well as age and sex, the effect of absolute income may be even negative. At the light of this evidence, research has recently moved its interest towards the effect exerted by relative rather than absolute income. Given the phenomenon of adaptation, individuals are thought of being only temporarily influ-

[15] Marginal partial effects computed for HPPOP as in (1.4).

enced by variations in their income, even when highly positive. This might explain why, despite the significative increase in income, people rate themselves as being as happy as always. The relative position in the social ladder, proxied by relative income, could explain the existence of frustrated achievement or constant self-reported levels of happiness corresponding to higher incomes.

Our work investigates the role of relative income for satisfaction with life making use of frontier econometric methods. Indeed, our primary concern is to perform an analysis tailored on the data at hand, as robust as possible, and taking into due consideration the possible problems arising from subjective micro-data on personal well-being. Furthermore, we try to compute the reference income embedding two distance dimensions between individuals, namely age-cohort and geographical sub-region. Whether the happiness paradox can be explained by the relationship between relative income and satisfaction is still an open debate. Nevertheless, we argue that a further step is moved towards the comprehension of people's psychology and their perception of what money can buy, based on the conviction that the strategy used is very appropriate for the treatment of such data. With this purpose in mind, we implement an heteroskedastic pooled panel ordered probit with 'quasi-fixed' effects, extending the method *à la* Mundlak (1978) to a non-linear setting where the homoskedasticity assumption is relaxed.

Our analysis is based on the assumption that self-reported life satisfaction is a valid measure for well-being, and that current happiness predicts future behavior. In accordance to a number of studies pursued for other countries, we find that health, employment and marital status are very important predictors of well-being. On the one hand, happiness appears to be decreasing in absolute income, even for people that rate themselves as completely satisfied with their life. On the other hand, relative income, i.e., the ratio between household income and average household income in the neighborhood, seems to impact positively on the probability of the self-rated happiest categories. The relative size of their effects is positive, this meaning that the positive impact of an increase in income with respect to the reference one overcomes the effect exerted by absolute income. Furthermore, the effect is asymmetric affecting the poor more than the rich. Our results lead to conclude that relative income should be accounted for when exploring what actually affects people's behavior and their perception of life satisfaction. This could represent a key for the solution of the happiness paradox.

Bibliography

[1] Alesina, A., R. Di Tella, and R. MacCulloch (2004). "Inequality and happiness: Are Europeans and Americans different?" *Journal of Public Economics 88*(9-10), 2009-2042.

[2] Alvarez, R. and J. Brehm (1997). "Are Americans ambivalent towards racial policies," *American Journal of Political Science 41*, 345-374.

[3] Alvarez, R. and J. Brehm (1998). "Speaking in two voices: American equivocation about the internal revenue service," *American Journal of Political Science 42*, 418-452.

[4] Alvarez, R. and J. Brehm (2002). *Hard choices, easy answers.* Princeton, NJ: Princeton university Press.

[5] Blanchflower, D. and A. Oswald (2001). "Well-being over time in Britain and the USA," The Warwick Economics Research Paper Series (TWERPS) 616, University of Warwick, Department of Economics.

[6] Blanchflower, D. and A. Oswald (2004a). "Money, sex and happiness: An empirical study," *Scandinavian Journal of Economics 106*(3), 393–415.

[7] Blanchflower, D. and A. Oswald (2004b). "Well-being over time in Britain and the USA," *Journal of Public Economics 88*, 1359–1386.

[8] Boes, S., K. Staub, and R. Winkelmann (2010). "Relative status and satisfaction," *Economics Letters 109*, 168-170.

[9] Boes, S. and R. Winkelmann (2006). "Ordered response models," *Advances in Statistical Analysis 90*(1), 167-181.

[10] Busch, M. and R. Reinhardt (1999). "Industrial location and protection: The political and economic geography of U.S. nontariff barriers," *American Journal of Political Science 43*, 1028-1050.

[11] Clark, A., P. Frijters, and M. Shields (2008). "Relative income, happiness, and utility: An explanation for the Easterlin paradox and other puzzles," *Journal of Economic Literature 46*(1), 95-144.

[12] Clark, A. and A. Oswald (1994). "Unhappiness and unemployment," *Economic Journal 104*(424), 648-59.

[13] Dietz, R. D. (2002). "The estimation of neighborhood effects in the social sciences: An interdisciplinary approach," *Social Science Research 31*, 539-575.

[14] Dolan, P., T. Peasgood. and M. White (2008). "Do we really know what makes us happy? A review of the economic literature on the factors associated with subjective well-being," *Journal of Economic Psychology 21*(1), 94-122.

[15] Duesenberry, J. S. (1949). *Income, savings and the theory of consumer behavior.* Cambridge, MA: Harvard University Press.

[16] Dynan, K. and E. Ravina (2007). "Increasing income inequality, external habits and self-reported happiness," *American Economic Review 97*(2), 226-231.

[17] Easterlin, R. A. (1974). *Does economic growth improve the human lot? Some empirical evidence.* In P. A. David and M.W. Reder (Eds.), Nations and Households in Economic Growth: Essays in Honor of Moses Abramowitz, New York and London. Academic Press.

[18] Easterlin, R. A. (1995). "Will raising the incomes of all increase the happiness of all?" *Journal of Economic Behavior and Organization 27*, 35-48.

[19] Easterlin, R. A. (2001). "Income and happiness: Towards a unified theory," *Economic Journal 101*, 465-484.

[20] Easterlin, R. A. (2003). Explaining happiness. *Proceedings of the National Academy of Sciences 100*, 11176-11183.

[21] Ferrer-i Carbonell, A. (2005). "Income and well-being: An empirical analysis of the comparison income effect," *Journal of Public Economics 89*(5-6), 997-1019.

[22] Ferrer-i Carbonell, A. and P. Frijters (2004). "How important is methodology for the estimates of the determinants of happiness?" *Economic Journal 114*(497), 641-659.

[23] Firebaugh, G. and L. Tach (2000). "Income and happiness in the United States," Technical report, Harvard University.

[24] Frey, B. and A. Stutzer (1999). "Measuring preferences by subjective well-being," *Journal of Institutional and Theoretical Economics (JITE) 155*(4), 755–778.

[25] Frey, B. and A. Stutzer (2000). "Happiness, economy and institutions," *Economic Journal 110*(466), 918–38.

[26] Frijters, P., J. Haisken-DeNew and M. Shields (2004a). "Investigating the patterns and determinants of life satisfaction in Germany following reunification," *Journal of Human Resources 39*(3).

[27] Frijters, P., J. Haisken-DeNew and M. Shields (2004b). "Money does matter! Evidence from increasing real income and life satisfaction in East Germany following reunification," *American Economic Review 94*(3), 730-740.

[28] Gabel, M. (1998). "Economic integration and mass politics: Market liberalization and public attitudes in the European Union," *American Journal of Political Science 42*, 936-953.

[29] Graham, C. (2008). *The economics of happiness.* Houndmills, England: Macmillan. The New Palgrave Dictionary of Economics.

[30] Gravelle, H. and R. Sutton (2009). "Income, relative income, and self-reported health in Britain 1979-2000," *Health Economics 18*, 125-145.

[31] Hagerty, M. and R. Veenhoven (2003, October). "Wealth and happiness revisited - growing national income does go with greater happiness," *Social Indicators Research 64*(1), 1-27.

[32] Headey, B., R. Muffels and M. Wooden (2008). "Money does not buy happiness: Or does it? A reassessment based on the combined effects of wealth, income and consumption," *Social Indicators Research 87*(1), 65-82.

[33] Hsiao, C. (1986). *Analysis of panel data.* Cambridge, USA: Cambridge University Press. Econometric Society Monograph.

[34] Kahneman, D., E. Diener, and N. Schwarz (2003, January). *Well-Being: Foundations of hedonic psychology.* Russell Sage Foundation Publications.

[35] Keele, L. and D. Park (2006). "Ambivalent about ambivalence: A re-examination of heteroskedastic probit models," Technical report.

[36] Krutz, G. (2005). "Issues and institutions: Winnowing in the US congress," *American Journal of Political Science 49*, 436-447.

[37] Lee, T. (2002). *Mobilizing public opinion: Black insurgency and racial attitudes in the civil rights era.* Chicago: University of Chicago Press.

[38] Lelkes, O. (2006). "Knowing what is good for you: Empirical analysis of personal preferences and the 'objective good'," *The Journal of Socio-Economics 35*(2), 285-307.

[39] Luttmer, E. (2005). "Neighbors as negatives: Relative earnings and well-being," *Quarterly Journal of Economics 120*, 963-1002.

[40] Maddala, G. (1983). *Limited dependent and qualitative variables in econometrics.* Cambridge: Cambridge University Press.

[41] McBride, M. (1978). “Relative-income effects on subjective well-being in the cross-section,” *Journal of Economic Behaviour and Organization 45*, 251-278.

[42] Mundlak, Y. (1978). “On the pooling of time series and cross section data,” *Econometrica 46*, 69-85.

[43] Oswald, A. (1997). “Happiness and economic performance,” *Economic Journal 107*, 1815-1831.

[44] Sacks, D., B. Stevenson and J. Wolfers (2010). “Subjective well-being, income, economic development and growth,” NBER Working Papers 16441, National Bureau of Economic Research, Inc.

[45] Stevenson, B. and J. Wolfers (2008). “Economic growth and subjective well-being: Reassessing the Easterlin paradox,” Brookings Papers on Economic Activity 39(1), 1-102.

[46] Williams, R. (2010). “Fitting heterogeneous choice models with oglm,” *Stata Journal 10*(4), 540-567.

[47] Williams, R. (2011). “Oglm: Stata module to estimate ordinal generalized linear models,” Mimeo.

[48] Winkelmann, L. and R. Winkelmann (1998). “Why are the unemployed so unhappy? Evidence from panel data,” *Economica 65*(257), 1-15.

A Dataset Features and Statistical Package

Quoting the official BHPS web site "The British Household Panel Survey began in 1991 and is a multi-purpose study whose unique value resides in the fact that:

- it follows the same representative sample of individuals – the panel – over a period of years;
- it is household-based, interviewing every adult member of sampled households;
- it contains sufficient cases for meaningful analysis of certain groups such as the elderly or lone parent families.

The wave 1 panel consists of some 5,500 households and 10,300 individuals drawn from 250 areas of Great Britain". From Wave 7 (1997), there is oversampling of low income people for comparability with ECPH. "Moreover, many more observations have been sampled for Scotland and Wales. Additional samples of 1,500 households in each of Scotland and Wales were added to the main sample in Wave 9 (1999), and in 2001 a sample of 2,000 households was added in Northern Ireland, making the panel suitable for UK-wide research".

Data in each wave are organized in different macro-groups: INDSAMP includes all sampled individuals (either respondents or not), INDALL is an individual level record for all members of the household, corresponding to the household grid, INDRESP includes responding individuals only. The same applies to household-specific data, collected into HHSSAMP, HHSAMP and HHRESP. Hence, when extracting the individual interview outcome (IVFIO) from INDSAMP/HHSAMP, we are taking more observations than those that we have in INDRESP/HHRESP. They are dropped when dropping according to IVFIO (we drop all the observations where the interview outcome was not 1, i.e. all the non-full interviews). Also, in order to maintain comparability with previous waves and random sampling, we keep only observations belonging to the original sample (MEMORIG=1 for INDRESP and HHORIG=1 for HHRESP), disregarding the data added from 1997, 1999 and 2001 mentioned before.

Here follows a list of BHPS codes for the raw variables used in our analysis, in alphabetical order:

Raw Data			
age	age from birth	biographical	continuos
fihhyl	annual household labor income	derived	continuous
fihhynl	annual household non-labor income	derived	continuous
hgemp	In paid employment - household grid	self-reported	binary
hllte	health no hindrance daily activities	self-reported	binary
hlstat	health over last 12 months	self-reported	1-5 ordered
lfsato	satisfaction with life overall	self-reported	1-7 ordered
mastat	marital status	biographical	5 different stati
nchild	number of own children in household	biographical	continuous
race	ethnicity	biographic	5 different races
region	region / metropolitan area	biographical	18 UK sub-regions
sex	gender	biographical	

By means of STATA, the PPOP model has been estimated using the standard command `oprobit`. For the HPPOP model, instead, we have made use of a STATA module by Williams (2010, 2011), known as `oglm`.

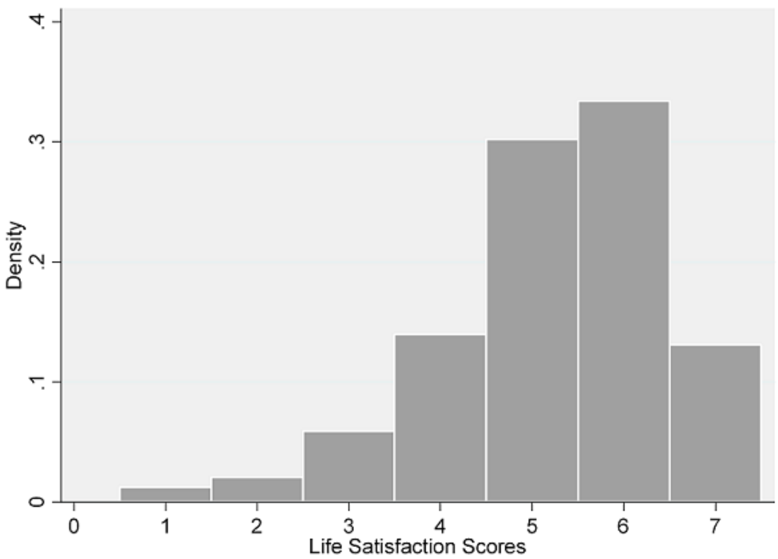

Figure 1: Density of Life Satisfaction Responses

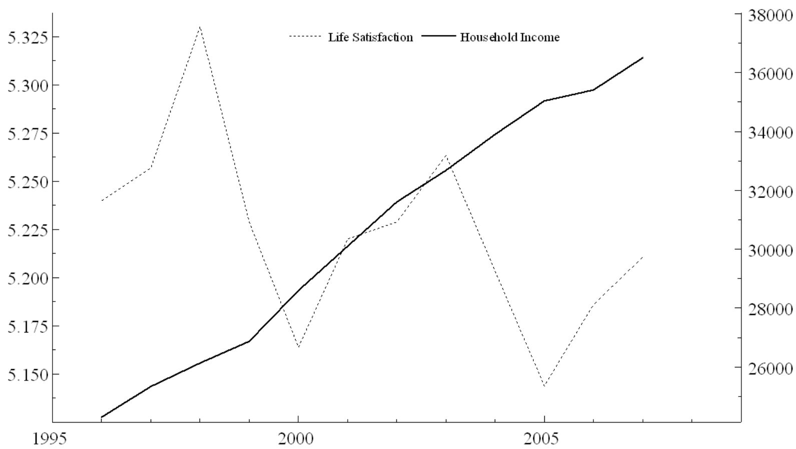

Figure 2: Average Real Household Income and Life Satisfaction Series

Table 1. Transition Matrix for Life Satisfaction, 1996-2007

		Life Satisfaction in t							
		1	**2**	**3**	**4**	**5**	**6**	**7**	Total
Life Satisfaction in $t-1$	**1**	30.16	14.43	15.52	14.86	12.68	6.56	5.79	100
	2	8.40	18.86	25.79	20.97	15.65	7.50	2.82	100
	3	2.81	9.20	25.38	29.39	21.89	9.11	2.21	100
	4	1.68	3.17	12.22	32.21	34.12	13.40	3.18	100
	5	0.36	1.07	4.58	15.64	46.36	28.19	3.79	100
	6	0.24	0.44	1.57	5.89	26.15	54.93	10.78	100
	7	0.68	0.43	1.27	3.83	9.81	30.28	53.70	100
	Total	1.2	2.05	5.94	13.96	30.57	33.66	12.61	100

Table 2. BHPS Descriptive Statistics, 1996-2007

Variable	Obs	Mean	St.Dev.
Life Satisfaction	91494	5.22	1.25
Household Income	93145	30345.45	22830.35
Age	92870	45.07	18.51
Life Satisfaction if income > average in the neighborhood	38259	5.36	1.12
Life Satisfaction if income < average in the neighborhood	52675	5.13	1.33
Life Satisfaction if income > average in the sample	37462	5.29	1.10
Life Satisfaction if income < average in the sample	54032	5.18	1.35
Life Satisfaction if younger than average	48722	5.17	1.18
Life Satisfaction if older than average	40923	5.30	1.33
Life Satisfaction if good health status	58284	5.45	1.08
Life Satisfaction if bad health status	33163	4.84	1.43
Life Satisfaction if employed	57654	5.24	1.12
Life Satisfaction if unemployed	33840	5.20	1.46
Life Satisfaction if woman	49450	5.22	1.29
Life Satisfaction if man	42040	5.23	1.21
Life Satisfaction if married or living-as-couple	60601	5.31	1.19
Life Satisfaction if divorced, widowed or single	30654	5.05	1.36
Life Satisfaction if have children	26167	5.12	1.20
Life Satisfaction if do not have children	65367	5.26	1.27

Notation in Tables 3, 4, 5 and 6:

- HPPOP= Heteroskedastic Pooled Panel Ordered Probit; PPOP= Pooled Panel Ordered Probit.

- Dependent Variable LIFE SATISFACTION naturally coded; score 1=very unsatisfied, score 7=completely satisfied.

- 'hhincome' refers to household labor and non-labor income.

- 'comparison income' is determined by age-cohort (16-25, 26-35, 36-45, 46-55, 56-65, 66-75, 75>), and sub-region (Inner London, Outer London, Rest of South East, South West, East Anglia, East Midlands,West Midlands Conurbation,Rest of West Midlands, Greater Manchester, Merseyside, Rest of North West, South Yorkshire, West Yorkshire, Rest of Yorks and Humberside, Tyne and Wear, Rest of North, Wales, Scotland).

- 'Deprivation' is $D{\cdot}\ln$(hhincome/comparison income), where

$$D = \begin{cases} 1 \text{ if hhincome} \leq \text{comp. income} \\ 0 \text{ otherwise} \end{cases}$$

- *Sigma=exp($\mathbf{z}_{nt}\boldsymbol{\vartheta}$).

- AIC= Akaike Information Criterion; HQ= Hannan-Quinn Information Criterion; SC= Schwarz Information Criterion.

Table 3. Estimation Results (Subjective Health)

Life Satisfaction in the UK,1996-2007

Individual Effects Total Std. Deviation	YES 0.3674	NO	YES 1.0124	NO
ln(hhincome)	-0.0275***	-0.0649***	-0.0681***	-0.190***
	(0.00821)	(0.0104)	(0.0226)	(0.0282)
ln(hhincome/comparison income)	0.0157*	0.0916***	0.0475*	0.278***
	(0.00949)	(0.0131)	(0.0265)	(0.0342)
Deprivation	0.0207**	-0.00563	0.0416**	-0.0404
	(0.00806)	(0.00946)	(0.0208)	(0.0267)
Health	0.0973***	0.222***	0.274***	0.651***
	(0.00721)	(0.0152)	(0.00933)	(0.0128)
Married	0.0746***	0.114***	0.220***	0.337***
	(0.00776)	(0.00930)	(0.0177)	(0.0171)
Children	-0.00204	-0.0128**	-0.00793	-0.0400**
	(0.00518)	(0.00545)	(0.0150)	(0.0160)
Woman	0.0258***	0.0205***	0.0667***	0.0529***
	(0.00534)	(0.00512)	(0.0146)	(0.0146)
Employed	0.0176***	0.0358***	0.0440***	0.103***
	(0.00583)	(0.00618)	(0.0164)	(0.0164)
Age	-0.0389***	-0.0393***	-0.112***	-0.122***
	(0.00438)	(0.00377)	(0.0103)	(0.00774)
Age×Age/100	0.0806***	0.0731***	0.230***	0.231***
	(0.00887)	(0.00762)	(0.0204)	(0.0166)
Age×Age×Age/1000	-0.00544***	-0.00378***	-0.0155***	-0.0124***
	(0.000595)	(0.000460)	(0.00136)	(0.00108)
Ethnicity: White	0.0265***	0.0237***	0.0735***	0.0635***
	(0.00694)	(0.00666)	(0.0192)	(0.0188)
Ethnicity: Black	0.0322**	0.0297**	0.0999***	0.0865**
	(0.0129)	(0.0127)	(0.0365)	(0.0364)
Ethnicity: Asian	-0.00381	-0.00785	-0.000625	-0.0141
	(0.0190)	(0.0185)	(0.0528)	(0.0524)
Ethnicity: Chinese	-0.0393	-0.0346	-0.0899	-0.0707
	(0.0775)	(0.0764)	(0.218)	(0.213)
Cut Point 1	-2.295***	-1.837***	-6.616***	-5.359***
	(0.210)	(0.153)	(0.438)	(0.291)
Cut Point 2	-2.127***	-1.674***	-6.167***	-4.922***
	(0.203)	(0.145)	(0.437)	(0.290)
Cut Point 3	-1.921***	-1.477***	-5.602***	-4.372***
	(0.194)	(0.136)	(0.436)	(0.289)
Cut Point 4	-1.686***	-1.251***	-4.941***	-3.728***
	(0.185)	(0.127)	(0.436)	(0.289)
Cut Point 5	-1.372***	-0.948***	-4.044***	-2.851***
	(0.174)	(0.116)	(0.436)	(0.289)
Cut Point 6	-0.977***	-0.563***	-2.934***	-1.757***
	(0.163)	(0.106)	(0.436)	(0.290)
ln(sigma*)	ln(hhincome) -0.123*** (0.00598)	ln(hhincome) -0.125*** (0.00591)	-	-
	Woman 0.0649*** (0.00999)	Woman 0.0674*** (0.00980)		
	Age 0.00346*** (0.000278)	Age 0.00367*** (0.000273)		
	White 0.0587*** (0.00905)	White 0.0426*** (0.00883)		
	Black -0.0101 (0.0274)	Black -0.000474 (0.0266)		
	Asian 0.0559 (0.0436)	Asian 0.0470 (0.0423)		
	Chinese -0.0469 (0.167)	Chinese -0.0574 (0.168)		
Year Dummies	YES	YES	YES	YES
Geographical Dummies	YES	YES	YES	YES
Observations	91068	91068	91068	91068
AIC: $[-\frac{2}{N}\cdot loglik + 2\frac{k}{N}]$	2.9644	2.9984	2.9900	3.7448
HQ: $[-\frac{2}{N}\cdot loglik + 2\frac{k}{N}\cdot \ln(\ln(N))]$	2.9662	2.9998	2.9914	3.2658
SC: $[-\frac{2}{N}\cdot loglik + \frac{k}{N}\cdot \ln(N)]$	2.9704	3.0028	2.9950	7.1357

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 4. Estimation Results (Objective Health)				
Life Satisfaction in the UK,1996-2007				
Individual Effects	YES	NO	YES	NO
Total Std. Deviation	0.3501		1.0716	
ln(hhincome)	-0.0294***	-0.0418***	-0.0840***	-0.131***
	(0.00767)	(0.00979)	(0.0219)	(0.0289)
ln(hhincome/comparison income)	0.0174**	0.0819***	0.0600**	0.274***
	(0.00876)	(0.0126)	(0.0256)	(0.0349)
Deprivation	0.0191***	-0.0150	0.0363*	-0.0861***
	(0.00741)	(0.00925)	(0.0201)	(0.0272)
ADL	-0.0474***	-0.145***	-0.133***	-0.427***
	(0.00731)	(0.0130)	(0.0195)	(0.0257)
Married	0.0672***	0.106***	0.212***	0.331***
	(0.00710)	(0.00893)	(0.0171)	(0.0177)
Children	0.00159	-0.00640	0.00415	-0.0202
	(0.00475)	(0.00527)	(0.0145)	(0.0164)
Woman	0.0162***	0.0149***	0.0423***	0.0400***
	(0.00498)	(0.00490)	(0.0151)	(0.0151)
Employed	0.0490***	0.0568***	0.151***	0.180***
	(0.00663)	(0.00690)	(0.0172)	(0.0171)
Age	-0.0447***	-0.0397***	-0.138***	-0.132***
	(0.00446)	(0.00377)	(0.0101)	(0.00798)
Age×Age/100	0.0806***	0.0733***	0.246***	0.249***
	(0.00853)	(0.00756)	(0.0199)	(0.0171)
Age×Age×Age/1000	-0.00544***	-0.00380***	-0.0166***	-0.0135***
	(0.000573)	(0.000454)	(0.00133)	(0.00111)
Ethnicity: White	0.0235***	0.0232***	0.0703***	0.0648***
	(0.00642)	(0.00629)	(0.0188)	(0.0186)
Ethnicity: Black	0.0258**	0.0251**	0.0820**	0.0728**
	(0.0120)	(0.0118)	(0.0362)	(0.0360)
Ethnicity: Asian	0.00421	0.00384	0.0291	0.0207
	(0.0176)	(0.0174)	(0.0523)	(0.0522)
Ethnicity: Chinese	-0.0776	-0.0611	-0.168	-0.130
	(0.0769)	(0.0739)	(0.222)	(0.218)
Cut Point 1	-1.899***	-1.701***	-5.801***	-5.288***
	(0.193)	(0.148)	(0.452)	(0.299)
Cut Point 2	-1.748***	-1.552***	-5.380***	-4.870***
	(0.187)	(0.141)	(0.452)	(0.299)
Cut Point 3	-1.565***	-1.374***	-4.850***	-4.346***
	(0.179)	(0.133)	(0.451)	(0.298)
Cut Point 4	-1.356***	-1.169***	-4.231***	-3.732***
	(0.172)	(0.124)	(0.451)	(0.298)
Cut Point 5	-1.075***	-0.892***	-3.381***	-2.889***
	(0.163)	(0.114)	(0.451)	(0.298)
Cut Point 6	-0.713***	-0.536***	-2.308***	-1.822***
	(0.155)	(0.104)	(0.451)	(0.298)
ln(sigma*)	ln(hhincome) -0.131***	ln(hhincome) -0.132***	-	-
	(0.00613)	(0.00607)		
	Woman 0.0666***	Woman 0.0685***		
	(0.0102)	(0.0101)		
	Age 0.00396***	Age 0.00396***		
	(0.000281)	(0.000279)		
	White 0.0485***	White 0.0416***		
	(0.00873)	(0.00860)		
	Black 0.00297	Black 0.00158		
	(0.0261)	(0.0259)		
	Asian 0.0518	Asian 0.0499		
	(0.0422)	(0.0428)		
	Chinese 0.000190	Chinese -0.0250		
	(0.184)	(0.182)		
Year Dummies	YES	YES	YES	YES
Geographical Dummies	YES	YES	YES	YES
Observations	91108	91108	91108	91108
AIC: $[-\frac{2}{N}\cdot loglik + 2\frac{k}{N}]$	3.0458	3.0581	3.0760	3.0886
HQ: $[-\frac{2}{N}\cdot loglik + 2\frac{k}{N}\cdot \ln(\ln(N))]$	3.0476	3.0595	3.0775	3.0898
SC: $[-\frac{2}{N}\cdot loglik + \frac{k}{N}\cdot \ln(N)]$	3.0517	3.0626	3.0811	3.0924

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 5. Average Marginal Partial Effects (Subjective Health)

	MODEL			
	(1) HPPOP with FE	**(2) HPPOP no FE**	**(3) PPOP with FE**	**(4) PPOP no FE**
ln(hhincome) MPE's				
1	-0.00388***	-0.00235***	0.00130***	0.00417***
2	-0.00474***	-0.00110	0.00225***	0.00668***
3	-0.00535**	0.00460*	0.00591***	0.0168***
4	0.00832**	0.0252***	0.0104***	0.0284***
5	0.0414***	0.0524***	0.00714***	0.0194***
6	0.00746	-0.0133**	-0.0140***	-0.0383***
7	-0.0432***	-0.0655***	-0.0130***	-0.0371***
ln(hhincome/comparison income) MPE's				
1	-0.000746*	-0.00509***	-0.000906*	-0.00612***
2	-0.00147*	-0.00930***	-0.00157*	-0.00980***
3	-0.00392*	-0.0236***	-0.00413*	-0.0246***
4	-0.00685*	-0.0397***	-0.00726*	-0.0417***
5	-0.00468*	-0.0269***	-0.00499*	-0.0284***
6	0.00923*	0.0534***	0.00981*	0.0561***
7	0.00843*	0.0512***	0.00905*	0.0544***
Deprivation MPE's				
1	-0.000982***	0.000313	-0.000794**	0.000887
2	-0.00194***	0.000571	-0.00138**	0.00142
3	-0.00516***	0.00145	-0.00361**	0.00356
4	-0.00901***	0.00244	-0.00636**	0.00605
5	-0.00616***	0.00165	-0.00437**	0.00412
6	0.0121***	-0.00328	0.00859**	-0.00814
7	0.0111***	-0.00315	0.00793**	-0.00790
Obs	91068	91068	91068	91068

*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 6. Average Marginal Partial Effects (Objective Health)

	MODEL			
	(1) HPPOP with FE	**(2) HPPOP no FE**	**(3) PPOP with FE**	**(4) PPOP no FE**
ln(hhincome) MPE's				
1	-0.00521***	-0.00451***	0.00223***	0.00363***
2	-0.00496***	-0.00351***	0.00321***	0.00510***
3	-0.00406*	-0.000436	0.00754***	0.0118***
4	0.0109***	0.0167***	0.0122***	0.0188***
5	0.0427***	0.0465***	0.00821***	0.0126***
6	0.00912**	0.00210	-0.0164***	-0.0252***
7	-0.0485***	-0.0568***	-0.0170***	-0.0267***
ln(hhincome/comparison income) MPE's				
1	-0.00121**	-0.00605***	-0.00159**	-0.00760***
2	-0.00201**	-0.00977***	-0.00230**	-0.0107***
3	-0.00475**	-0.0227***	-0.00538**	-0.0247***
4	-0.00762**	-0.0360***	-0.00869**	-0.0393***
5	-0.00512**	-0.0241***	-0.00587**	-0.0265***
6	0.0102**	0.0483***	0.0117**	0.0529***
7	0.0105**	0.0504***	0.0121**	0.0559***
Deprivation MPE's				
1	-0.00134***	0.00110	-0.000964*	0.00239***
2	-0.00221***	0.00178	-0.00139*	0.00336***
3	-0.00523***	0.00415	-0.00326*	0.00776***
4	-0.00839***	0.00657	-0.00526*	0.0124***
5	-0.00564***	0.00440	-0.00355*	0.00833***
6	0.0112***	-0.00881	0.00708*	-0.0166***
7	0.0116***	-0.00920	0.00736*	-0.0176***
Obs	91108	91108	91108	91108

*** p<0.01, ** p<0.05, * p<0.1

Chapter 2

Eating Behavior and Social Interactions from Adolescence to Adulthood

Coauthored with Luisa Corrado

Abstract

This paper analyzes the importance of social ties for eating behavior of US youth. We propose a novel approach that addresses identification of social endogenous effects. We overcome the problem of measuring the separate impact of endogenous and contextual effects on individual body mass index in a dynamic linear-in-means model, where individual- and group-specific unobservable effects are controlled for. We show that the main drivers of eating behavior are habituation and imitation effects. Imitation effects explain most of the variation in body mass index of individuals who were normal-weight and overweight during adolescence. Obese adolescents, instead, become future obese adults through wrong habits enforced by imitative behavior.

2.1 Introduction

Overweight and obesity are social plagues of modern societies. According to the World Health Organization (WHO), there are more than one billion overweight adults in our globe, at least three hundred million of them obese, and figures are even worse for children and adolescents. This is a multifaceted condition with social and psychological dimensions in all ages and socioeconomic groups. The rising epidemic surely reflects significant changes in eating behavioral patterns of communities: over-consumption of carbohydrates and saturated fats as well as scarce physical activity are important causes. However, social and peer effects likely act as drivers of such a large-scale phenomenon, that seems to occur independent of cultural, economic, and environmental circumstances (Cohen-Cole, 2006; Christakis and Fowler, 2007; Trogdon et al., 2008; Fortin and Yazbeck, 2011). Indeed, many epidemic phenomena occur because they spread within social groups through the homogenization of behaviors among individuals of the same group (Manski, 2000): hence, the propensity of a person to behave in a certain way varies positively with the dominant behavior in her group, similar to informal enforcement mechanisms or social norms (Kandori, 1992; Bernheim, 1994).[1]

The aim of the present paper is to estimate the impact of social and peer effects on eating behavior of US adolescents who transition into adulthood.[2] By means of a novel, yet simple empirical approach we overcome the problem of identifying the impact of social endogenous effects on individual Body Mass Index (BMI). Furthermore, we examine to what extent adulthood BMI status depends on habituation and imitation during adolescence.

Adopting the notion first introduced by Manski (1993), we can identify three types of group effects impacting on individual behavior: endogenous effects, which occur when individual behavior varies with the behavior of the group; contextual effects, that arise

[1] Peer effects have been extensively examined both in education (Bénabou, 1993) and in psychology (Brown, 1990; Brown et al. 1996). For a review of the literature on social interaction effects see Brock and Durlauf (2001).

[2] The expression 'eating behavior' comprises all the actions having influence on body weight, e.g. quantity and quality of food, physical exercise, and lifestyle-related issues.

when peer group characteristics directly affect individual behavior; correlated or group (unobservable) effects, that arise because group members share a common environment or common latent traits that affect their individual behavior. Analyzing the statistical effect of social interactions is generally challenging due to a special kind of identification problem, the so called reflection problem (Manski, 1993): in a linear-in-means model of social interactions, the distinct role of endogenous and contextual effects may be difficult to disentangle because such effects co-move. Since the work of Manski (1993), many are the studies that tackle the estimation of peer effects (see Brock and Durlauf, 2001; Moffitt, 2001 for a review of the literature). Recent empirical work seems characterized by writhed frameworks that require information on the network structure (Bramoullé et al., 2009; Petacchini et al., 2010; Fortin and Yazbeck, 2011; Corrado and Fingleton, 2012) or on out of group effects (Cohen-Cole, 2006).[3] Different from this literature, we resort to a simple and reasonable framework to identify social endogenous effects in a linear-in-means model. Specifically, we estimate a dynamic linear-in-means model that allows individual behavior to linearly depend on individual past behavior as well as on group-specific effects, which include some group observable characteristics and the expected aggregate behavior of the others in the group. Such an assumption makes sense when not only choice is thought of as being the result of social and peer effects, but also of past behavior. Habituation as well as social behavior are possible determinants of individual choice, especially in the case of eating decisions. Furthermore, our econometric strategy allows us to control for endogeneity, individual and group heterogeneity by exploiting stationarity restrictions of a system GMM estimator augmented to control for individual- and group-specific unobservable effects.

We make use of the National Longitudinal Study of Adolescent Health (Add Health)

[3] Specifically, Cohen-Cole (2006) uses out of group effects to identify endogenous and contextual effects. However the method does not take into account individual- and group-specific unobservable effects. In another recent paper De Giorgi et al. (2010) show that, in a context where peer groups do not overlap fully, it is possible to identify all the relevant parameters of the standard linear-in-means model of social interactions. The Instrumental Variable approach proposed in their paper, though, while properly accounting for correlated unobservable effect at the group level, does not provide a separate estimation of unobservable effects at the individual and group level.

dataset, a (US) representative sample of adolescents who transition into early adulthood for which information on demographic, health and socioeconomic status is registered along four waves, from 1994 to 2008.[4] We are able to study the behavioral causes of overweight and obesity among teenagers, and the effects of such behavior during their transition to adulthood. In contrast to Cohen-Cole and Fletcher (2008), who also focus on obesity and social interactions using the Add Health dataset, the results of our dynamic linear-in-means model show that the tendency of individuals to become overweight is the outcome of both social effects and past individual behavior. Cohen-Cole and Fletcher (2008) general finding is that social interactions with closest peers are not significant once fixed effects for social groups and individuals are accounted for. In our estimation, instead, social endogenous effects - i.e., the tendency to be affected by the behavior of others in the same school - are still present even after accounting for school and individual effects. We estimate that a 1% variation in average group BMI produces a 0.44% variation in current BMI status. Such finding is in line with the evidence reported in Christakis and Fowler (2007) and Trogdon et al. (2008). The former study analyzes data on a social network of people pertaining to the Framingham Offspring Study, finding that a person's chance of becoming obese increases by 57% if a friend became obese in a given interval. The latter makes use of Add Health data to estimate a coefficient of friends' average BMI of about 0.50. However, on the one hand Christakis and Fowler (2007) make use of a statistical method which ignores school and individual effects, and are criticized by Cohen-Cole and Fletcher (2008) as their results may not merely represent endogenous social interactions. On the other hand, Trogdon et al. (2008) base their identification of peer effects on an instrumental variable estimation in a cross-sectional analysis, where the instruments are obesity of (two self-nominated) friends' parents, and friends' birth weight; in their study a friend-selection effect might be driving results, rendering the appropriateness of the instrumental variables questionable.

[4] The time distance between each wave is not homogeneous. The first and the second waves are two consecutive years, the third is 6 years later than the second, and the fourth 6 years later than the third. We have weighted the sample in order to account for this difference in gaps between waves (Cf. Appendix B).

Our study differs from the ones just mentioned in many aspects, which are considered crucial for results to be reliable. First, we believe that schoolmates represent the best approximation of a potential reference group, as these are individuals who adolescents compare and interact with in their everyday life, especially during meals. Second, looking at the OLS estimations proposed by Christakis and Fowler (2007) and Cohen-Cole and Fletcher (2008), the presence of a lagged dependent (or independent) variable in a social interactions model can lead to substantial biases in the estimation, unless properly addressed.[5] In this respect, misspecification of the model or of the error structure can lead to very large biases and thus incorrect inference. Our proposed GMM approach conciliates the different positions. Indeed, the econometric framework and estimation strategy used aim at overcoming the limits of the previous works, while relying on plausible hypotheses about habituation and reference groups. Third, we are able to trace out eating behavioral patterns from adolescence to adulthood. We show that obese teenagers become obese adults picking up wrong habits which are enforced by imitative behavior; the coefficient of autocorrelation for this category is 0.97 and the one related to group BMI is greater than one. The story seems different for adults who were normal-weight and overweight during adolescence: their adult outcome is not highly correlated with the past; rather, the role of peers at school has a crucial importance for their current BMI.

Finally, we deal with missingness in the dependent variable (BMI) and in the most important economic variable, income, by replacing missing values using a multiple-imputation method. This has a dramatic effect on the coefficient of average BMI when we consider estimation results for different weight categories. Specifically, our results show that for individuals who were overweight and obese adolescents there is a marked increase in the group effect coefficient. Such result is supportive of the hypothesis that certain categories of individuals are less likely to report their weight so that average group effects will be downsized if missingness in the dependent variable is not properly

[5] Liu et al. (2006) find evidence of significant bias in estimation relating to the dynamic role of social interactions by making use of simulation techniques.

accounted in the estimation.

The paper is structured as follows. Section two illustrates the main identification issue arising in linear-in-means models of social effects and shows how to resolve the identification problem in panel data using the lagged endogenous variable as an internal instrument. Section three describes the system GMM estimation strategy employed in the paper. Section four describes the Add Health dataset and presents the main results. Section five concludes.

2.2 Linear-in-Means Models of Social Interactions

2.2.1 The Linear-in-Means Model and the Reflection Problem

The baseline LMM is conceptually very simple. Usually not derived from any predefined individual decision problem, this model allows individual behavior to linearly depend on some individual-specific characteristics as well as on group-specific factors, which include some group observable characteristics and the expected aggregate behavior of the others in the group.[6] This makes it easily interpretable as a regression model, and therefore interesting to the econometrician. However, as pointed out by Manski (1993), the LMM suffers from a special kind of identification problem - the so called reflection problem - due to difficulties in disentangling two different group-effects, namely contextual and endogenous effects. Therefore, in such a framework measuring the impact of social interactions is typically challenging.

Consider the simple version of the model, where estimation concerns are not yet addressed. Assume to have G non-overlapping, *a priori* determined groups, each of them made of N^g individuals. Individual choice is assumed to be the result of the following process:

[6] LMM can be the result of an optimal decision problem framed around agent's choice, as illustrated in Brock and Durlauf (2001).

$$y_{ig} = a + y^e_{ig}\beta + \mathbf{x}'_g\boldsymbol{\gamma} + \mathbf{r}'_{ig}\boldsymbol{\delta} + \varepsilon_{ig} \text{ , where } \begin{array}{c} g = 1, \ldots, G \\ i = 1, \ldots, N^g \end{array} . \tag{2.1}$$

The individual-specific terms are defined by a $r \times 1$ vector of observable characteristics, $\mathbf{r}_{ig}$, and ε_{ig}, a random and unobservable scalar assumed to be independent and identically distributed across individuals. As to group-specific factors, these are divided into a $k \times 1$ vector of predetermined characteristics, $\mathbf{x}_g$, and the expected average choice in the group, y^e_{ig}. These two terms are conceptually different, the former being interpreted as contextual effects and the latter as an endogenous effect, and those exist under the condition that β is non-zero and $\boldsymbol{\gamma}$ has at least a non-zero element. The key effect is exerted by y^e_{ig}, since it creates reciprocal reactions between individual decisions.

Using expected average behavior rather than the realized one is merely due to analytical convenience. This is a reasonable assumption when the behaviors of the rest of group are not directly observable - i.e., in large groups. When it comes to empirical analysis, such an assumption presupposes a restriction on the way individuals form expectations about the average choice in their group. Specifically, expectations are supposed to be consistent with the structure of the choices in the model, or self-consistent. This means that the perceived average choice is equivalent to the mathematical conditional expectation of the average choice, y^e_g, given the information set of each individual. The information set includes values of r_{ig} for other individuals within i's group, as well as the equilibrium expected choice level that occurs for her group. Individuals are assumed to be unable to observe the choices of others, y_{-ig}, or their random payoff terms ε_{ig}. Alternative information assumptions will not affect the qualitative properties of the model. For the LMM, self-consistency amounts to:

$$y^e_{ig} = y^e_g = \frac{a + \mathbf{x}'_g\gamma + \mathbf{r}'_g\boldsymbol{\delta}}{1-\beta} = \frac{a + \mathbf{x}'_g\gamma}{1-\beta} + \frac{\mathbf{r}'_g\boldsymbol{\delta}}{1-\beta}, \tag{2.2}$$

where $\mathbf{r}_g$ is the average of $\mathbf{r}_{ig}$ within group g.

Notice that such an assumption on the aggregate outcome implies a unique equilibrium: there exists only one expected average choice level that is consistent with the model, given individual and group characteristics. Therefore, equation (2.2) maps these characteristics into a single y_g^e.

An identification problem in this framework could arise because endogenous and contextual effects may co-move. Indeed, under the self-consistency assumption, the contextual variables determine the endogenous variable, as indicated by condition (2.2). Given that the identification failure is a consequence of the correlation, by construction, between the endogenous and the contextual effects, Manski (1993) renamed it 'reflection problem', which is not too dissimilar from the basic identification problem in linear regressions with linearly dependent covariates. Manski's original argument is that every contextual effect might be defined as the average of a corresponding individual characteristic. For example, if one controls for student's maternal education one also introduces average (school) maternal education so that $\mathbf{x}_g = \mathbf{r}_g$. Condition (2.2) becomes

$$y_g^e = \frac{a + \mathbf{x}_g' (\boldsymbol{\gamma} + \boldsymbol{\delta})}{1 - \beta}, \tag{2.3}$$

meaning that the regressor $y_{ig}^e = y_g^e$ in (2.1) is linearly dependent on the regressors a and $\mathbf{x}_g$ in (2.1), so the parameters are not identified. Substituting (2.3) into (2.1):

$$y_{ig} = \frac{a}{1 - \beta} + \frac{\beta}{1 - \beta} \mathbf{x}_g' (\boldsymbol{\gamma} + \boldsymbol{\delta}) + \mathbf{r}_{ig}' \boldsymbol{\delta} + \varepsilon_{ig}. \tag{2.4}$$

We can therefore state the following two remarks on the identification of social interaction effects in a LMM:

Remark 1 *In the empirical model (2.1) the set of regressors* $(1, y_g^e, \mathbf{x}_g, \mathbf{r}_{ig})$ *requires the estimation of* $2 + k + r$ *parameters.*

Remark 2 *Assuming reflection* $\mathbf{r}_g = \mathbf{x}_g$ *in the reduced form (2.4) the set regressors* $(1, \mathbf{x}_g, \mathbf{r}_{ig})$ *allows us to identify* $1 + k + r$ *parameters. Hence, the endogenous effect parameter,* β*, remains unidentified.*

It is then clear why in the LMM framework identification of parameters is a major challenge. In the remainder of this section we show how to achieve identification of the endogenous effect parameter, β.

2.2.2 An AR(1) Linear-in-Means Model: Breaking the Reflection Problem

We discuss a dynamic LMM of social interactions, and show how the reflection problem can be broken. Consider a case in which the econometrician has access to a grouped panel, with G non-overlapping groups ($g = 1, \ldots, G$) of individuals and N^g individuals ($i = 1, \ldots, N^g$) sampled in the g^{th} group. The following autoregressive model generates the observed data:

$$y_{t,ig} = a + y_{t-1,ig}\varphi + y^e_{t,ig}\beta + \mathbf{x}'_{t,g}\boldsymbol{\gamma} + \mathbf{r}'_{t,ig}\boldsymbol{\delta} + \varepsilon_{t,ig} \tag{2.5}$$

In practice, the set of individual-specific attributes supposed to be determining individual behavior at time t is assumed to depend on past period choice, $y_{t-1,ig}$. Such an assumption makes sense when not only is choice thought of as being the result of contemporaneous exogenous characteristics, but also of a certain past behavior that could play a role in actual choice. Extending the example on peer effects and students' obesity, we use student's body mass index in the previous period, $y_{t-1,ig}$, as an internal instrument to resolve the reflection problem since it will be orthogonal to the error term. The use of internal instruments to solve endogeneity problems is advocated for example by Lewbel (1997). Lewbel's idea is that when the endogenous regressor has a skewed distribution certain transformations of the data, including using lagged endogenous effects, provide a set of valid instruments.

The self-consistency condition in this case is:

$$y^e_{t,ig} = y^e_{t,g} = \frac{a + y_{t-1,g}\varphi + \mathbf{x}'_{t,g}\boldsymbol{\gamma} + \mathbf{r}'_{t,g}\boldsymbol{\delta}}{1-\beta} = \frac{a + \mathbf{x}'_{t,g}\boldsymbol{\gamma}}{1-\beta} + \frac{y_{t-1,g}\varphi + \mathbf{r}'_{t,g}\boldsymbol{\delta}}{1-\beta}. \tag{2.6}$$

The term $y_{t-1,g}$ is the average choice in the group in $t-1$, which enlarges the individual information set among the observable effects. Therefore, even under the assumption $\mathbf{x}_{t,g} = \mathbf{r}_{t,g}$, there is an additional element, $y_{t-1,g}$, which allows identification. Indeed, the social equilibrium equation is:

$$y^e_{t,g} = \frac{a + \mathbf{x}'_{t,g}(\boldsymbol{\gamma} + \boldsymbol{\delta})}{1-\beta} + \frac{y_{t-1,g}\varphi}{1-\beta}. \tag{2.7}$$

Substituting the social equilibrium into (2.5) yields:

$$y_{t,ig} = \frac{a}{1-\beta} + y_{t-1,ig}\varphi + \frac{\beta}{1-\beta}\mathbf{x}'_{t,g}(\boldsymbol{\gamma} + \boldsymbol{\delta}) + \frac{\beta\varphi}{1-\beta}y_{t-1,g} + \mathbf{r}'_{t,ig}\boldsymbol{\delta} + \varepsilon_{t,ig}. \tag{2.8}$$

Clearly, the model is now identified.

Proposition 1 *In the empirical model (2.5) the set of regressors* $(1, y_{t-1,ig}, y_{t,g}, \mathbf{r}_{t,ig}, \mathbf{x}_{t,g})$ *requires the estimation of* $(3 + r + k)$ *parameters.*

Proposition 2 *Assuming reflection* $\mathbf{r}_g = \mathbf{x}_g$ *in the reduced form (2.8) the set regressors* $(1, y_{t-1,ig}, y_{t-1,g}, \mathbf{r}_{t,ig}, \mathbf{x}_{t,g})$ *allows to identify* $(3 + r + k)$ *parameters. Hence, all the parameters in the empirical equation (2.5) are identified and the ratio of the two coefficients* $\frac{\beta\varphi}{1-\beta}$ *and* φ *gives the endogenous effect* β*.*

The model avoids the linear dependence between $y_{t,g}$, $\mathbf{x}_{t,g}$ and $\mathbf{r}_{t,g}$ since we have the average action of the group in the previous period, $y_{t-1,g}$, as an additional regressor. This implies that $y_{t,g}$ depends on the entire history of $\mathbf{x}_{t,g}$ and $\mathbf{r}_{t,g}$ resolving the contemporaneous correlation with the same variables. Once the correlation is resolved, we can get an efficient and consistent estimation of all the parameters. Specifically, in the following section we illustrate how to estimate the social interaction parameters in the empirical equation (2.5).

2.3 Estimation

We consider the following econometric framework:

$$\begin{aligned} y_{t,ig} &= y_{t-1,ig}\varphi + y^{e}_{t,ig}\beta + \mathbf{x}'_{t,g}\boldsymbol{\gamma} + \mathbf{r}'_{t,ig}\boldsymbol{\delta} + e_{t,ig},\ |\varphi| < 1 \\ e_{t,ig} &= \alpha_g + u_{t,ig}, \\ u_{t,ig} &= f_i + \varepsilon_{t,ig} \end{aligned} \tag{2.9}$$

where we allow for individual-specific effects, captured by f_i as well as for group-specific effects, α_g; $\varepsilon_{t,ig}$ is an individual-specific random disturbance.[7] Appendix A demonstrates that system (2.9) accounts for correlated effects both at the individual and group level so that α_g and f_i can be treated as random.[8]

In order to account for the presence of endogeneity, we assume that:

$$E\left[\mathbf{h}_{t,ig}\ \varepsilon_{s,ig}\right] \neq 0, \quad \mathbf{h}_{t,ig} = \left[\mathbf{x}_{t,g}, \mathbf{r}_{t,ig}\right] \tag{2.10}$$

$i = 1, ..., N^g$, $g = 1, ..., G$, and $s \leq t$. This assumption allows both for contemporaneous correlation between current disturbances and covariates and feedbacks from past shocks into the current value of the covariates. Moreover, the following assumptions hold:

$$\begin{aligned} E\left[\varepsilon_{t,ig} \middle| \mathbf{X}_t\right] &= 0 \\ Var\left[\varepsilon_{t,ig} \middle| \mathbf{X}_t\right] &= \sigma^2_\varepsilon \end{aligned}$$

where $\mathbf{X}_{t,ig} = [y_{t-1,ig}, y_{t,g}, \mathbf{x}_{t,g}, \mathbf{r}_{t,ig}]$.

We assume that f_i and $\varepsilon_{t,ig}$ are independently distributed across individuals and

[7] Notice that $a = \alpha_g + f_i$ as in equation (2.5).

[8] Most of the current research on social interaction effects (see De Giorgi et al. 2010 among others) is also accounting for potential correlated (unobservable) effects at the group level, without estimating these effects.

have a familiar error structure in which:

$$E[f_i] = 0,\ E[\varepsilon_{t,ig}] = 0,\ E[f_i\ \varepsilon_{t,ig}] = 0 \quad \text{for } t = 2, ..., T, \quad i = 1, ..., N^g, \quad g = 1, ..., G$$

and

$$E[\varepsilon_{t,ig}\ \varepsilon_{s,ig}] = 0, \quad \forall t \neq s. \tag{2.11}$$

In addition, we impose the initial condition

$$E[y_{1,ig}\ \varepsilon_{t,ig}] = 0 \text{ for } t = 2, ..., T,\ i = 1, ..., N^g,\ g = 1, ..., G \tag{2.12}$$

Conditions (2.11) and (2.12) imply the following moment $m = 0.5(T-1)(T-2)$ conditions:

$$E[y_{t-s,ig}\ \Delta\varepsilon_{t,ig}] = 0 \text{ for } t = 3, ..., T,\ s \geq 3$$

First difference GMM can poorly behave when time series are highly persistent, as lagged levels of the series provide only weak instruments for subsequent first differences. In addition, first differencing would lead to lose substantial information from contextual effects, which are somewhat time-invariant. Therefore, we resort to a more efficient GMM estimator that exploits stationarity restrictions. Bond et al. (2001a) show that this system GMM estimator provides more reasonable estimates than first-differenced GMM.[9] Blundell and Bond (1998) consider the additional assumption that

$$E[f_i\ \Delta y_{2,ig}] = 0, \text{ for } i = 1, ..., N^g \text{ and } g = 1, ..., G \tag{2.13}$$

[9] We have four waves and 4443 respondents therefore we use the Arellano-Bond estimator which was designed for small T large N panels. The second lag is required, because it is not correlated with the current error term, while the first lag is. This is also shown by the Arellano-Bond test for autocorrelation which has a null hypothesis of no autocorrelation and it is applied to the differenced residuals. The test for AR (1) process in first differences usually rejects the null hypothesis (as in our results reported in Table 6).

This further assumption implies additional $T-2$ linear moment conditions:

$$E\left[u_{t,ig}\ \Delta y_{t-1,ig}\right]=0, \text{ for } t=3,...,T,\ i=1,...,N^g,\ g=1,...,G \qquad (2.14)$$

These allow us to use lagged first-differences of the series as instruments for the equation in levels, as suggested by Arellano and Bover (1995).

2.4 Data and Results

2.4.1 The National Longitudinal Study of Adolescent Health

The National Longitudinal Study of Adolescent Health (Add Health)[10] is a (US) nationally representative, school-based survey of youth. The study was designed to determine how peers (within family, schools, neighborhoods, and communities) as well as individual characteristics influence health behaviors and therefore health outcomes.

While initially focused on adolescents only, in later phases the study analyzes health and health behaviors during the transition from adolescence into early adulthood. Indeed, in the first years of adulthood the young develop habits, and choose their lifestyle so that future health and well-being are strongly affected by such behaviors. It is therefore possible to study what happens during the transition to adulthood, as well as to explore early behavioral causes of adult chronic diseases.

The survey is made of four waves. In $1994-1995$ a random sample of $7th$ to $12th$ grade students from schools across the country was selected. About $90,000$ young individuals participated by filling out a brief questionnaire at school. Afterwards, at-home interviews with students and their parents were conducted. Students were

[10] This research uses data from Add Health, a program project directed by Kathleen Mullan Harris and designed by J. Richard Udry, Peter S. Bearman, and Kathleen Mullan Harris at the University of North Carolina at Chapel Hill, and funded by grant P01-HD31921 from the Eunice Kennedy Shriver National Institute of Child Health and Human Development, with cooperative funding from 23 other federal agencies and foundations. Special acknowledgment is due Ronald R. Rindfuss and Barbara Entwisle for assistance in the original design. Information on how to obtain the Add Health data files is available on the Add Health website (http://www.cpc.unc.edu/addhealth) and from Harris et al. (2009). No direct support was received from grant P01-HD31921 for this analysis.

interviewed again in their homes one year later (1996). School administrators provided information about the schools participants attended and existing data were compiled to describe neighborhoods and communities (in both waves 1994 – 1995 and 1996). In the last two waves (2001 – 2002; 2007 – 2008) participants in the first in-home interview were re-interviewed at ages 18 to 26, and again at ages 24 to 32.

The survey contains information on demographics, family life and background, school and academic outcomes, and health behaviors (drug use, smoking, pregnancy, etc.). For this research, the desired sample is the one relative to the in-home survey of the public-use data sets.

The reader is cross-referred to Appendices B and C for further details on design, weighting and missing information of data at hand.

2.4.2 The Dependent Variable

Our dependent variable is BMI, constructed using self-reported height and weight.[11] BMI is an index of weight-for-height which is age-independent and the same for both sexes. It is computed as weight in kilograms divided by the square of height in metres (kg/m^2) and it is standardly used to classify underweight, overweight and obesity.[12]

Table 1 shows the international classification of underweight, overweight and obesity according to BMI, as reported on the WHO studies.[13]

Table 1 about here

Based on this classification, we constructed a transition matrix for BMI in order to analyze the dynamic behavior of the variable in our sample.

[11] We make use of self-reported height and weight because Add Health wave 1 lacks information on measured height and weight. However, it has been shown that BMI computed using self-reported variables is highly correlated with BMI generated using measured height and weight ($r = 0.92$), and correctly classifies 96% as to obesity status (Goodman et al., 2000).

[12] BMI is not a direct measure of body fatness. However, it parallels changes obtained by direct measures of body fat such as underwater weighing and dual energy x-ray absorptiometry (DXA), therefore it can be considered as a proxy for measures of body fat.

[13] http://apps.who.int/bmi/index.jsp?introPage=intro.html

Table 2 about here

First of all, we notice that more than 50 percent of individuals has a normal body weight, while the probability of facing an overweight individual is about 25% and an obese one about 17%.[14] Probabilities located on the main diagonal are quite high, meaning that BMI is highly autocorrelated, especially for normal-weight and heavily obese people. Such a finding strongly corroborates the validity of our empirical specification which includes lagged BMI among the set of regressors, given our hypothesis that habituation effects as well as imitation effects explain current BMI.

2.4.3 Descriptive Statistics

Tables 3 to 5 show the summary statistics of the sample under analysis.

Tables $3-5$ about here

Table 3 displays the descriptive statistics relative to our dependent variable of interest, BMI, and control variables for the whole sample covering all waves. Figure 1 shows that average BMI is close to the threshold between normal weight and overweight - as predicted by the transition matrix. Having a look at the distribution of BMI (Figure 1) we realize that the modal bins are BMI$=20-22$ and BMI$=22-24$, meaning that normal-weight individuals are those for whom frequency is highest. Furthermore, the distribution appears to be right-skewed, signaling a majority of overweight and obese individuals in the sample observed. Household income levels (Figure 2) are in line with those reported by the US Census Bureau. Discrepancies between our data and the US Census Bureau data probably lie in the very high percentage of missing values, a problem addressed in the estimation.[15] Furthermore, in wave four self-reported income information comes in range format. Thus, in order to achieve a longer panel specification for household

[14] Obtained by summing the percentages of all the obese categories, i.e. 11.03%, 4.28% and 3.31%.
[15] Cf. Appendix C.

income, we take the average of each income brackets as the point information related to each individual. This might explain why we observe observations clustered around some values, as clearly visible in Figure 2.

Concerning the sample composition of some characteristics of interest, we observe that the proportion of females is slightly larger than males, registering 57% of counts. Also, the two wider ethnic groups are white and African American, while American Indian and Asian groups have a very small impact on the ethnical composition. The vast majority of individuals is in a good to excellent health status, while only 28% of the sample lives in a completely urban city, and 29% lives in a geographical area with low unemployment rate (though this variable shows a high percentage of missing values). Finally, the figure on parental education shows a low percentage of college graduate, both on the mother and on the father side.

All variables but income have a percentage of missing values of about 20%, which we consider acceptable and equally distributed across characteristics. We decide to deal with income missingness, instead, as it seems to be quite significant, and with BMI missingness, as it is our dependent variable. Specifically, given that peer-effects are derived as averages of individual BMI by dealing with missingness in this variable we explicitly take into account that certain categories may be less likely than others to report their weight, which can bias the definition of average (group) BMI. Details about the procedure are described in Appendix C.[16]

Tables 4 and 5 are informative about variations in average BMI depending on certain characteristics, for both the entire sample and the adolescent subsample (first two waves only) respectively. The only difference for adolescents is that average BMI is in general lower. In both tables the most important figures are the correlation of poorer health statuses with higher average BMI, and the correspondence of lower BMI statuses to higher parental education. Also, those with a household income greater than the median show a higher BMI on average. Therefore, what comes out is that parental education,

[16] Missingness here is not due to design effects, as data have been previously weighted and therefore adjusted to account for those.

income (possibly correlated with parental education), and health are important factors for individual eating behavior.

2.4.4 Reference Groups

A crucial issue in the analysis of social endogenous effects in eating behavior is the definition of reference groups. Many papers attempting to address the complexities of social interactions in obesity rely on the nomination of adolescents' closest peers or on family history, and this is always subject to selection problems that the authors do not seem to address (e.g., the most important ones, Christakis and Fowler, 2007; Cohen-Cole and Fletcher, 2008; Fowler and Christiakis, 2008; Trogdon et al., 2008).[17] Besides, it could be restrictive to consider self-nominated friends or family as the only plausible reference group, especially for phenomena like overweight or obesity which may depend on social norms and acceptance in a broader context. Rather, we believe that schoolmates better fit the potential reference group adolescents compare and interact with. Indeed, interconnections between members of the same school may determine mutual influence through a variety of factors, e.g., food quality and quantity, time spent to exercise, appearance, etc. It is likely that contextual effects (those exerted by environmental factors) on eating behavior are common to schoolmates, and may drive similarities in individual behavior - therefore in their body weight.

Hence, our peer groups correspond to all the individuals belonging to the same school, meaning that endogenous effects measure the propensity to become overweight due to a direct interaction within the school. Such a choice is consistent with our dynamic analysis of eating behavioral patterns from adolescence to adulthood, mostly because it is believed that what affects the transition of body weight into early adulthood is behavior during adolescence (e.g., *inter alia*, Kemper et al., 1999; Sun Guo et al., 2002; Kvaavik et al., 2003; Gordon-Larsen et al., 2004) which in turn depends on schoolmates

[17] Cohen-Cole and Fletcher (2008) and Christakis and Fowler (2007) consider the data on obesity status for an individual (in their terminology, an "Ego") at a given point in time and estimate its relationship to the obesity status of a friend ("Alter").

behavior (cf. Section 4.6).

2.4.5 Estimation Results

In this section we report results produced by estimating the system defined in (2.9).[18]

A premise is due at this stage. We make clear to the reader that the lagged dependent variable on the right hand side does not refer to the value of the dependent variable the year before, as the gaps between waves are not homogeneous; rather, that embeds all past history up to the previous wave. In particular, wave one and two are consecutive years registering information on adolescents, wave three is 6 years later than wave two and includes data on early adults, wave four is again 6 years later than wave three and contains information on adults. Data have been purposely weighted to account for uneven time gaps, and a dummy variable for being adolescent (observations registered in waves one and two) has been included to capture variation due to being part of the adolescent cohort versus belonging to the adult cohort.

Our specification allows us to investigate the hypothesis that obesity can spread through peers versus the claim that obesity is essentially an individual outcome linked to personal and family history. We also establish whether peer effects may be stronger for obese pupils compared to the non-obese counterparts.

As pointed by Cohen-Cole and Fletcher (2008), in order to avoid spurious conclusions on the role exerted by group behavior the estimation should include contextual effects.[19] In other terms both individual and group behavior can be affected by exposure to common influences: for example, the opening of a fast food, gym or recreational area near a school could simultaneously affect the weight of all pupils in the same school.

[18] The Arellano–Bond estimators is available for Stata 9.0 as proprietor program written by Roodman (2006) (called `xtabond2`). See http://ideas.repec.org/c/boc/bocode/s435901.html.

[19] The paper by Cohen-Cole and Fletcher (2008) argues that previous studies on the spread of obesity (Christakis and Fowler, 2007) do not include a sufficiently broad set of contextual effects to account for a range of hypothesized causes of obesity, therefore overstating the endogenous effect. Corrado and Fingleton (2012) also suggest that the significance of a spatially lagged dependent variable involving network dependence and spatial externalities may be misleading, since it may be simply picking up the effects of omitted spatially dependent variables, incorrectly suggesting the existence of a spillover mechanism.

Since access to such facilities may be linked to the socio-demographic characteristics of the adolescents in the same school, we include the average school values for household income, age, gender, ethnicity and parental education in the estimation.[20,21] Therefore, as in Cohen-Cole and Fletcher (2008), we consider a time-dependent set of school specific covariates, $\mathbf{x}_{t,g}$. These represent a much richer set of controls to absorb the average change in social context experienced by all individuals in the sample. They can also be interpreted as school-specific trends which account for environmental factors shared by adolescents in the same school. Clearly, more environmental confounders may exist which are positively correlated with an individual's BMI. We therefore enrich our instrumental variables set by adding two location-specific variables indicating whether the neighborhood where the individual resides is characterized by a low unemployment rate, and whether the adolescent lives in a completely urban area.[22,23] These environmental confounders reflect the social context of the geographical area where the respondents reside and represent a valid set of instruments since they are likely to be correlated with both individual and group BMI but neither with the unobserved individual propensity or tolerance to become overweight, nor with unobservable effects at the school level.

We employ a system GMM estimation which uses the levels equation (2.5) to obtain a system of two equations: one differenced and one in levels. Additional instruments can be obtained by adding the second equation so that variables in levels can be instrumented with their own lags. This usually increases efficiency. The set of endogenous variables $[y_{t-1,ig}, y_{t,g}]$ includes lag individual BMI, $y_{t-1,ig}$, and contemporaneous group BMI, $y_{t,g}$;

[20] In system GMM, one can include time-invariant regressors, which would disappear in difference GMM. Asymptotically, this does not affect the coefficients estimates for other regressors. This is because all instruments for the levels equation are assumed to be orthogonal to the fixed effects, thus to all time-invariant variables; in expectation, removing them from the error term does not affect the moments that are the basis for identification.

[21] We also performed estimation with average effects at the school level and centered effects at the individual level in order to account for potential collinearity among regressor and the results were similar. These additional results are available on request.

[22] The definition of neighborhood follows a geographical criterion as such community variables are based on state, county, tract, and block group levels derived from addresses.

[23] Our estimation also accounts of wave effects, through the inclusion of time and adolescent dummies.

these are instrumented with GMM style instruments, i.e., third and fourth lags of the endogenous variables $[y_{t-3,ig}, y_{t-4,ig}, y_{t-3,g}, y_{t-4,g}]$. The exogenous variables chosen as set of standard instruments $[\mathbf{r}_{t,ig}, \mathbf{x}_{t,g}, \mathbf{z}_t]$ include the exogenous controls, $\mathbf{r}_{t,ig}$, their school average, $\mathbf{x}_{t,g} = \mathbf{r}_{t,g}$, and two additional instruments, $\mathbf{z}_t$, characterizing the macro-area where each adolescent lives (urban and employment rate).

In order to compare the results at hand with previous findings by Cohen-Cole and Fletcher (2008) we also address the issue of missing data. In the dataset we register $3,372$ missing observations for income and $1,654$ missing observation for BMI. We use a Multiple Imputation method to estimate these missing values as described in Appendix C, because we expect missingness at random to be explained by covariates included in our model (e.g., ethnicity or gender).

The third column in Table 6 reports the estimates for the system defined in (2.9) where missingness in income and BMI are accounted for.[24] Results show that current BMI is affected both by past individual decisions and social behavior. In fact, an increase by 1% in past BMI leads to an increase in current BMI by 0.83%. This result is very much in line with the evidence from the transition probabilities in Table 2 where BMI is highly autocorrelated, especially for heavily obese people. Looking at peer effects, we can see that an increase by 1% in the average BMI leads to an increase in current BMI by 0.44%. Both the Sargan test and the Hansen test indicate that the instruments chosen as a group are exogenous. Looking at the significance of other controls, we find that adolescent of Asian ethnicity tend to experience a lower BMI than their White and Black counterparts. Other studies also show that the prevalence of overweight and obesity among Asian Americans is much lower than the national average and all other main racial/ethnic groups (Gordon-Larsen et al., 2003; Popkin and Udry, 1998). In addition, adolescents belonging to Black ethnic groups have a higher BMI. This result is also in line with other evidence using Add Health data showing that lower socioeconomic status and minority population groups have less access to physical activity facilities, which in

[24] We use a robust standard errors estimation where the standard covariance matrix is robust to panel-specific autocorrelation and heteroskedasticity. We also bootstrap the standard error and find no difference with the robust standard errors. The results are available on request.

turn is associated with decreased physical activity and increased overweight (Gordon-Larsen et al., 2006). We also find that obesity is less widespread among adolescents whose father gained a college education. There are different channels through which parental education can affect their children's health. Education might have a direct impact on child health because it helps parents to make better health investments for themselves and their children. Alternatively, education can affect child health indirectly. An increased level of education can give access to more skilled work with higher earnings and these resources could be used to invest in health (Case et al., 2002; Lindeboom et al., 2009). In the presence of assortative mating, individuals with a higher level of education also marry partners with higher levels of education, which positively affects family income. In this respect, public health strategies aimed at preventing obesity may need to target families of low socioeconomic status early in children's lives, in order to counteract the adverse effect of poor socioeconomic status on parental health and eating decisions.[25]

If we consider imputed data for income only the qualitative results do not change insofar past individual behavior still dominates. In this case an increase by 1% in past individual BMI leads to an increase by 0.67% in current individual BMI whereas an increase in average BMI leads to an increase in current individual BMI by 0.46%. It is worth stressing that when we impute data for BMI alongside income the coefficient for Black ethnicity is now significant. This seems to support the evidence that data are not missing completely at random, and that weight self-reported information might be dependent on individual ethnicity with Black being less likely to report their weight than individuals of other ethnicities. The results also show that dealing with missingness of both income and BMI increases by 76% the coefficient of average BMI (from 0.25 to

[25] We also estimate a model omitting the health dummies and their group averages among the set of regressors (available on request). This exclusion has the effect of amplifying the impact of lagged individual BMI on individual BMI and to downsize the effect of average BMI. An increase by 1% in past BMI leads now to an increase in current BMI by 0.94%. Whereas an increase by 1% in average BMI leads to an increase by 0.19% in current BMI. We therefore opt to include the health dummy variable among the set of regressors since the significance of lagged individual BMI may be simply pick up an omitted variable problem. Note that the potential endogeneity of the health variables is controlled by the use of lagged endogenous instruments in the system GMM estimation.

0.44). In Table 6 (Model 1) we might erroneously understate the effect exerted by peers if BMI missingness in not properly addressed.

2.4.6 The Role of Habituation and Imitation in Obesity Behavior

In this section we want to assess how habituation and imitative behavior influence the behavior of adults who were normal-weight, overweight and obese adolescents.[26]

We note from Table 2 that BMI is highly autocorrelated, especially for normal-weight and heavily obese people. In this instance, personal history and personality traits may dominate upon the influence of the reference group. We therefore estimate model (2.9) for each BMI category, paying attention to endogenous sample selection arising from selecting categories based on the dependent variable. Hence, we split the sample according to the BMI status in wave 1 and keep individuals in the same strata. This allows us to clearly understand how behaviors during adolescence contribute to adult outcomes. The results for normal, overweight and obese adolescents are reported in Tables 7, 8 and 9 respectively.

We find that personal history (lagged BMI) does not matter for individuals who were normal and overweight adolescents; rather, they seem to be affected mainly by their reference group behavior. For those who were obese in adolescence, instead, habituation is certainly a fundamental driver of current BMI, though the effect exerted by social ties is explosive. In practice, BMI status during adulthood is due to both past behavior and group behavior for individuals who experienced obesity when adolescent - therefore they are obese adults (Whitlock et al., 2005)- but peer effects outnumber habituation effects.

Table 7, Model 3 (benchmark model) shows that individual behavior is dominated by the influence from peers for the sample of normal-weight individuals when teenagers. In this instance, for any 1% increase in average BMI we expect about 0.37% increase in individual BMI, whereas the coefficient for past BMI is not statistically significant. Results show that normal-weight adolescents tend to develop a social behavioral pattern

[26] Specifically, we focus on the International Classification of Weight according to the WHO, as reported in Table 1.

in eating, perhaps related to social inclusion (e.g., Falkner at al, 2001; Chen and Brown, 2005). Table 8 shows the same behavior for individuals who were overweight during their adolescence: a 1% increase in average BMI leads to an increase in individual BMI by 0.67% (Model 3). Interestingly, gender plays a role in explaining eating behavioral patterns of normal-weight adolescents, registering a negative relationship with individual BMI, while the coefficient for the gender dummy is insignificant for overweight adolescents. Ethnicity, instead, does not seem a decisive driver of differences in BMI status.

For obese adolescents the story is different. A pattern of self-weight-manteinance behavior is observed, possibly supported by patterns of wrong behavioral routines such as unhealthy eating habits and scarce exercise. However, the influence of peers at the school level is now stronger, even explosive. Table 9 (Model 3) reports that an increase by 1% in average BMI leads to an increase by 1.32% in current BMI. The habituation effect is lower - still very high in absolute terms - leading to a 0.97% increase in BMI for a 1% increase in aggregate BMI. This means that obese adolescents become future obese adults through wrong habits enforced by imitative behavior. As stressed by Christakis and Fowler (2007), having obese school contacts might change a person's tolerance for being obese or might influence his or her adoption of specific behaviors (e.g., smoking, eating, and exercising). The fact that adolescents' appearance and behaviors are influenced by the appearance and behaviors of those around them suggests that weight gain in one person might influence weight gain in others. In addition to such strictly social effects, it is plausible that physiological imitation might occur (Fogassi et al., 2005); areas of the brain that correspond to actions such as eating food may be stimulated if these actions are observed in others and this possibility is higher within restricted social environments, such as schools, where individuals spend most of their time. Moreover, the positive effect of the gender dummy on BMI of obese individuals could signal that efforts to prevent obesity should not ignore the central role of cognitive factors, as often obese young women lack motivation, and personality traits may dominate over external factors (Andajani-Sutjahjo et al., 2004).

Finally, it is also important to stress that dealing with missingness of both income and BMI has a dramatic effect on the coefficient of average BMI. In Table 7 (Model 1) we might erroneously think the habituation effect to be explosive and understate the effect exerted by peers if BMI missingness in not properly addressed. Table 8 shows that for overweight adolescents, when missing data for income and BMI (Model 3) are replaced by multiple imputation there is an increase by 33% in the coefficient of the group effect measured by average BMI (from 0.501 to 0.667). Table 9 (Model 1) shows that for obese individuals the peer effect would disappear when missingness is not taken into account, while it represents a key factor in delineating eating behavioral patterns for obese adolescents. In addition the results also show that the magnitude of lagged individual BMI in Models 1 and 2 of all tables may be misleading, since it may be simply picking up the effects of missing data. This evidence is generally not captured by a typical full sample estimation (Table 6) where we can observe that the coefficients of average BMI is rather stable across the three models.

2.5 Conclusions

Personal and family history, the impact of the social context where each individual lives as well as endogenous effects induced by interactions with peer groups are all possible determinants of eating behavior. One of the econometric challenges is to identify the separate impact of the endogenous and contextual effects, and to break the so called reflection problem (Manski, 1993). The dynamic linear-in-means model proposed in the paper allows us to estimate all social effects and to control for individual- and group-specific unobservable effects, by exploiting stationarity restrictions of a system GMM estimator. Our results show that individuals tend to become overweight mainly due to habituation and social effects, even properly accounting for contextual effects. In particular, imitative behavior seems to explain a relevant part of variation in body mass index of all individuals in the sample, though habituation plays the most important role. Individuals who were normal-weight and overweight during adolescence are not influenced by past behavior; rather, they are affected by average behavior in their reference group. Obese adolescents, instead, become future obese adults, showing a high persistence in their body mass index status; their wrong habits are enforced by imitative behavior, as peer effects impact dramatically on current weight status.

Such peer effects are of obvious policy significance, though rarely taken into account by policy makers or even by entities with a collective perspective. The implication is that group-level interventions may be more successful and more efficient than individual interventions, as a social multiplier effect takes place. This means that clinical and policy interventions may be more cost-effective than policy-makers have previously supposed. We are facing a health problem characterized by a imitative, therefore multiplicative, dimension. Public health policy makers should implement urgent and targeted actions for preventing this epidemic to spread further.

Despite the advantage of being able to identify peer groups at the level of schoolmates, data used in our study have some limitations. First, the longitudinal analysis is conducted in a panel dataset with non-homogeneous gaps relative to adolescents who become adults. Second, friends outside of school and romantic partners are not captured. Future research is needed to better understand the mechanisms behind the influence of peers on weight. Several candidates exist, such as peer influence on weight loss attempts, physical activity, and perceptions of own body weight. The causal mechanisms should also be consistent with effects of higher moment of the BMI distribution. Knowledge about the source of peer influence on weight and the size of any social multipliers will improve implementation and evaluation of policies aimed at reducing overweight or obesity in adolescence, hence in adulthood.

Bibliography

[1] Anderson, T.W., and Hsiao, C. (1982). "Formulation and estimation of dynamic models using panel data," *Journal of Econometrics*, 18: 47–82.

[2] Andajani-Sutjahjo, S., Ball K., Warren, N., Inglis V. and Crawford D. (2004). "Perceived personal, social and environmental barriers to weight maintenance among young women: A community survey," *International Journal of Behavioral Nutrition and Physical Activity* $1-15$.

[3] Arellano, M., and Bond, S. (1991). "Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations," *Review of Economic Studies* 58: 277–97.

[4] Arellano, M., and Bond, S. (1998). *Dynamic Panel data estimation using DPD*98 *for Gauss: A guide for users.*

[5] Arellano, M., and Bover, O. (1995). "Another look at the instrumental variables estimation of errorcomponents models," *Journal of Econometrics* 68: 29–51.

[6] Baum, C., Schaffer, M., and Stillman, S. (2003). "Instrumental variables and GMM: Estimation and testing," *Stata Journal* 3(1): $1-31$.

[7] Beck, T., and Levine, R. (2004). "Stock markets, banks, and growth: Panel evidence," *Journal of Banking and Finance* 28(3): 423–42.

[8] Bénabou, R. (1993). "Workings of a city: Location, education, and production," *Quarterly Journal of Economics*, CVIII, 619–652.

[9] Blundell, R., and Bond, S. (1998). "Initial conditions and moment restrictions in dynamic panel data models," *Journal of Econometrics* 87: 11–143.

[10] Bond, S., Hoeffler, A., and Temple, J. (2001a). "GMM estimation of empirical growth models," Economics Papers 2001−W21, Economics Group, Nuffield College, University of Oxford.

[11] Bond, S., Bowsher, C., and Windmeijer, F. (2001b). "Criterion-based inference for GMM in autoregressive panel data models," *Economics Letters*, Elsevier, 73(3): 379 − 388, December.

[12] Bond, S. (2002). "Dynamic panel data models: A guide to micro data methods and practice," Working Paper 09/02. Institute for Fiscal Studies. London.

[13] Brock, W. A., and Durlauf, S. N. (2001). "Discrete choice with social interactions," *Review of Economic Studies*, 68(2), 235 − 260.

[14] Brown, B. (1990). "Peer groups and peer cultures," in *At the Threshold*, ed. by S. Feldman, and G. Elliott, Cambridge, MA. Harvard University Press.

[15] Brown, B., Clasen, D., and Eicher, S. (1986). "Perceptions of peer pressure, peer conformity dispositions, and self-reported behavior among adolescents," *Developmental Psychology*, 22(4): 521 − 530

[16] Bernheim, D. (1994). "A theory of conformity," *Journal of Political Economy*, 102: 841 − 877.

[17] Bramoullé, Y., Djebbari, H., and Fortin, B. (2009). "Identification of peer effects through social networks," *Journal of Econometrics*, 150(1): 41–55.

[18] Burke, M. A., and Heiland, F. (2007). "Social dynamics of obesity," *Economic Inquiry*, 45, 571–591.

[19] Case, A., M. Lubotsky and C. Paxson (2002). "Economic Status and Health in Childhood: The Origins of the Gradient," *American Economic Review*, 92: 1308 − 1334.

[20] Chamberlain, G. (1982). *Multivariate Regression Model for Panel Data*, Journal of Econometrics, $18, 5 - 46$.

[21] Chamberlain, G. (1984). "Panel Data" in *Handbook of Econometrics* ed. by Z. Griliches and M.D. Intriligator, Amsterdam, North Holland Publishing Co, pp. $1247 - 1318$.

[22] Chen, E.Y. and Brown, M. (2005). "Obesity stigma in sexual relationships," *Obesity Research*, 13(8): $1393 - 7$.

[23] Christakis N. A., and Fowler, J. (2007). "The spread of obesity in a large social network over 32 years," *The New England Journal of Medicine*, 357: $370 - 379$.

[24] Cohen-Cole, E. (2006). "Multiple groups identification in the linear-in-means model," *Economic Letters*, 92(2): 157–162.

[25] Cohen-Cole, E., and Fletcher, J. (2008). "Is obesity contagious? Social networks VS. environmental factors in the obesity epidemic," *Journal of Health Economics*, 27(5): $1382 - 1387$.

[26] Corrado, L. and B. Fingleton (2012). "Where is the Economics in Spatial Econometrics?," *Journal of Regional Science*, Volume 51, Issue 2, pp. $1 - 30$.

[27] Crawford, D. and Ball, K. (2002). "Behavioural determinants of the obesity epidemic," *Asia Pacific Journal of Clinical Nutrition*, 11(Suppl 8): S718 $-$ 721.

[28] De Giorgi, G., Pellizzari, M., and Redaelli, S. (2010). "Identification of social interactions through partially overlapping peer groups," *American Economic Journal: Applied Economics*, 2(2): $241 - 75$.

[29] Falkner N.H., Neumark-Sztainer, D., Story, M., Jeffery, R.W., Beuhring, T. and Resnick, M.D. (2001). "Social, educational, and psychological correlates of weight status in adolescents," *Obesity Research*, 9(1):$32 - 42$.

[30] Fogassi, L., Ferrari, P.F., Gesierich. B., Rozzi, S., Chersi, F., Rizzolatti, G. (2005). "Parietal lobe: from action organization to intention understanding," *Science*, 308: 662 – 7.

[31] Fortin, B. and Yazbeck M. (2011). "Peer Effects, Fast Food Consumption and Adolescent Weight Gain," Centre Interuniversitaire sur le Risque, les Politiques Èconomiques et l'Emploi, Cahier de recherche/Working Paper 11 – 03.

[32] Fowler, J., and Christakis, N. A. (2008). "Estimating peer effects on health in social networks; A response to Cohen-Cole and Fletcher; and Trogdon, Nonnemaker, and Pail," *Journal of Health Economics*, 27, 1400 – 1405.

[33] Goodman, E., Hinden, B. R., and Khandelwal, S. (2000). "Accuracy of teen and parental reports of obesity and body mass index," *Pediatrics*, 106: 52 – 58.

[34] Gordon-Larsen, P., Adair, L.S., Nelson, M.C., and Popkin, B.M. (2004). "Five-year obesity incidence in the transition period between adolescence and adulthood: the National Longitudinal Study of Adolescent Health," *American Journal of Clinical Nutrition*, 80(3): 569 – 575.

[35] Gordon-Larsen, P., Nelson M.C., Page P., et al. (2006). "Inequality in the built environment underlies key health disparities in physical activity and obesity," *Pediatrics*, 117: 417 – 24.

[36] Harris, K.M., Halpern, C.T., Whitsel, E., Hussey, J., Tabor, J., Entzel, P. and Udry, J.R. (2009). "The National Longitudinal Study of Adolescent Health: Research Design" website: http://www.cpc.unc.edu/projects/addhealth/design.

[37] Harris, K. M. (2011). "Design Features of Add Health," Carolina Population Center. website: www.cpc.unc.edu/projects/addhealth/data/guides/.

[38] Kandori, M. (1992). "Social norms and community enforcement," *Review of Economic Studies*, 59: 63 – 80.

[39] Kemper, H.C.G., Post, G.B., Twisk, J.W.R., and van Mechelen, W. (1999). "Lifestyle and obesity in adolescence and young adulthood: Results from the Amsterdam Growth And Health Longitudinal Study (AGAHLS)," *International Journal of Obesity*, 23(3): s34−s40.

[40] Kvaavik, E., Tell, G.S., and Klepp, K. (2003). "Predictors and Tracking of Body Mass Index From Adolescence Into Adulthood. Follow-up of 18 to 20 Years in the Oslo Youth Study," *Arch Pediatr Adolesc Med*, 157:1212 − 1218.

[41] Lewbel, A. (1997). "Constructing Instruments for Regressions with Measurement Error when no Additional Data are Available with an Application to Patents and RD," *Econometrica*, 65(5): 1201 − 1214.

[42] Lindeboom M., Nozal A. L. and van der Klaauw, B. (2009). "Parental education and child health: Evidence from a schooling reform", *Journal of Health Economics*, 28(1): 109 − 131.

[43] Liu, X., Lee, L.F., Kagel, J. (2006). "Dynamic Discrete Choice Models with Lagged Social Interactions: With an Application to a Signaling Game Experiment," Ohio State; Mimeo.

[44] Manski, C. F. (1993). "Identification of endogenous social effects: The reflection problem," *Review of Economic Studies*, 60(3): 531 − 542.

[45] Manski, C. F.(2000). "Economic analysis of social interactions," *Journal of Economic Perspectives*, 14(3): 115 − 136.

[46] Moffitt, R. A. (2001). "Policy interventions, low-level equilibria, and social interactions" in S. N. Durlauf and H. P. Young (Eds.), *Social dynamics*, pp. 45–82. Cambridge, MA: MIT Press.

[47] Gordon-Larsen P., Adair L.S., and Popkin, B.M. (2003). "The relationship of ethnicity, socioeconomic factors, and overweight in US adolescents." *Obesity Research*, 11: 121 − 9.

[48] Popkin, B.M. and Udry, J.R. (1998). "Adolescent obesity increases significantly in second and third generation U.S. immigrants: the National Longitudinal Study of Adolescent Health," *Journal of Nutrition*, 128: 701 – 6.

[49] Sterne J., White I., Carlin J., Spratt M., Royston P., Kenward M., Wood A., Carpenter J. (2009). "Multiple Imputation for Missing Data in Epidemiological and Clinical Research: Potential and Pitfalls," *British Medical Journal*, 338, *b*2393.

[50] Sun Guo, S., Wu, W., Chumlea, W.C., and Roche, A.F. (2002). "Predicting overweight and obesity in adulthood from body mass index values in childhood and adolescence," *American Journal of Clinical Nutrition*, 76(3):653 – 658.

[51] Trogdon, J., Nonnemaker, J., and Pail, J. (2008). "Peer effects in adolescent overweight," *Journal of Health Economics*, 27, 1388–1399.

[52] Tourangeau, R. and Shin, H. (1999). "National Longitudinal Study of Adolescent health: Grand Sample Weights," National Opinion Research Center and Carolina Population Center. website: http://www.cpc.unc.edu/projects/addhlth/.

[53] Whitlock E.P., Williams, S.B., Gold, R., Smith, P.R. and Shipman, S.A. (2005). "Screening and interventions for childhood overweight: a summary of evidence for the US Preventive Services Task Force," *Pediatrics*, 116(1).

A GMM and Chamberlain's Correlated Effects Approach in Linear Panel Data Models

Consider system (2.9)

$$
\begin{aligned}
y_{t,ig} &= y_{t-1,ig}\varphi + y^{e}_{t,ig}\beta + \mathbf{x}'_{t,g}\boldsymbol{\gamma} + \mathbf{r}'_{t,ig}\boldsymbol{\delta} + e_{t,ig},\ |\varphi| < 1 \qquad (2.15)\\
e_{t,ig} &= \alpha_g + u_{t,ig},\\
u_{t,ig} &= f_i + \varepsilon_{t,ig}
\end{aligned}
$$

By recursion we can write (for $t = 1, ..., T$) :

$$
\begin{aligned}
y_{t,ig} = {} & \left(1 + \varphi + ... + \varphi^{t-1}\right) f_i + \left(1 + \varphi + ... + \varphi^{t-1}\right) \alpha_g + \varphi^t y_{0,ig} + \qquad (2.16)\\
& +[y^{e}_{t,ig}\beta + y^{e}_{t-1,ig}\beta\varphi + ... + y^{e}_{1,ig}\beta\varphi^{t-1}]\\
& +[\mathbf{x}'_{t,g}\boldsymbol{\gamma} + \mathbf{x}'_{t-1,g}\boldsymbol{\gamma}\varphi + ... + \mathbf{x}'_{1,g}\boldsymbol{\gamma}\varphi^{t-1}] +\\
& +[\mathbf{r}'_{t,ig}\boldsymbol{\delta} + \mathbf{r}'_{t-1,ig}\boldsymbol{\delta}\varphi + ... + \mathbf{r}'_{1,ig}\boldsymbol{\delta}\varphi^{t-1}] +\\
& +[\varepsilon_{t,ig} + \varphi\varepsilon_{t-1,ig} + ... + \varphi^{t-1}\varepsilon_{1,ig}]
\end{aligned}
$$

This transformation links system (2.9) to Chamberlain's method (1982, 1984) to deal with correlated effects in dynamic linear panel data models. In fact, we can write system (2.16) in compact form as:

$$E[y_{t,ig} \mid \mathbf{W}_i] = \mathbf{W}_i'\, \boldsymbol{\Pi} + \eta(f_i \ + \alpha_g)$$

where $\mathbf{W}_i = \left[y_{0,ig},\ y^{e}_{t,ig}, ..., y^{e}_{1,ig},\ \mathbf{x}_{t,g}, ..., \mathbf{x}_{1,g},\ \mathbf{r}_{t,ig}, ..., \mathbf{r}_{1,ig}\right]$. The $\boldsymbol{\Pi}$ matrix is defined in terms of the coefficients of the linear predictors of the dependent variable at each period given all explanatory variables at all periods. For the individual effect, f_i, and the group effect, α_g, we therefore have:

$$E[f_i, \alpha_g \mid \mathbf{W}_i] = 0$$

Given the equivalence with system (2.16) both f_i and α_g can therefore be treated as random individual- and group-specific effects also in the original system (2.9).

B Design and Weighting

The Add Health Study is a US representative, probability-based survey of adolescents in grades 7 through 12 conducted between 1994 and 1995, and extended to 2008 with three in-home interviews. The sample design used to collect the data embeds a certain degree of complexity which should be accounted for. Indeed, failing at considering such complexity may result in biased parameter estimates and incorrect variance estimates. Hence, we corrected for design effects and unequal probability of selection, according to what is suggested in the Add Health user guides.[27]

We exploit the longitudinal feature of the dataset, keeping the strength of its innovative design. With the longitudinal data from adolescence, the third and four in-home interviews allow "researchers to map early trajectories out of adolescence in health, achievement, social relationships, and economic status and to document how adolescent experiences and behaviors are related to decisions, behavior, and health outcomes in the transition to adulthood. The fundamental purpose of this [...] follow-up was to understand how what happens in adolescence is linked to what happens in the transition to adulthood when adolescents begin to negotiate the social world on their own and develop their expectations and goals for their future adult roles." (Harris, 2011). Data have been appropriately weighted to correct for time gaps in their longitudinal format. For details on the Add Health weighting scheme, the reader is cross-referred to Tourangeau and Shin (1999).

[27] http://www.cpc.unc.edu/projects/addhealth/data/guides

C Missing data

There are several reasons why the data may be missing. We say that data are "missing completely at random" if the probability that an observation is missing is not related to the value of that observation or to the value of any other variable. In this case the design power is lower, but the estimated parameters are not biased. However, this data feature is not very common.

In other cases data may be classified as "missing at random". For data to be missing at random, missingness should not depend on the value of the missing observation after controlling for another variable. The type of missingness should be dealt with in order to produce relatively unbiased estimates. Both these types of missingness are said to be "ignorable", but the latter needs to be addressed in some way.

Finally, we could have "missing not at random" data, i.e. data for which missingness depends on the value of the missing observation. Under such circumstances, the only way to obtain unbiased estimates is to write a model that takes missing data in due account. Clearly, this could be a rather difficult task as we rarely know what the missingness model is.

Concerning our case, we consider the variables of the dataset to display missing-at-random or missing-completely-at-random values. For example, on the one hand, income self-reported information might be dependent on individual ethnicity: black people could be less likely to report their income than white individuals. The black probably have lower incomes than the white, and it would at first appear that missingness on income is related to the value of income itself. But the data would still be missing at random if the conditional probability of missingness were unrelated to the value of income within each ethnic group. On the other hand, missing values on gender, for example, could be considered as being missing completely at random.

We decide to deal with missingness by applying a Multiple Imputation method to two variables of interest, namely household income and BMI (Estimation Tables: Model

2, income only imputed; Model 3 income and BMI imputed).[28] Multiple Imputation, involves estimating what the missing values would be, and then using those "imputed values" in the solution. Obviously, in this case we have selected the variables for which it could make sense to expect missingness at random to be explained by other variables included in our model. Income is the variable showing the highest number of missing values; moreover, for the reasoning just explained, we believe that such a missingness is not completely at random. Therefore, we perform a multiple imputation on income only first, and on income together with BMI later. Indeed, body weight is another variable showing a high level of missing cases. Therefore, we decide to impute BMI as it is our dependent variable, since peer effects are directly generated from it.

We make use of Multiple Imputation by Chained Equations (MICE)(Sterne et al., 2009). For a set of variables, $x_1; \ldots; x_k$ some or all of which have missing values, the MICE algorithm initially fills all missing values at random. The first variable with at least one missing value, e.g., x_1, is then regressed on the other variables, $x_2; \ldots; x_k$. The estimation is restricted to individuals with observed x_1. Missing values in x_1 are replaced by simulated draws from the posterior predictive distribution of x_1, an important step known as proper imputation. The next variable with missing values, say x_2, is regressed on all the other variables, $x_1; x_3; \ldots x_k$. Estimation is restricted to individuals with observed x_2 and uses the imputed values of x_1. Again, missing values in x_2 are replaced by draws from the posterior predictive distribution of x_2. The process is repeated for all other variables with missing values in turn (cycle), for about ten cycles.[29] The entire procedure is repeated independently M times, yielding M imputed datasets.[30]

[28] Our criteria for choosing to deal with missingness are, first the (potentially reasonable) source of missingness -i.e., completely random or random-, second the percentage of missing cases – i.e., if greater than 20%-, and last the importance of such missing information for the model (e.g., our dependent variable). The percentage of missing values per variable are reported in Table 3.

[29] Because each variable is imputed using its own imputation model, MICE can handle different variable types (for example, continuous, binary, unordered categorical, ordered categorical).

[30] Standard texts on MI suggest that small numbers of imputed datasets ($M = 3$ to 5) are adequate.

Table 1: International Classification of Weight According to BMI

Classification	Category	BMI (Principal Cut-off Points)
Underweight	1	<18.50
Normal Range	2	18.50-24.99
Overweight (Pre-Obese)	3	25.00-29.99
Obese Class I	4	30.00-34.99
Obese Class II	5	35.00-39.99
Obese Class III	6	≥40

Note. Source: Adapted from WHO, 1995, WHO, 2000 and WHO 2004.

Table 2: Transition matrix of BMI categories

		BMI in t						
		1	2	3	4	5	6	Total
BMI in $t-1$	1	45.09	52.32	2.18	0.24	0.12	0.06	100
	2	2.31	74.22	20.15	2.91	0.32	0.09	100
	3	0.31	13.55	55.74	24.96	4.64	1.07	100
	4	0.14	2.06	15.64	48.29	24.97	8.92	100
	5	0.00	0.81	5.81	19.92	37.80	35.57	100
	6	0.29	0.00	0.29	4.34	15.03	80.06	100
	Total	5.43	51.20	24.76	11.03	4.28	3.31	100

Note. Categories defined as in Table 1.

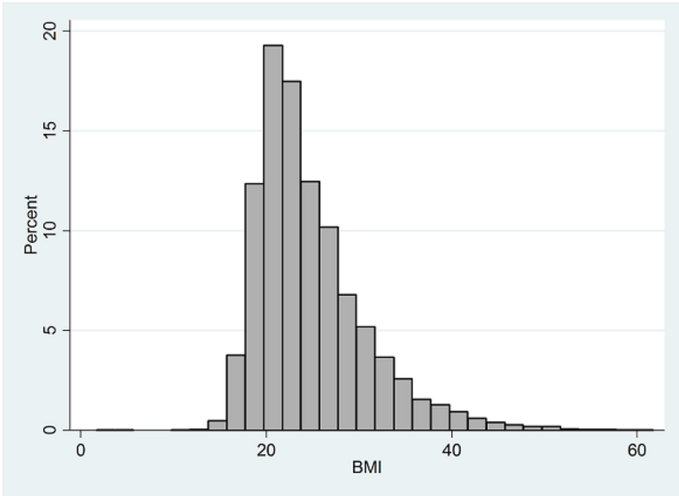

Figure 1: BMI distribution. Bin width: 2 units.

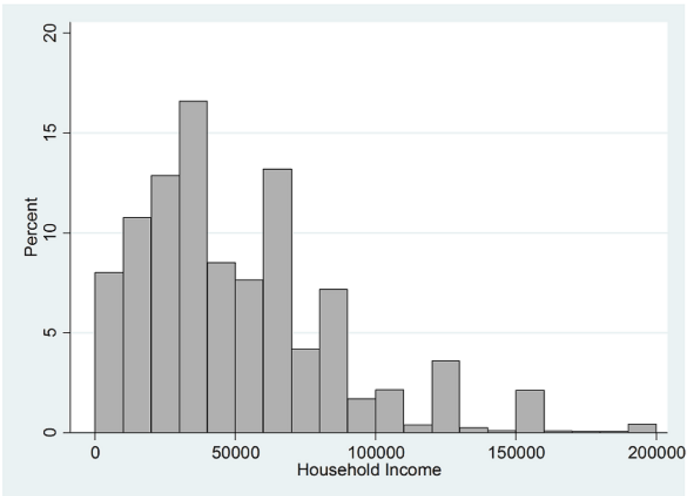

Figure 2: Household income distribution. Bin width: 10,000 US Dollars.

Table 3: Summary Statistics (Waves I-IV, total sample)

Variable	Definition	Mean	SD	Missing Values
BMI	weight (kg)/height^2 (m)	24.74	5.90	19%
Household Income	Total income before taxes	52,320.90	50,626.62	39%
Median of Household Income	40,000			
1st percentile	70			
99th percentile	200,000			
Ethnicity (%)	Race as observed by interviewer			11%
White		60.84	48.81	
Black	African American	21.94	41.39	
American Indian	Native American.	1.05	10.18	
Asian		3.24	17.71	
Other		2.31	15.02	
Self-reported Health (%)				17%
Excellent		22.80	41.95	
Very Good		33.04	47.03	
Good		21.48	41.07	
Fair		5.19	22.18	
Poor		0.44	6.62	
Age	Age from birth	20.70	5.53	17%
Gender (%)	Proportion of females	57.09	49.50	20%
Mother Education (%)	Proportion of college graduate (or certified 4 years in college)	9.63	29.50	18%
Father Education (%)	Proportion of college graduate (or certified 4 years in college)	9.45	29.95	18%
Urban	Completely urban city	28.51	45.15	0%
Unemployment Rate	Low unemployment rate (vs Medium and High)	29.24	45.49	46%

Table 4: Average BMI by Subsample (Waves I-IV, total sample)

Variable	Mean	SD
BMI if Male	24.44	4.98
BMI if Female	23.95	4.99
BMI if White	24.70	5.75
BMI if Black	25.86	6.57
BMI if American Indian	29.37	8.91
BMI if Asian	23.28	4.85
BMI if Other	23.53	4.75
BMI if Mother went to college	22.58	4.70
BMI if Father went to college	22.45	4.60
BMI if Excellent health	23.31	4.61
BMI if Very good health	24.32	5.46
BMI if Good health	26.41	6.83
BMI if Fair health	28.31	8.06
BMI if Poor health	30.07	10.20
BMI if Living in completely urban city	22.95	4.76
BMI if Living in not completely urban city	22.67	4.73
BMI if Low unemployment rate	24.78	5.71
BMI if Medium-High unemployment rate	24.77	5.97
BMI if Household income>Median	24.39	5.50
BMI if Household income<Median	25.23	6.31

Table 5: Average BMI by Subsample (Waves I-II, adolescent sample)

Variable	Mean	SD
BMI	22.73	4.73
BMI if Male	22.57	4.60
BMI if Female	22.82	4.81
BMI if White	22.43	4.51
BMI if Black	23.43	5.13
BMI if American Indian	26.37	7.47
BMI if Asian	21.56	3.81
BMI if Other	23.28	4.65
BMI if Mother went to college	22.17	4.28
BMI if Father went to college	21.80	4.02
BMI if Excellent health	21.71	3.72
BMI if Very good health	22.35	4.24
BMI if Good health	23.78	5.35
BMI if Fair health	25.64	6.82
BMI if Poor health	26.92	8.55
BMI if Living in completely urban city	22.95	4.76
BMI if Living in not completely urban city	22.67	4.73
BMI if Low unemployment rate	22.53	4.46
BMI if Medium-High unemployment rate	22.78	4.81
BMI if Household income>Median	22.38	4.39
BMI if Household income<Median	23.14	5.05
BMI if Smoked at least 1 cigarette daily for 30 days	24.26	5.41
BMI if Smokers in household	22.98	4.93
BMI if Own decision in diet	22.90	4.73
BMI if Have dinner with parents frequently	23.03	4.63

Table 6: Estimates using full sample

Dependent variable: ln(BMI)						
Variables	Model 1		Model 2		Model 3	
	Coefficient	SE	Coefficient	SE	Coefficient	SE
$\ln(\text{BMI})_{t-1}$	0.981^{***}	(0.214)	0.675^{***}	(0.165)	0.839^{***}	(0.154)
Average ln(BMI)	0.259^{*}	(0.151)	0.469^{***}	(0.091)	0.440^{***}	(0.084)
Household Income	0.003^{*}	(0.002)	0.001	(0.002)	0.000	(0.002)
Age	−0.003	(0.003)	0.001	(0.003)	0.000	(0.003)
Woman	−0.005	(0.004)	−0.002	(0.003)	0.004	(0.004)
Ethnic Group: White	−0.019	(0.046)	-0.060^{**}	(0.028)	-0.094^{**}	(0.047)
Ethnic Group: Black	−0.037	(0.049)	−0.047	(0.029)	-0.082^{*}	(0.048)
Ethnic Group: American Indian	−0.032	(0.049)	-0.074^{*}	(0.038)	-0.108^{**}	(0.055)
Ethnic Group: Asian	−0.012	(0.046)	-0.078^{***}	(0.029)	-0.102^{**}	(0.047)
Ethnic Group: Other	−0.029	(0.034)	-0.060^{**}	(0.030)	−0.081	(0.050)
Maternal Education (college)	0.005	(0.011)	−0.004	(0.010)	0.001	(0.009)
Paternal Education (college)	−0.009	(0.006)	-0.011^{*}	(0.006)	-0.011^{**}	(0.005)
Health Status (Excellent to Fair)	−0.019	(0.012)	-0.031^{***}	(0.007)	-0.025^{***}	(0.007)
First-2-waves Dummy	−0.006	(0.006)	−0.004	(0.007)	0.001	(0.007)
Average Household Income	0.003	(0.005)	0.000	(0.003)	0.002	(0.003)
Average Age	−0.001	(0.017)	-0.007^{***}	(0.002)	-0.009^{***}	(0.002)
Average presence of women	0.001	(0.017)	−0.003	(0.016)	0.006	(0.018)
Ethnic Group: Average White	−0.220	(0.320)	−0.180	(0.273)	−0.211	(0.301)
Ethnic Group: Average Black	−0.238	(0.321)	−0.202	(0.272)	−0.241	(0.301)
Ethnic Group: Average American Indian	−0.183	(0.320)	−0.188	(0.274)	−0.245	(0.301)
Ethnic Group: Average Asian	−0.200	(0.325)	−0.158	(0.274)	−0.188	(0.302)
Ethnic Group: Average Other	−0.216	(0.322)	−0.153	(0.273)	−0.165	(0.304)
Average Maternal Education (college)	-0.140^{**}	(0.059)	−0.064	(0.063)	−0.070	(0.057)
Average Paternal Education (college)	0.041	(0.042)	0.040	(0.042)	0.038	(0.041)
Average Health Status (Excellent to Fair)	0.038^{*}	(0.021)	0.035^{**}	(0.017)	0.054^{***}	(0.020)
Observations	6,598		10,700		10,677	
Number of individuals	4,095		4,655		4,646	
Number of Instruments	31		31		31	
Arellano-Bond Test for AR(1) in first Differences: $\Pr > z$	$9.16e^{-07}$		$2.06e^{-08}$		0	
Arellano-Bond Test for AR(1) in first Differences: z	−4.909		−5.607		−6.665	
Hansen test of overid. restrictions: $\Pr > \chi^2$	0.298		0.888		0.341	
Hansen test of overid. restrictions: Degrees of Freedom	3		3		3	
Hansen test of overid. restrictions: χ^2	3.684		0.635		3.349	
Sargan test of overid. restrictions: $\Pr > \chi^2$	0.308		0.683		0.386	
Sargan test of overid. restrictions: Degrees of Freedom	3		3		3	
Sargan test of overid. restrictions: χ^2	3.600		0.683		3.034	

Standard errors in parentheses:.$^{***}p < 0.01,^{**}p < 0.05,^{*}p < 0.1$

NOTES. Models differ because of missing values treatment. Model 1: Missing values are not treated; Model 2: income missing values replaced by multiple-imputed values; Model 3: BMI and income missing values replaced by multiple-imputed values.

For further details on missing values treatment see Appendix A, section A.1.2.

Our estimation also accounts of wave effects, through the inclusion of time and adolescent dummies.

Table 7: Estimates using sample of individuals who are normal-weight during adolescence

Dependent variable: ln(BMI)

Variables	Model 1		Model 2		Model 3	
	Coefficient	SE	Coefficient	SE	Coefficient	SE
$\ln(\text{BMI})_{t-1}$	1.269***	(0.446)	1.183***	(0.429)	0.588	(0.673)
Average ln(BMI)	0.194*	(0.115)	0.269***	(0.079)	0.369***	(0.106)
Household Income	0.004	(0.003)	0.002	(0.002)	−0.002	(0.002)
Age	−0.002	(0.002)	−0.002	(0.002)	−0.000	(0.003)
Woman	−0.008	(0.006)	−0.010**	(0.005)	−0.006*	(0.004)
Ethnic Group: White	−0.036	(0.066)	−0.055	(0.040)	−0.059**	(0.026)
Ethnic Group: Black	−0.023	(0.066)	−0.051	(0.040)	−0.049*	(0.027)
Ethnic Group: American Indian	−0.042	(0.071)	−0.058	(0.056)	−0.048	(0.044)
Ethnic Group: Asian	−0.040	(0.067)	−0.045	(0.045)	−0.062**	(0.030)
Ethnic Group: Other	−0.033	(0.067)	−0.093**	(0.042)	−0.058*	(0.030)
Maternal Education (college)	−0.004	(0.012)	−0.001	(0.010)	−0.014	(0.011)
Paternal Education (college)	−0.008	(0.008)	−0.004	(0.006)	−0.008	(0.010)
Health Status (Excellent to Fair)	−0.020**	(0.010)	−0.018**	(0.008)	−0.029***	(0.008)
First-2-waves Dummy	0.001	(0.010)	−0.003	(0.009)	−0.006	(0.008)
Average Household Income	−0.002	(0.007)	−0.003	(0.004)	0.001	(0.004)
Average Age	−0.005*	(0.003)	−0.009***	(0.002)	−0.008***	(0.002)
Average presence of women	0.035	(0.026)	0.026	(0.023)	−0.001	(0.020)
Ethnic Group: Average White	−0.356	(0.438)	−0.353	(0.468)	0.001	(0.357)
Ethnic Group: Average Black	−0.379	(0.440)	−0.374	(0.470)	−0.017	(0.357)
Ethnic Group: Average American Indian	−0.287	(0.436)	−0.336	(0.469)	−0.010	(0.351)
Ethnic Group: Average Asian	−0.376	(0.447)	−0.366	(0.472)	0.012	(0.364)
Ethnic Group: Average Other	−0.356	(0.446)	−0.271	(0.473)	0.076	(0.356)
Average Maternal Education (college)	−0.071	(0.085)	−0.042	(0.088)	−0.074	(0.129)
Average Paternal Education (college)	0.046	(0.055)	−0.002	(0.047)	0.010	(0.061)
Average Health Status (Excellent to Fair)	0.043*	(0.026)	0.019	(0.024)	0.047**	(0.021)
Observations	4,183		6,591		6,620	
Number of individuals	2,581		2,867		2,881	
Number of instruments	31		31		31	
Arellano-Bond Test for AR(1) in first Differences: $\Pr > z$	0.00369		0.00186		0.222	
Arellano-Bond Test for AR(1) in first Differences: z	−2.904		−3.111		−1.221	
Hansen test of overid. restrictions: $\Pr > \chi^2$	0.754		0.900		0.737	
Hansen test of overid. restrictions: Degrees of Freedom	3		3		3	
Hansen test of overid. restrictions: χ^2	1.195		0.583		1.266	
Sargan test of overid. restrictions: $\Pr > \chi^2$	0.777		0.941		0.702	
Sargan test of overid. restrictions: Degrees of Freedom	3		3		3	
Sargan test of overid. restrictions: χ^2	1.101		0.396		1.417	

Standard errors in parentheses:.$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

NOTES. Sample split according to the BMI status in the first wave (individuals are kept in the same strata).

Models differ because of missing values treatment. Model 1: Missing values are not treated; Model 2: income missing values replaced by multiple-imputed values; Model 3: BMI and income missing values replaced by multiple-imputed values.

For further details on missing values treatment see Appendix A, section A.1.2.

Our estimation also accounts of wave effects, through the inclusion of time and adolescent dummies.

Table 8: Estimates using sample of individuals who are overweight during adolescence

Dependent variable: ln(BMI)

Variables	Model 1		Model 2		Model 3	
	Coefficient	SE	Coefficient	SE	Coefficient	SE
$\ln(\mathrm{BMI})_{t-1}$	0.097	(0.610)	−0.195	(0.923)	−1.281	(1.841)
Average ln(BMI)	0.501**	(0.155)	0.599***	(0.122)	0.667***	(0.196)
Household Income	0.007*	(0.004)	0.005	(0.003)	0.002	(0.010)
Age	0.001	(0.006)	−0.002	(0.006)	0.004	(0.009)
Woman	−0.001	(0.012)	−0.001	(0.018)	−0.015	(0.023)
Ethnic Group: White	0.055	(0.072)	0.040	(0.042)	0.010	(0.064)
Ethnic Group: Black	0.080	(0.063)	0.075*	(0.039)	0.064	(0.058)
Ethnic Group: American Indian	0.038	(0.077)	0.043	(0.050)	−0.016	(0.099)
Ethnic Group: Asian	0.004	(0.081)	0.045	(0.056)	0.047	(0.088)
Ethnic Group: Other	0.082	(0.071)	0.073*	(0.044)	0.023	(0.069)
Maternal Education (college)	0.051	(0.034)	0.039	(0.053)	0.069	(0.072)
Paternal Education (college)	−0.036	(0.036)	−0.042	(0.053)	−0.112	(0.113)
Health Status (Excellent to Fair)	−0.034**	(0.014)	−0.026**	(0.011)	−0.034***	(0.012)
First-2-waves Dummy	−0.003	(0.017)	−0.025	(0.022)	0.003	(0.037)
Average Household Income	0.012	(0.012)	0.005	(0.009)	0.011	(0.014)
Average Age	−0.015**	(0.007)	−0.002	(0.007)	0.001	(0.024)
Average presence of women	−0.054	(0.069)	−0.012	(0.045)	0.023	(0.089)
Ethnic Group: Average White	−0.418	(0.938)	−0.061	(0.654)	0.396	(0.988)
Ethnic Group: Average Black	−0.426	(0.938)	−0.089	(0.652)	0.356	(0.985)
Ethnic Group: Average American Indian	−0.346	(0.944)	−0.084	(0.672)	0.463	(1.058)
Ethnic Group: Average Asian	−0.252	(0.946)	0.021	(0.662)	0.528	(1.007)
Ethnic Group: Average Other	−0.468	(0.936)	−0.141	(0.658)	0.376	(1.002)
Average Maternal Education (college)	−0.209	(0.206)	−0.042	(0.088)	0.117	(0.332)
Average Paternal Education (college)	0.213	(0.246)	−0.002	(0.047)	0.413	(0.392)
Average Health Status (Excellent to Fair)	0.086	(0.057)	0.021	(0.045)	0.047	(0.065)
Observations	942		1,515		1,542	
Number of individuals	577		668		669	
Number of instruments	31		31		31	
Arellano-Bond Test for AR(1) in first Differences: $\Pr > z$	0.256		0.706		0.839	
Arellano-Bond Test for AR(1) in first Differences: z	−1.137		−0.377		0.204	
Hansen test of overid. restrictions: $\Pr > \chi^2$	0.933		0.646		0.997	
Hansen test of overid. restrictions: Degrees of Freedom	3		3		3	
Hansen test of overid. restrictions: χ^2	0.436		1.659		0.0548	
Sargan test of overid. restrictions: $\Pr > \chi^2$	0.413		0.334		0.936	
Sargan test of overid. restrictions: Degrees of Freedom	3		3		3	
Sargan test of overid. restrictions: χ^2	2.864		3.403		0.420	

Standard errors in parentheses:.$^{***}p < 0.01$,$^{**}p < 0.05$,$^{*}p < 0.1$

NOTES. Sample split according to the BMI status in the first wave (individuals are kept in the same strata).

Models differ because of missing values treatment. Model 1: Missing values are not treated; Model 2: income missing values replaced by multiple-imputed values; Model 3: BMI and income missing values replaced by multiple-imputed values.

For further details on missing values treatment see Appendix A, section A.1.2.

Our estimation also accounts of wave effects, through the inclusion of time and adolescent dummies.

Table 9: Estimates using sample of individuals who are obese during adolescence

Dependent variable: ln(BMI)

Variables	Model 1		Model 2		Model 3	
	Coefficient	SE	Coefficient	SE	Coefficient	SE
$\ln(\text{BMI})_{t-1}$	0.664^{*}	(0.399)	1.020^{***}	(0.333)	0.968^{***}	(0.358)
Average ln(BMI)	1.166	(0.828)	1.017	(0.690)	1.321^{**}	(0.611)
Household Income	0.006	(0.011)	−0.001	(0.010)	0.017	(0.013)
Age	-0.013^{*}	(0.007)	−0.006	(0.009)	−0.002	(0.009)
Woman	0.030	(0.020)	0.048^{**}	(0.021)	0.051^{**}	(0.022)
Ethnic Group: White	−3.489	(4.463)	0.404	(4.176)	−1.117	(3.767)
Ethnic Group: Black	−3.486	(4.445)	0.383	(4.174)	−1.116	(3.767)
Ethnic Group: American Indian	−3.483	(4.396)	0.359	(4.140)	−1.139	(3.705)
Ethnic Group: Asian	−3.521	(4.419)	0.224	(4.127)	−1.200	(3.727)
Ethnic Group: Other	−3.493	(4.478)	0.403	(4.186)	−1.132	(3.791)
Maternal Education (college)	−0.059	(0.060)	0.023	(0.031)	0.025	(0.032)
Paternal Education (college)	0.066	(0.044)	0.034	(0.032)	0.013	(0.032)
Health Status (Excellent to Fair)	−0.031	(0.032)	0.016	(0.034)	−0.003	(0.036)
First-2-waves Dummy	−0.043	(0.031)	−0.035	(0.035)	−0.034	(0.038)
Average Household Income	−0.005	(0.019)	−0.002	(0.015)	−0.016	(0.013)
Average Age	−0.002	(0.009)	−0.006	(0.012)	−0.013	(0.012)
Average presence of women	−0.130	(0.095)	−0.067	(0.102)	−0.023	(0.095)
Ethnic Group: Average White	0.902	(2.801)	−3.245	(2.168)	−2.809	(2.564)
Ethnic Group: Average Black	0.883	(2.806)	−3.305	(2.161)	−2.874	(2.568)
Ethnic Group: Average American Indian	0.802	(2.759)	−3.458	(2.134)	−3.019	(2.500)
Ethnic Group: Average Asian	1.032	(3.048)	−3.318	(2.319)	−2.738	(2.730)
Ethnic Group: Average Other	0.935	(2.891)	−3.390	(2.182)	−2.922	(2.626)
Average Maternal Education (college)	−0.166	(0.300)	−0.105	(0.324)	−0.203	(0.295)
Average Paternal Education (college)	0.168	(0.200)	0.133	(0.224)	0.336	(0.215)
Average Health Status (Excellent to Fair)	0.635	(0.980)	−0.311	(0.686)	0.222	(0.746)
Observations	399		650		657	
Number of individuals	247		272		270	
Number of instruments	31		31		31	
Arellano-Bond Test for AR(1) in first Differences: $\Pr > z$	0.0135		0.00627		0.00214	
Arellano-Bond Test for AR(1) in first Differences: z	−2.470		−2.733		−3.070	
Hansen test of overid. restrictions: $\Pr > \chi^2$	0.0589		0.165		0.222	
Hansen test of overid. restrictions: Degrees of Freedom	2		2		2	
Hansen test of overid. restrictions: χ^2	5.664		3.601		3.012	
Sargan test of overid. restrictions: $\Pr > \chi^2$	0.0541		0.325		0.320	
Sargan test of overid. restrictions: Degrees of Freedom	2		2		2	
Sargan test of overid. restrictions: χ^2	5.833		2.250		2.276	

Standard errors in parentheses:.$^{***}p < 0.01, ^{**}p < 0.05, ^{*}p < 0.1$

NOTES. Sample split according to the BMI status in the first wave (individuals are kept in the same strata).

Models differ because of missing values treatment. Model 1: Missing values are not treated; Model 2: income missing values replaced by multiple-imputed values; Model 3: BMI and income missing values replaced by multiple-imputed values.

For further details on missing values treatment see Appendix A, section A.1.2.

Our estimation also accounts of wave effects, through the inclusion of time and adolescent dummies.

Chapter 3

Ordinal Welfare Comparisons with Multiple Discrete Indicators: A First Order Dominance Approach and Application to Child Poverty

World Development, **vol. 40(11), pages 2290-2301.**

Coauthored with Channing Arndt, M. Azhar Hussain, and Lars Peter Østerdal (Main Authors)
Maimuna Ibraimo, and Pham Lan Huong

Abstract

We develop an approach for making welfare comparisons between populations with multidimensional discrete well-being indicators observed at the micro level. The approach is rooted in the concept of multidimensional first order dominance. It assumes that, for each indicator, the levels can be ranked ordinally from worse to better; however, no assumptions are made about relative importance of any dimension nor about complementarity/substitutability relationships between dimensions. We also introduce an efficient algorithm for determining dominance and employ a bootstrap approach that permits cardinal rankings of populations. These approaches are applied to household survey data from Vietnam and Mozambique.

3.1 Introduction

Appropriate poverty measurement remains an active area of research. Traditional models of social welfare and inequality assume one-dimensional indicators, usually based on monetary variables (e.g., Sen, 1973). Nevertheless, poverty (or welfare) has long been recognized as a multidimensional phenomenon. Motivated by the desire to consider more dimensions in analyzing social welfare, poverty, and inequality (e.g., Sen, 2006; UNDP, 1990), recent literature has frequently focused on multidimensional measures of poverty. For example, Alkire and Foster (2011), Roelen and Gassmann (2008) and Rippin (2010) discuss weighting schemes to aggregate across multiple indicators of poverty and well-being. Application of a weighting scheme is very convenient and can be easily justified when a reasonably high degree of consensus exists on the appropriate values for weights. In the absence of such a consensus, application of methods that require weighting schemes can quickly become problematic as alternative weighting schemes may alter conclusions with respect to the welfare rankings of populations. In these cases, it is useful to consider what can be said concerning the welfare status of two populations without making recourse to a weighting scheme.

In response to the challenge of limiting the imposition of subjective assumptions, other contributions have focused on development of "robust" methods for comparing population welfare, poverty, and/or inequality with multidimensional data. These methods allow for valid comparisons across broad classes of underlying social welfare functions. Following the seminal work by Atkinson and Bourguignon (1982, 1987) and Bourguignon (1989), recent contributions include Duclos, Kahn, and Younger (2006, 2007), Bourguignon and Chakravarty (2003), Crawford (2005), Grab and Grimm (2007), Gravel, Moyes, and Tarroux (2009), Batana and Duclos (2010), Gravel and Mukhopadhyay (2010), and Muller and Trannoy (2011) among others. Still, these contributions apply conditions that are typically formulated in terms of specified signs on the second or higher order cross-derivatives of the underlying social welfare functions. In this paper, we consider the problem of making welfare comparisons between populations in a situation where only ordinal information is available at the micro level in terms of multidimensional (discrete) well-being indicators. The term "ordinal" here means that, for each well-being indicator, the levels can be ranked from worse to better. However, no assumptions are made about the strength of preference for each dimension, nor about the relative desirability of changes between levels within or between dimensions or the complementarity/substitutability between the dimensions.

To accomplish this, we draw upon a concept known in the literature as multidimensional first order dominance (henceforth, FOD). This concept allows us to make welfare comparisons between two populations on the basis of a series of (binary or multileveled) ordinal welfare indicators. [1] In addition, we introduce a rapid and reliable algorithm for empirically determining whether one population dominates another on the basis of available binary indicators by drawing upon linear programming theory.

The FOD approach obviates the need for the analyst to apply an (arbitrary) weighting scheme across multiple criteria or to impose conditions on the social welfare function, which can be a considerable advantage. However, as with any other "robust" method, this gain comes at some cost. First, the procedure may be unable to determine any difference between two populations. In other words, it can happen that population A does not dominate population B and population B does not dominate population A. Hence, the welfare ranking, based on FOD, is indeterminate. Second, as a pure binary indicator, the FOD procedure provides no sense as to the degree of dominance (or similarity) between two populations. Assume population A dominates population B. Without additional information, one does not know whether population A dominates population B by a considerable degree, such that the conclusion of dominance remains even if "large" declines in the individual welfare indicators of population A occur, or whether the conclusion of dominance rests on a knife's edge such that even a "small" decline in any one welfare indicator for population A would lead to an indeterminate outcome.

[1] Note that the analysis conducted here focuses on relative welfare/ poverty. There is no attempt to define a threshold below which some share of the population is considered poor.

We mitigate these costs through the application of a bootstrap approach (technical details are presented in Appendix B). In short, repeated bootstrap samples are drawn from the comparator populations, which are often themselves samples of larger populations. When these repeated bootstrap samples are compared, the final output can be interpreted as an empirical probability that population A dominates population B. These probabilities yield significantly more information than the static application of FOD. For example, we may find that occasionally population A dominates population B and occasionally the inverse occurs but most of the time the results are indeterminate. Or, we might find that A dominates B almost always. Or, we may find that the probability that A dominates B is fairly high while the probability that B dominates A is very low or zero. These cases correspond with the conclusions rough equality of A and B, solid dominance of A over B, and likely dominance of A over B respectively. Finally, if one is willing to accept the probability that A dominates B as a cardinal measure of welfare, one can then easily derive measures that yield cardinal welfare rankings across multiple populations (e.g., all provinces in a country or all countries in a region). Hence, without imposing weights on the various chosen binary welfare indicators that determine all results, one can cardinally rank populations by welfare status.

These approaches are applied to data from Vietnam and Mozambique with a focus on the distribution and evolution of child poverty through space and time. These countries were chosen because they are surprisingly close Asian and African analogs. In addition, they both exhibited rapid rates of economic growth over the periods considered. The focus on child poverty relates to the strong preference for a multidimensional view when evaluating the welfare of children (Roelen, Gassmann, and de Neubourg, 2010). It also permits comparison with existing studies in Mozambique and Vietnam that have employed multidimensional indicators.

In Vietnam, it is well accepted that most objective welfare indicators have been improving on average, including multidimensional child poverty measures (Roelen, 2010); however, the distribution of gains is increasingly in focus. As will be shown, the FOD approach is particularly well suited to considering whether gains are broad-based and to making comparisons across sub-groups (such as regions). In Mozambique, current debate centers on the recent stagnation in measured consumption poverty (DNEAP, 2010). While Arndt et al. (in press) show that this stagnation is consistent with an array of economic indicators, the multidimensional analysis conducted here provides a valuable additional perspective and complements existing deprivation based studies of child poverty (UNICEF, 2006, 2011).

The remainder of this article is laid out as follows. Section 2 provides a technical review of the multidimensional first order dominance methodology. Section 3 introduces our case countries, Vietnam and Mozambique, and presents the binary welfare indicators employed to measure child welfare. Section 4 presents results, and Section 5 presents concluding remarks and directions for future research.

3.2 Multidimensional First Order Dominance

FOD comparisons provide a way of comparing multidimensional well-being without relying on *ad hoc* assumptions about individual well-being or social welfare. FOD can be characterized in several equivalent ways, as reviewed in the following paragraphs. [2]

3.2.1 Definitions

Much research into the nature of distributional dominance concepts has been conducted, and the theory is by now well developed (see e.g., Marshall and Olkin, 1979; Müller and Stoyan, 2002; Shaked and Shanthikumar, 2007 for general treatments). The traditional criterion for one distribution being unambiguously "better" than another is that of first order dominance, also known as the *usual (stochastic) order* in the stochastic dominance literature.

We start by reviewing the classical theory of one-dimensional first order dominance. For our purpose, we can focus on a model with only a finite set of possible outcomes for each individual in the population. Assume, therefore, that the distribution of well-

[2] See also Østerdal (2010) for a discussion.

being of some population is described by probability mass function f over a finite set of real-valued outcomes X (i.e., $\sum f(x) = 1$ and $f(x) \geqslant 0$ for all x in X), and another population is described by the probability mass function g. In this one-dimensional case, f first order dominates g if any of the following (equivalent) conditions (a), (b) and (c) hold:

(a) g can be obtained from f by a finite sequence of bilateral transfers of density to less desirable outcomes.

(b) Social welfare is at least as high for f as for g for any nondecreasing additively separable social welfare function; i.e., $\sum_{x \in X} f(x) w(x) \geqslant \sum_{x \in X} g(x) w(x)$ for any nondecreasing real function w.

(c) $F(x) \leqslant G(x)$ for all x in X, where $F(\cdot)$ and $G(\cdot)$ are the cumulative distribution functions corresponding to f and g.

Intuitively, we could think of condition (a) as one distribution FOD another if one could hypothetically move from one population distribution to the other by iteratively shifting population mass in the direction from a better outcome to a worse outcome. Thus, whenever we are able to observe FOD between two population distributions, the dominating population is unambiguously "better off."

This fundamental characterization can be extended to a multidimensional setting (e.g., Grant, 1995; Lehmann, 1955; Levhari, Paroush, and Peleg, 1975; Strassen, 1965). Suppose now that f and g denote multidimensional probability mass functions over a finite subset X of R^n. Then, f first order dominates g if one of the following three equivalent properties (A)–(C) hold.[3]

(A) g can be obtained from f by a finite number of shifts of density from one outcome to another that is worse.

[3] The equivalence between (B) and (C) was proved by Lehmann (1955) and rediscovered in economics by Levhari et al. (1975). The equivalence between (A) and (C) has been obtained as a corollary of Strassen's Theorem (Strassen, 1965), cf. e.g., Kamae, Krengel, and O'Brien (1977). Østerdal (2010) provides a constructive direct proof of the equivalence between (A) and (C).

(B) $\sum_{x\in X} f(x)\, w(x) \geqslant \sum_{x\in X} g(x)\, w(x)$ for any nondecreasing real-valued function w.

(C) $\sum_{x\in X} f(x) \leqslant \sum_{x\in X} g(x)$ for any comprehensive set $Y \subseteq X$.[4]

Again, notice that (A) provides perhaps the most intuitively appealing definition.

3.2.2 Checking FOD in practice

For empirical work, it is important to be able to determine in an "efficient" way whether one distribution dominates another. In principle, one can check for FOD by directly checking all the inequalities in (C). This is a simple but generally inefficient method, as the number of inequalities to be checked is very large if you have many dimensions and levels. Algorithms dealing with first order dominance have been invented, though most of them are only built for the one-dimensional case (e.g., Bawa, Lindenberg, and Rafsky, 1979; Fishburn and Lavalle, 1995). Preston (1974) and Hansel and Troallic (1978) assert that an algorithm for finding the maximum flow in a properly defined network can be used to determine dominance. More usefully, for the multivariate discrete case, Mosler and Scarsini (1991) and Dyckerhoff and Mosler (1997) show from definition (A) that first order dominance corresponds to a linear program that has a feasible solution. Hence, first order dominance can be verified using a linear programming package. We operationalize the linear programming technique in GAMS (GAMS Development Corporation, 2008). In our experience, FOD is rapidly and robustly verified using the CONOPT solver (Drud, 2008). We provide an example linear program for the three dimensional case in Appendix A. Extension to higher dimensions is straightforward.

3.2.3 Illustration of FOD with binary indicators

To illustrate the concept, let us consider a hypothetical example of two binary $0-1$ variables (dimensions) A and B, i.e., $n = 2$ and $X = \{(0,0);(0,1);(1,0);(1,1)\}$.[5] In every dimension, it is useful to think of the outcome "1" as the good outcome

[4] Y is comprehensive if $x \in Y$, $y \in X$, and $y \leq x$ implies $y \in Y$.

[5] An empirical illustration to the 2×2 case is presented in Sonne-Schmidt, Tarp, and Østerdal (2011).

(nondeprived) and "0" as the bad outcome (deprived). Thus, the outcome $(0,0)$ for a person means she is deprived in both dimensions; the outcome $(0,1)$ means she is deprived in the first dimension and nondeprived in the second dimension, and so forth.

Let f and g be two probability mass functions on X, defined as indicated in Table 1. (The percentages in bold at the right side and bottom of the table indicate the marginal distributions).

Table 1 about here

When analyzing each dimension separately, distribution g will appear to be better than distribution f because, for each dimension, a higher share of the population is not deprived (60% versus 50%). However, the welfare ranking of distributions g and f is indeterminate according to the FOD criterion. Formally, this can be seen with reference to (C). We have that f first order dominates g if and only if the following four inequalities (i)–(iv) are jointly satisfied:

(i) $g(0,0) \geqslant f(0,0)$,

(ii) $g(0,0) + g(0,1) \geqslant f(0,0) + f(0,1)$,

(iii) $g(0,0) + g(1,0) \geqslant f(0,0) + f(1,0)$, and

(iv) $g(0,0) + g(1,0) + g(0,1) \geqslant f(0,0) + f(1,0) + f(0,1)$.

Here, f does not FOD g, nor does g FOD f, since we have $g(0,0) > f(0,0)$ but, for example, $f(0,0) + f(0,1) > g(0,0) + g(0,1)$.

Intuitively, no distribution is dominant, since f would be better in the case where what matters most is to minimize the share of population deprived in both dimensions, while g would be better in the case where what matters most is, for example, the share of population not deprived in dimension A.

Let us now consider the probability mass function, h, given in Table 1. Comparing distributions h and g; h does not FOD g, nor does g FOD h, since we have $g(0,0) >$

$h(0,0)$ but $h(0,0)+h(1,0)+h(0,1) > g(0,0)+g(1,0)+g(0,1)$. Intuitively, FOD does not occur since h would be better if what matters most is minimization of the share of population deprived in both dimensions, while g would be better if what matters most is maximization of the share of the population not deprived in either dimension.

However, h FOD f. This is immediately verified from checking the four inequalities in (C) listed above. An intuitive way of seeing this (by reference to condition (A)), is to observe that we can obtain f from h by moving some probability mass (10 percentage points) from the outcome $(1,1)$ to $(0,0)$.

From the examples, it can also be seen that the FOD criterion differs from the criteria for robust welfare comparisons of the Atkinson–Bourguignon type as invoked by Atkinson and Bourguignon (1982, 1987), Bourguignon (1989), Bourguignon and Chakravarty (2003), Duclos et al. (2006, 2007) and others (see the Introduction for further references). The latter are instances of what is also known as orthant stochastic orderings (Dyckerhoff and Mosler, 1997). Orthant orderings make stronger assumptions about the underlying social welfare function than the FOD criterion.[6] In its primary variant (assuming "substitutability" between dimensions), f orthant dominates g if and only if $\sum_{y \leq x} g(x) \geqslant \sum_{y \leq x} f(x)$ for any $x \in X$. Note that this criterion is less restrictive than (C) and hence orthant dominance may appear when condition (A) (or (B)) is violated. For the 2×2 case, conditions (i)–(iii) (without condition (iv)) are necessary and sufficient for orthant dominance. In our example, h orthant dominates g. Hence, the FOD criterion differs from orthant dominance orderings even in the 2×2 case.

3.3 Case Countries and Choice of Welfare Indicators

Vietnam and Mozambique are in focus for the empirical analysis. Arndt, Garcia, Tarp, and Thurlow (2010) describe a number of similarities between Vietnam and Mozambique.

[6] A possible source of confusion is that in the multidimensional context the term "first order dominance" has been used with different meanings. In particular, in the economics literature, orthant stochastic orderings of the Atkinson and Bourguignon type for welfare comparisons are often referred to as first order dominance criteria. (Second- and higher order dominance criteria are then derived from further assumptions on the underlying social welfare function.)

In terms of geography, they are both long relatively thin countries with substantial coastline. In terms of recent history, both have conducted socialist experiments and endured brutal and extended periods of warfare. In addition, both formally adopted a much more market oriented economic approach in the same year, 1986. Since the early 1990s, both Vietnam and Mozambique have been among the fastest growing economies in the world. There are structural similarities as well. In both countries, about 70% of the population is rural. Also, the composition of value added across sectors is surprisingly similar (Arndt et al., 2010). Finally, both Vietnam and Mozambique receive significant external resources. Mozambique has been, since the early 1990s, one of the largest aid recipients in the world on a per capita basis. At the same time, Vietnam has been one of the largest aid recipients in absolute terms. When aid to Vietnam is combined with offshore oil revenues, the per capita value of these resources is roughly similar between the two countries.

There are also important differences. Economic takeoff began in earnest earlier and from a more developed base in Vietnam. As a result, Vietnam is richer. Population size differs dramatically with the Vietnamese population being about four times larger than the population of Mozambique. At the same time, land area is smaller in Vietnam. Vietnam is one of the most densely populated countries in the world while population density in Mozambique is relatively sparse. Finally, while both countries are investing heavily in education, Vietnam began its economic takeoff with much higher levels of educational attainment and these differences persist. Other social indicators, such as the infant mortality rate and access to health care services, are generally much better in Vietnam for similar reasons.

As highlighted in the Introduction, current debate in Vietnam tends to center around the distribution of gains. In Mozambique, there is considerable interest in determining whether the recent stagnation in consumption poverty is being accompanied by a slowdown or stagnation in other measures of welfare. The FOD analysis here contributes to both of these debates.

3.3.1 Multidimensional child poverty in Vietnam

As indicated, following the Doi Moi reforms, the country experienced rapid economic growth that allowed a reduction in monetary poverty. However, little was known about the specific situation of Vietnamese children until the works by MOLISA, University of Maastricht, and UNICEF (2008) (UNICEF (2008), henceforth), Roelen (2010), and Roelen, Gassmann, and de Neubourg (2009, 2010). These studies apply an outcome- and deprivation-based approach to estimating child poverty in Vietnam and obtain broadly comparable results.

The UNICEF (2008) report employs Multiple Indicator Cluster Surveys (MICS) and Vietnam Households Living Standard Survey (VHLLS) data from 2006 to produce two Vietnam-specific outcome measures, namely the Child Poverty Rate (CPR) and the Child Poverty Index (CPI). The former is a headcount measure referred to the proportion of poor children, the latter is an index calculated at the regional level. To this end, the authors select six domains (with one or more indicators each) for their child poverty approach along the lines of the works by Biggeri (2007) and Alkire (2008): education poverty, health poverty, shelter poverty, water and sanitation poverty, social inclusion and protection poverty, and child labour. Thus, the indicators are aggregated over attributes per individual into the two outcome measures, first at the domain level and then at the overall level.[7]

According to the CPR, $31 - 37\%$ of all children below 16 years of age are poor, with a marked difference in poverty incidence between rural and urban areas and vast heterogeneity - therefore inequality - among different geographic regions. No gender differences are found. The CPI, for its part, indicates that the best performing region is Red River Delta, while the worst performing one is the North West region. [8]

[7] The poverty criterion used to compute the CPR is deprivation in at least two domains, where the deprivation thresholds partly match those defined in the Bristol Study (Gordon, Nandy, Pantazis, Pemberton, and Townsend, 2003a, 2003b). The overall proportion of poor children is determined both at the regional and national level. As to the CPI, instead, average indicator poverty rates per domain are aggregated by dividing the sum of squared domain scores by the number of domains. Indeed, such index can be considered a squared domain severity index.

[8] Roelen et al. (2009) develop a similar analysis with MICS 2006 data. They design two Vietnam-

Roelen et al. (2010) also estimate child poverty incidence, depth, and severity in Vietnam. The authors calculate the poverty headcount based on a dual cut-off identification strategy, a child poverty gap measure (percentage of deprivations over the maximum number of observable deprivations) and a child poverty severity measure. Findings show that 37% of children are poor, with an average poverty gap of 21%. High poverty incidence is correlated with deeper and more severe poverty. Rural areas are poorer than urban areas, and poverty incidence varies among geographic regions. The results also suggest variability in poverty depth and severity by age group: youngest and oldest children are subject to deeper and more severe poverty.[9]

3.3.2 Multidimensional childhood poverty in Mozambique

UNICEF (2006, 2011) takes a human rights-based approach to childhood poverty in Mozambique. These studies examine deprivation-based child poverty in the dimensions of water, sanitation, shelter, education, health, nutrition, and information (the Bristol Indicators, cf. Gordon et al., 2003a, 2003b). A number of data sets were used to capture a broader picture of childhood poverty: MICS 2008 (also used here), IOF 2008–09 and the National Child Mortality Study 2009. Overall, Mozambique experienced reductions in child and maternal mortality, stunting, and increase in school enrollment. At the same time, child poverty and deprivation remains high with improvements threatened by the AIDS pandemic. In both studies, absolute poverty is defined as having two or more severe deprivations in any of the mentioned deprivation domains. The studies lend some support to the view that deprivation-based poverty (at least two deprivations) and consumption-based poverty (being below the poverty line) are not necessarily highly correlated. Deprivation-based poverty (at least two severe deprivations) decreased from 59% in 2003 to 48% in 2008, which is in contrast to consumption-based poverty that

specific outcome measures, namely the Child Vulnerability to Poverty Rate (CVPR) and the Child Vulnerability to Poverty Index (CVPI), computed in the same way as the CPR and CPI, respectively. The only difference with the UNICEF (2008) study is the inclusion of a leisure poverty domain and their findings are very similar.

[9] However, this might be the byproduct of differences in age groupspecific indicators.

shows stagnation at around 55% of the population from 2002–03 to 2008–09.

3.3.3 Welfare indicators

To consider the living standards of Vietnamese and Mozambican children through time and across space, we choose five main indicators of welfare in the spirit of the severe deprivation notion of the Bristol Indicators (cf. Gordon et al., 2003a, 2003b). They are defined as follows:

- *Severe water deprivation*: Children who only have access to surface water for drinking or for whom the nearest source of water is not within 15 minutes from their dwelling.

- *Severe sanitation facilities deprivation*: Children who have no access to any kind of improved latrine or toilet.

- *Severe shelter deprivation*: Children living in dwellings with more than five people per room (severe overcrowding) or with no flooring material (e.g., a mud floor).

- *Severe education deprivation*: Children who had never been to school and were not currently attending school.

- *Severe information deprivation*: Children who belong to a household where there is not access to a TV set nor to a radio.

For Vietnam, we use the MICS from 2000 and from 2006. For Mozambique we use the Demographic and Health Survey (DHS) from 2003 and the MICS from 2008.[10]

[10] The definitional and operational differences between the DHS and MICS surveys are (relatively) small. To assure comparability, consultations were conducted between the MICS and DHS teams (Gordon, Nandy, Pantazis, Pemberton, & Townsend, 2010). Both surveys were conducted by the National Institute of Statistics.

3.4 Results

3.4.1 Descriptive Statistics

For the purposes of the analysis presented here, we focus on children aged 7–17. For this age group, we consider the five indicators of well-being presented above. The percentage of children not deprived in each dimension is presented in Table 2.

Table 2 about here

With five binary indicators, the number of possible welfare indicator combinations is $2^5 = 32$. Hence, in the analysis, each comparator population is divided into 32 sub-groups. Due to the properties of the approach, a population may fail to register improvement through time due to stagnation or regress in only a small number of sub-groups. This demanding characteristic renders the approach particularly well suited to determining whether gains are broad based. Similar logic applies to regional comparisons at the same point in time. The share of children falling into each combination of welfare indicators is presented in Table 3.

Table 3 about here

The first row of the table shows the share of the population characterized by severe deprivation in all dimensions. This result has a very small probability in Vietnam. In Mozambique, it is about 7% in 2003 with substantial improvement by 2008. The bottom row of the table illustrates the probability of a child not being deprived in any of the five dimensions. Here, the gain in Vietnam is impressive registering an absolute increase of about 26% points, corresponding to a relative change of 100% between the two waves. Mozambique also registers improvement in the final row (child not deprived in all dimensions) though the improvement is marginal.

3.4.2 FOD comparisons

Tables 4 and 5 illustrate the temporal FOD comparisons for Vietnam and Mozambique respectively.

Table 4 and 5 about here

In Vietnam, advance in well-being is registered at the national level, in rural zones, and in two regions using the static approach. The bootstrap confirms that the advances at the national level and in rural zones are robust. Advance in the Mekong River Delta is also robust while advance in the South East region is somewhat more likely than an indeterminate outcome. Positive (empirical) probability of advance is also registered in urban zones and three additional regions. There is essentially no probability of regression through time in any region. These results provide evidence that, at the national level and frequently on a region by region basis, gains over 2000–2006 period were reasonably broad based.

Mozambique registers fewer gains through time. As in Vietnam, there is essentially no evidence of regression through time. Nevertheless, only one province, Niassa, exhibits gains through time using the static approach over the 2003–2008 period. Niassa was also one of the best performing provinces in terms of poverty headcount using consumption as a metric over the same period (DNEAP, 2010). Nevertheless, the bootstrap indicates that this gain is only somewhat more likely than an indeterminate outcome. There is positive probability of advance at the national level, in rural zones, and in six of 11 provinces. However, these probabilities tend to be quite small. Zambézia province registers about a one in four chance of advance through time. These results point in broadly the same directions as the consumption poverty measures with the exception of Zambézia, which exhibited an increase in consumption based poverty.

FOD comparisons are also possible across regions for a given point in time. Tables 6–9 show, for Vietnam, regional comparisons for the cases: static 2000, bootstrap 2000, static 2006, and bootstrap 2006 respectively.

Table 6 to 9 about here

Tables 10–13 show analogous results for Mozambique.

Table 10 to 13 about here

In each case, the row average and the column average are provided. The row (column) average provides the probability that the region dominates (is dominated by) another region.

Using these metrics, relatively well-off regions should have relatively large row averages while relatively poor regions should have relatively large column averages. In Vietnam, urban zones, the Red River Delta, and the South East are shown to be relatively well-off. On the other hand, the rural zones, North East, the North West, and the Central Highlands are shown to be relatively poor. Consistent with the temporal analysis, the Mekong River Delta is shown to be relatively poor in 2000 but improves to being neither particularly poor nor particularly well-off in 2006.

For Mozambique, the relatively well-off regions are the urban zones and Maputo Province and City. Relatively disfavored provinces include rural zones, Tete, Zambézia, Nampula, and Cabo Delgado. In the temporal analysis, Niassa registers improvement through time. This does not show up in the changes in the static spatial row/column averages through time as Niassa becomes dominated by the urban zone (Niassa is 77% rural) while remaining dominated by Maputo Province and Maputo City. Some progress is evident in the bootstrap where Niassa registers a small gain. Zambézia also exhibits a reasonable chance of temporal gain. This is more evident in the inter-regional comparisons. Zambézia Province is dominated by other provinces less frequently and less decidedly in 2008 compared with 2003. Despite these gains, Zambézia is the poorest province in both 2003 and 2008 using the column average as the metric.

Finally, the results provide some indication of trends in spatial inequality. If all regions were nearly equal, then it is likely that no region would dominate any other (at least it would not do so with a very high probability for the bootstrapping case). The values provided in the matrices shown in Tables 6–13 would then tend to be small and with roughly equivalent values for the row and column averages. The average of the row averages (which equals the average of the column averages) provides a measure of total overall registered probability of dominance in the static case and in the bootstrap case. An increase (decrease) in this value over time could be taken as an indication of an increase (decrease) in inequality across regions.

By this measure, spatial inequality in Vietnam has remained roughly the same (a small increase is found in both the static and the bootstrap cases). Overall, while the temporal FOD analysis indicates that the strong gains in average objective indicators are being reasonably shared across the population from a national perspective and in many regions, the pattern of improvement does not appear to be reducing existing (relatively large) spatial inequalities.

Turning to Mozambique and focusing on trends in spatial inequality, Mozambique registers a fairly dramatic decline in the probabilities of dominance in both the static and bootstrap cases. At the extremes, Maputo Province and City, while still performing relatively well, are less overwhelmingly dominant and Zambézia, while still performing relatively poorly, is less overwhelmingly dominated. In addition, the reductions in dominance do not occur uniquely at the extremes. Of the 14 regions considered (three aggregates and 11 provinces) and focusing on the bootstrap results, 11 exhibit reduced row averages and 11 (not the same 11) exhibit reduced column averages between 2003 and 2008. As a result, in many ways, the opposite conclusion to Vietnam pertains. While improvements in objective indicators are insufficiently strong or insufficiently broad based to register robust improvements in the temporal analysis, the pattern of gains on a regional basis tends to point toward a reduction in spatial inequalities.

3.5 Conclusions

The FOD criterion is a demanding test for the dominance of a population distribution relative to another in that only a minimum of highly plausible assumptions on underlying social welfare criteria are made. Despite the generality of the criterion, the empirical analysis illustrates that this criterion delivers useful comparisons of populations. This is particularly true in the context of child poverty where the application of one-dimensional (household) income-based welfare/inequality/poverty measures tend to provide too narrow a view given the importance of other indicators such as access to publicly provided goods and services.

While the theoretical underpinnings of the multidimensional FOD criterion have been known and appreciated in the stochastic dominance literature for around half a century, to our knowledge, an empirical implementation of the multidimensional FOD criterion for comparisons of actual population distributions has never been conducted prior to this study. Our findings provide strong evidence for broad based advance in the welfare of 7–17 year old in Vietnam using the five chosen indicator variables. Because these findings are based on the barest minimum of underlying assumptions, they lend strong support to the similar conclusions obtained by existing studies in Vietnam. Evidence for advance in Mozambique, on the other hand, is much more muted. This result is consistent with recent evidence on consumption based poverty. As pointed out by UNICEF (2011), some improvements have been registered. Nevertheless, these gains were insufficiently generalized across indicators and insufficiently broad based across the population to register as unambiguous improvement. Importantly, in neither country is there any evidence of regression through time. Finally, the FOD analysis provides a useful and novel perspective on inequality. In Vietnam, regional differences remain relatively constant. In Mozambique, evidence exists for a reduction in regional disparities between 2003 and 2008. In absolute terms, spatial inequalities remain pronounced in both countries.

Future research may take several directions. We shall only mention two:

First, in our analysis, we have focused on the Bristol Indicators for severe child deprivation (adapted to the context of the case countries and available data). While the welfare comparisons are robust for given indicators, changing the indicators themselves may, of course, change conclusions. Our empirical implementation strategy may be adapted to deal with additional (binary or multileveled) indicators. The number of inequalities to be tested for each pairwise comparison of distributions, however, increases dramatically with the addition of further levels or dimensions to the existing indicators and fewer FOD are to be expected. Future research may explore the value of expanding dimensions and levels in the FOD approach. Second, in the present paper, we focused on a single age group (children aged 7–17) and welfare comparisons within a single country. With the widespread availability of data from DHS (potentially supplemented by MICS), the possibility exists to compare target populations across countries. If children remained in focus, it would be possible to consider the evolution of the living conditions of children and develop indicators of the degree of inequality in important indicators of welfare across a broad array of countries.

Bibliography

[1] Alkire, S. (2008). "Choosing dimensions: The capability approach and multidimensional poverty" In N. Kakwani, and J. Silber (Eds.), *The many dimensions of poverty* (pp. 89–119). New York, NY: Palgrave-MacMillan.

[2] Alkire, S., and Foster, J. (2011). "Counting and multidimensional poverty measurement," *Journal of Public Economics*, 95(7–8), 476–487.

[3] Arndt, C., Garcia, A., Tarp, F., and Thurlow, J. (2010). "Poverty reduction and economic structure: Comparative path analysis for Mozambique and Vietnam," UNU-WIDER Working Paper WP2010/122.

[4] Arndt, C., Hussain, M. A., Jones, E. S., Nhate, V., Tarp, F., and Thurlow, J. (in press). "Explaining the evolution of poverty: The case of Mozambique," *American Journal of Agricultural Economics*.

[5] Atkinson, A. B., and Bourguignon, F. (1982). "The comparison of multidimensioned distributions of economic status," *Review of Economic Studies*, 49(2), 183–201.

[6] Atkinson, A. B., and Bourguignon, F. (1987). Income distribution and differences in need. In G. R. Feiwek (Ed.), Arrow and the foundations of the theory of economic policy. London, UK: Harvester Wheatsheaf.

[7] Batana, Y. M., and Duclos, J.-Y. (2010). Comparing multidimensional poverty with qualitative indicators of well-being. CIRPEE Working Paper 10-04.

[8] Bawa, V. S., Lindenberg, E. B., and Rafsky, L. C. (1979). An efficient algorithm to determine stochastic dominance admissible sets. Management Science, 25(7), 609–622.

[9] Biggeri, M. (2007). Choosing dimensions in the case of children's wellbeing. Maitreyee, (7), 4–7, Human Development and Capability Association.

[10] Bourguignon, F. (1989). Family size and social utility: Income distribution dominance criteria. Journal of Econometrics, 42(1), 67–80.

[11] Bourguignon, F., and Chakravarty, S. (2003). The measurement of multidimensional poverty. Journal of Economic Inequality, 1(1), 25–49.

[12] Crawford, I. (2005). A nonparametric test of stochastic dominance in multivariate distributions. DP 12/05, Department of Economics, University of Surrey.

[13] DNEAP. (2010). Poverty and well-being in Mozambique: The third national assessment. National Directorate of Studies and Policy Analysis, Ministry of Planning and Development, Maputo, Mozambique.

[14] Drud, A. (2008). CONOPT. ARKI Consulting and Development A/S, Bagsvaerd, Denmark.

[15] Duclos, J.-Y., Sahn, D. E., and Younger, S. D. (2006). Robust multidimensional poverty comparisons. Economic Journal, 116, 943–968.

[16] Duclos, J.-Y., Sahn, D. E., and Younger, S. D. (2007). Robust multidimensional poverty comparisons with discrete indicators of well-being. In S. P. Jenkins, and J. Micklewright (Eds.), Inequality and poverty reexamined. Oxford, UK: Oxford University Press.

[17] Dyckerhoff, R., and Mosler, K. (1997). Orthant orderings of discrete random vectors. Journal of Statistical Planning and Inference, 62(2), 193–205.

[18] Efron, B. (1979). Bootstrap methods: Another look at the jackknife. Annals of Statistics, 7(1), 1–26.

[19] Efron, B., and Tibshirani, R. J. (1993). An introduction to the bootstrap. New York, NY: Chapman and Hall.

[20] Fishburn, P. C., and Lavalle, I. H. (1995). Stochastic dominance on unidimensional grids. Mathematics of Operations Research, 20(3), 513–525.

[21] GAMS Development Corporation. (2008). GAMS– A user's guide. Washington, DC: GAMS Development Corporation.

[22] Gordon, D., Nandy, S., Pantazis, C., Pemberton, S., and Townsend, P. (2003a). Child poverty in the developing world. Bristol, UK: Policy Press.

[23] Gordon, D., Nandy, S., Pantazis, C., Pemberton, S., and Townsend, P. (2003b). The distribution of child poverty in the developing world. Bristol, UK: Centre for International Poverty Research.

[24] Gordon, D., Nandy, S., Pantazis, C., Pemberton, S., and Townsend, P. (2010). Using Multiple Indicator Cluster Survey (MICS) and Demographic and Health Survey (DHS) data to measure child poverty. UK: University of Bristol and London School of Economics.

[25] Grab, J., and Grimm, M. (2007). Robust multiperiod poverty comparisons. Document de Travail, DIAL, DT/2007-12.

[26] Grant, S. (1995). A strong (Ross) characterization of multivariate risk aversion. Theory and Decision, 38(2), 131–152.

[27] Gravel, N., Moyes, P., and Tarroux, B. (2009). Robust international comparisons of distributions of disposable income and regional public goods. Economica, 76, 432–461.

[28] Gravel, N., and Mukhopadhyay, A. (2010). Is India better off today than 15 years ago? A robust multidimensional answer. Journal of Economic Inequality, 8(2), 173–195.

[29] Hansel, G., and Troallic, J. P. (1978). Measures marginales et theore‘me de ford-fulkerson. Zeitschrift für Wahrsheinlichkeitstheorie und verwandte Gebiete, 43(3), 345–351.

[30] Kamae, T., Krengel, U., and O'Brien, G. L. (1977). Stochastic inequalities on partially ordered spaces. The Annals of Probability, 5(6), 899–912.

[31] Lehmann, E. L. (1955). Ordered families of distributions. The Annals of Mathematical Statistics, 26(3), 399–419.

[32] Levhari, D., Paroush, J., and Peleg, B. (1975). Efficiency analysis for multivariate distributions. Review of Economic Studies, 42(1), 87–91.

[33] Marshall, A. W., and Olkin, I. (1979). Inequalities: Theory of majorization and its applications. New York, NY: Academic Press.

[34] Mosler, K. C., and Scarsini, M. R. (1991). Stochastic orders and decision under risk. Heyward, California: Institute of Mathematical Statistics.

[35] Müller, A., and Stoyan, D. (2002). Comparison methods for stochastic models and risks. John Wiley and Sons.

[36] Muller, C., and Trannoy, A. (2011). A dominance approach to the appraisal of the distribution of well-being across countries. Journal of Public Economics, 95(3–4), 239–246.

[37] Østerdal, L. P. (2010). The mass transfer approach to multivariate discrete first order stochastic dominance: Direct proof and implications. Journal of Mathematical Economics, 46(6), 1222–1228.

[38] Preston, C. (1974). A generalization of the FKG inequalities. Communications in Mathematical Physics, 36(3), 233–241.

[39] Rippin, N. (2010). Poverty severity in a multidimensional framework: The issue of inequality between dimensions. Courant Research Centre: Poverty, Equity and Growth – Discussion Paper No. 47.

[40] Roelen, K. (2010). Multidimensional child poverty in Vietnam from a longitudinal perspective – Improved lives or impoverished conditions? Chronic Poverty Research Centre. Available from http://www.chronicpoverty.org/uploads/publication_files/roelen_child_poverty_vietnam.pdf.

[41] Roelen, K., and Gassmann, F. (2008). Measuring child poverty and wellbeing: A literature review. Maastricht Graduate School of Governance Working Paper Series No. 2008/WP001.

[42] Roelen, K., Gassmann, F., and de Neubourg, C. (2009). The importance of choice and definition for the measurement of child poverty—The case of Vietnam. Child Indicators Research, 2(3), 245–263.

[43] Roelen, K., Gassmann, F., and de Neubourg, C. (2010). Child poverty in Vietnam: Providing insights using a country-specific and multidimensional model. Social Indicators Research, 98(1), 129–145.

[44] Sen, A. (1973). On economic inequality. New York: Oxford University Press.

[45] Sen, A. (2006). Conceptualizing and measuring poverty. In D. Grusky, and R. Kanbur (Eds.), Poverty and inequality. Stanford University Press.

[46] Shaked, M., and Shanthikumar, J. G. (2007). Stochastic orders. Springer.

[47] Sonne-Schmidt, C., Tarp, F., and Østerdal, L. P. (2011). Ordinal multidimensional inequality: Theory and application to the 2 2 case. (A previous version entitled "Ordinal comparison of multidimensional deprivation: Theory and application" is available as Discussion Paper 08-33, Department of Economics, University of Copenhagen).

[48] Strassen, V. (1965). The existence of probability measures with give marginals. Annals of Mathematical Statistics, 36(2), 423–439.

[49] UNDP. (1990). Human development report. New York: Oxford University Press.

[50] UNICEF. (2006). Childhood poverty in Mozambique: A situation and trends analysis. Maputo, Mozambique.

[51] UNICEF. (2008). Children in Vietnam – Who and where are the poor. MOLISA, University of Maastricht, and UNICEF.

[52] UNICEF. (2011). Child poverty and disparities in Mozambique. Maputo, Mozambique.

A Example Linear Programming Formulation of FOD Test

We illustrate the linear program formulation of the FOD test by way of an example with three dimensions.

Define binary indices i, j, k, which each can take the value 0 or 1. The value 0 refers to deprived and the value 1 to not deprived for the three dimensions.

Define binary indices i', j', k', which are aliases of i, j and k, respectively.

For two populations A and B, let a_{ijk} and b_{ijk} be the share of the respective populations corresponding to the state of deprived and not deprived for the three indicators. So, for example, the value of a_{111} is the share of population A not deprived in any dimension while the value of b_{100} is the share of population B that is not deprived in the first dimension while deprived in all other dimensions.

Define the variable $x_{ijki\prime j\prime k\prime}$ which represents transfer of probability mass from outcome ijk to outcome $i'j'k'$.

Define Z as the set of source–destination pairs $(ijk, i'j'k')$ that move probability from preferred to less preferred outcomes. That is, if outcome ijk is the source of the transfer and outcome $i'j'k'$ is the destination, a legal transfer is where $i' \leq i$, $j' \leq j$, and $k' \leq k$. All three conditions must hold.

For example, $(111, 011)$ is an element of Z while $(001, 011)$ is not an element of Z.

Under these conditions, population A FOD population B if and only if the following linear program is feasible.

$$
\begin{aligned}
&\min y = 1 \\
&\text{subject to} \\
&a_{ijk} + \sum_{(i'j'k',ijk)\in Z} x_{i'j'k',ijk} + \sum_{(ijk,i'j'k')\in Z} x_{ijk,i'j'k'} = b_{ijk} \quad \forall i, j, k \\
&x_{ijk,i'j'k'} \geqslant 0,\; x_{ijk,ijk} = 0
\end{aligned}
$$

Extension to a higher dimension involves defining a new index l, its alias $l\prime$, and appropriately expanding the dimensions of all parameters, variables, and equations. The GAMS code for operationalization of the FOD test with up to 7 binary indicators is available from the authors upon request.

B The Bootstrapping

Bootstrapping is a general means of generating consistent estimates of an estimator's sampling distribution when an analytical solution cannot be derived or requires unreasonable assumptions (Efron, 1979; Efron and Tibshirani, 1993). It is based on repeated (J times) samples, drawn with replacement, of size K from the original sample data, of size N, where $K < N$. As the original sample size, N, increases, the bootstrap approach converges to Monte Carlo for fixed K. The primary assumption behind the bootstrap is that the distribution of the observed sample is a good approximation of the distribution of the population.

In our application, the bootstrap samples are drawn in a manner that mimics the stratified cluster sample design of the DHS and MICS surveys. That is, within each stratum, K clusters are randomly drawn, with replacement, where K is also the number of primary sampling units in the stratum (i.e., $K = N$). When a cluster is drawn, all of the households in that cluster are drawn. Because the bootstrap sampling is done with replacement, each cluster (and household) may appear one or more times in a given bootstrap sample, or not at all. The FOD analysis using the linear programming techniques discussed in the previous section is conducted for each bootstrap sample. The process is repeated $J = 1000$ times. The share of times where temporal and/or spatial dominance is discovered over the 100 bootstrap replications is then calculated for each result.

For the cases, like the ones considered in this article, where the populations being considered are in fact samples from larger populations (say A and B), the results of the bootstrap can be interpreted as probabilities for three possible outcomes: (i) A FOD B; (ii) B FOD A and (iii) an indeterminate outcome. It is, in this sense, a form of statistical inference analysis with respect to the static case. Development of more formal inference procedures is a potential topic for future research.

Table 1

The distribution for f				
f		Dimension B		Total
		0 (deprived)	1 (not deprived)	
Dimension A	0 (deprived)	25%	25%	50%
	1 (not deprived)	25%	25%	50%
Total		50%	50%	100%

The distribution for g				
g		Dimension B		Total
		0 (deprived)	1 (not deprived)	
Dimension A	0 (deprived)	30%	10%	40%
	1 (not deprived)	10%	50%	60%
Total		40%	60%	100%

The distribution for h				
h		Dimension B		Total
		0 (deprived)	1 (not deprived)	
Dimension A	0 (deprived)	15%	25%	40%
	1 (not deprived)	25%	35%	60%
Total		40%	60%	100%

Table 2

Children not deprived by welfare indicator, 7-17 years old (%).

	Vietnam		Mozambique	
	2000	2006	2003	2008
Water	75.7	87.8	37.6	33.3
Sanitation	37.1	70.9	52.7	60.0
Shelter	60.4	78.4	30.3	46.0
Education	96	98.2	76.0	88.4
Information	76.9	87.1	61.8	63.5

Source: Own calculations based on MICS 2 (2000) and 3 (2006) for Vietnam, and DHS 2003 and MICS 2008 for Mozambique.

Table 3

Children by combination of welfare indicators, 7-17 years old (%)

Welfare Indicator Combination					Vietnam			Mozambique		
Water	Sanita.	Shelter	Educa.	Inform.	2000	2006	Change	2003	2008	Change
0	0	0	0	0	1.33	0.22	-1.11	6.92	2.09	-4.83
0	0	0	0	1	1.13	0.08	-1.05	5.88	2.03	-3.86
0	0	0	1	0	5.44	1.24	-4.20	10.80	9.07	-1.73
0	0	0	1	1	8.30	1.68	-6.63	11.17	10.24	-0.94
0	0	1	0	0	0.13	0.02	-0.11	0.50	0.75	0.26
0	0	1	0	1	0.18	0.05	-0.12	0.61	1.03	0.42
0	0	1	1	0	1.97	0.85	-1.12	0.80	2.93	2.13
0	0	1	1	1	4.59	1.79	-2.80	1.33	5.61	4.28
0	1	0	0	0	0.00	0.03	0.03	2.60	1.27	-1.33
0	1	0	0	1	0.00	0.06	0.06	2.41	1.02	-1.39
0	1	0	1	0	0.07	0.40	0.33	5.19	6.83	1.65
0	1	0	1	1	0.22	1.26	1.04	8.43	8.96	0.53
0	1	1	0	0	0.00	0.00	0.00	0.14	0.24	0.10
0	1	1	1	0	0.00	0.05	0.05	0.30	0.55	0.25
0	1	1	0	1	0.12	0.51	0.39	1.35	3.36	2.01
0	1	1	1	1	0.85	3.98	3.13	3.99	10.74	6.75
1	0	0	0	0	0.35	0.35	0.00	0.88	0.76	-0.12
1	0	0	0	1	0.41	0.19	-0.22	0.98	0.25	-0.73
1	0	0	1	0	5.34	3.49	-1.85	2.05	1.72	-0.33
1	0	0	1	1	12.58	4.89	-7.69	2.44	1.57	-0.87
1	0	1	0	0	0.08	0.05	-0.03	0.06	0.15	0.09
1	0	1	0	1	0.23	0.23	-0.01	0.17	0.14	-0.03
1	0	1	1	0	4.19	2.25	-1.94	0.47	0.67	0.20
1	0	1	1	1	16.69	11.72	-4.97	2.19	0.94	-1.25
1	1	0	0	0	0.02	0.01	-0.01	0.77	0.28	-0.49
1	1	0	0	1	0.03	0.13	0.10	0.95	0.39	-0.56
1	1	0	1	0	1.04	1.08	0.03	2.89	3.08	0.19
1	1	0	1	1	3.39	6.51	3.11	5.34	4.39	-0.96
1	1	1	0	0	0.01	0.03	0.02	0.13	0.18	0.05
1	1	1	0	1	0.15	0.29	0.14	0.66	0.41	-0.25
1	1	1	1	0	3.00	2.36	-0.64	2.68	3.13	0.45
1	1	1	1	1	28.16	54.21	26.05	14.88	15.18	0.31
Total					100.00	100.00	100.00	100.00	100.00	100.00

Note: In the first five columns, a “0” means that the child is deprived and a “1” means that the child is not deprived with respect to a given of the five presented welfare indicators. Source: same of Table 2.

Table 4

Temporal FOD comparisons for Vietnam (Probabilities).

	Static Case	Bootstrap			
	2006	FOD 2000	Undecided	2000 FOD 2006	Total
National	1	1.00	0.00		1
Rural	1	1.00	0.00		1
Urban		0.30	0.70		1
Red River Delta			1		1
North East		0.14	0.86	0.00	1
North West		0.04	0.96	0.00	1
North Central Coast			1		1
South Central Coast			1		1
Central Highlands		0.30	0.70		1
South East	1	0.54	0.46		1
Mekong River Delta	1	0.98	0.02		1

Note: A “1” in the static case indicates that the region’s last year welfare level FOD the first year welfare level, while an empty cell indicates no domination.

In the bootstrap case a “1” indicates that all 1,000 bootstrap replications resulted in the mentioned domination, while a “1.00” indicates that there were between 995 and 999 dominations, an empty cell indicates that there were no dominations and finally a “0.00” indicates that there were between 1-4 dominations out of a total of 1,000 bootstrap replications. Source: Same as for Table 2.

Table 5

Temporal FOD comparisons for Mozambique (Probabilities).

	Static Case	Bootstrap			
		2008 FOD 2003	Undecided	2003 FOD 2008	Total
National		0.01	0.99		1
Rural		0.08	0.92		1
Urban		0.00	1.00		1
Niassa	1	0.53	0.47		1
Cabo Delgado		0.01	0.99		1
Nampula		0.01	0.99		1
Zambézia		0.24	0.76		1
Tete			1		1
Manica			1		1
Sofala		0.01	1.00		1
Inhambane			1		1
Gaza		0.01	0.99		1
Maputo Province		0.00	1.00		1
Maputo City			1		1

Note: Same as for Table 4. Source: Same as for Table 2.

Table 6

Spatial FOD comparisons for Vietnam, 2000

	National	Rural	Urban	RRD	NE	NW	NCC	SCC	CH	SE	MRD	Avg.
National		1			1	1					1	0.40
Rural						1						0.10
Urban	1	1			1	1			1	1	1	0.70
Red River Delta	1	1			1	1	1		1		1	0.70
North East						1						0.10
North West												0.00
North Central Coast									1			0.10
South Central Coast		1				1					1	0.30
Central Highlands												0.00
South East	1	1			1	1			1		1	0.60
Mekong River Delta												0.00
Average	0.30	0.50	0.00	0.00	0.40	0.70	0.10	0.00	0.40	0.10	0.50	0.30

Note: Same as for Table 4. Averages: The row averages are the fraction of times the row region dominates other regions.
The column averages indicate the fraction of times the column region is dominated by other regions. Source: Same as for Table 2.

Table 7

Bootstrap spatial FOD comparisons for Vietnam, 2000 (Probabilities)

	National	Rural	Urban	RRD	NE	NW	NCC	SCC	CH	SE	MRD	Avg.
National		1			0.47	0.96			0		0.66	0.31
Rural					0.60	0.78					0.26	0.11
Urban	1	1			0.98	1	0.09	0.23	0.91	0.47	1	0.67
Red River Delta	0.99	1			0.94	1	0.39	0.40	0.99	0.02	0.96	0.67
North East	0.01	0.04				0.67	0		0.01		0.15	0.10
North West											0.00	0.00
North Central Coast	0.01	0.07			0.08	0.28		0.01	0.31		0.09	0.10
South Central Coast	0.12	0.52			0.31	0.93			0.18		0.72	0.30
Central Highlands	0.00	0.01			0.02	0.22					0.10	0.03
South East	0.92	0.98			0.88	1			0.62		0.98	0.54
Mekong River Delta						0.04						0.00
Average	0.31	0.46	0.00	0.00	0.37	0.69	0.05	0.06	0.30	0.05	0.49	0.28

Note: Same as for Table 4. Totals: See explanation in Table 6. Source: Same as for Table 2.

Table 8

Spatial FOD comparisons for Vietnam, 2006

	National	Rural	Urban	RRD	NE	NW	NCC	SCC	CH	SE	MRD	Avg.
National		1			1	1						0.30
Rural						1						0.10
Urban	1	1			1	1	1	1	1		1	0.80
Red River Delta	1	1			1	1	1	1	1			0.70
North East						1						0.10
North West												0.00
North Central Coast		1			1	1			1			0.40
South Central Coast					1	1						0.20
Central Highlands						1						0.10
South East	1	1			1	1			1			0.50
Mekong River Delta						1						0.10
Average	0.30	0.50	0.00	0.00	0.60	1.00	0.20	0.20	0.40	0.00	0.10	0.33

Note: Same as for Table 4. Totals: See explanation in Table 6. Source: Same as for Table 2.

Table 9

Bootstrap spatial FOD comparisons for Vietnam, 2006 (Probabilities)

	National	Rural	Urban	RRD	NE	NW	NCC	SCC	CH	SE	MRD	Avg.
National		1			0.46	1			0.02			0.25
Rural					0.27	0.99			0.01			0.13
Urban	1	1			0.98	1	0.72	0.48	0.98	0.01	0.51	0.67
Red River Delta	1	1			0.98	1	0.82	0.87	1	0.02	0.13	0.68
North East	0.00	0.00				0.59	0		0.01			0.06
North West												0.00
North Central Coast	0.08	0.41			0.73	1		0	0.30			0.25
South Central Coast	0.01	0.07			0.36	0.97	0		0.35			0.18
Central Highlands					0.03	0.92						0.10
South East	0.69	0.90			0.84	1	0.06	0.24	0.78		0.00	0.45
Mekong River Delta					0.05	0.87			0.02			0.09
Average	0.28	0.44	0.00	0.00	0.47	0.93	0.16	0.16	0.35	0.00	0.06	0.29

Note: Same as for Table 4. Totals: See explanation in Table 6. Source: Same as for Table 2.

Table 10

Spatial FOD comparisons for Mozambique, 2003

	Nat	Rur	Urb	NI	CD	NA	ZA	TE	MA	SO	IN	GA	MP	MC	Avg.
National		1					1								0.15
Rural							1								0.08
Urban	1	1			1	1	1	1	1	1					0.62
Niassa															0.00
Cabo Delgado							1								0.08
Nampula															0.00
Zambézia															0.00
Tete							1								0.08
Manica		1					1								0.15
Sofala							1								0.08
Inhambane		1					1								0.15
Gaza		1			1		1								0.23
Maputo Province	1	1		1	1	1	1	1		1	1	1			0.77
Maputo City	1	1	1	1	1	1	1	1	1	1	1	1			0.92
Average	0.23	0.54	0.08	0.15	0.31	0.23	0.85	0.23	0.15	0.23	0.15	0.15	0.00	0.00	0.30

Note: Same as for Table 4. Totals: See explanation in Table 6. Source: Same as for Table 2.

Table 11

Bootstrap spatial FOD comparisons for Mozambique, 2003

	Nat	Rur	Urb	NI	CD	NA	ZA	TE	MA	SO	IN	GA	MP	MC	Avg.
National		1				0.02	1	0.01							0.16
Rural							0.47								0.04
Urban	1	1		0.47	0.93	1	1	1.00	0.43	0.96	0.39	0.12			0.64
Niassa							0.18								0.01
Cabo Delgado		0.06		0.00			0.65	0.00							0.05
Nampula		0.04					0.48								0.04
Zambézia															0.00
Tete		0.19					0.87								0.08
Manica		0.95			0.02	0.01	0.99	0.11							0.16
Sofala		0.03				0.01	0.94								0.07
Inhambane	0.02	0.96		0.03	0.14	0.04	1.00	0.07							0.17
Gaza	0.03	0.97		0.05	0.37	0.16	1	0.20			0.10				0.22
Maputo Province	1	1	0.19	0.96	1	1	1	1	0.40	0.80	1	0.99			0.80
Maputo City	1	1	0.98	0.87	1	1	1	1	1	1	1.00	0.99	0.19		0.93
Average	0.23	0.55	0.09	0.18	0.27	0.25	0.81	0.26	0.14	0.21	0.19	0.16	0.01	0.00	0.26

Note: Same as for Table 4. Totals: See explanation in Table 6. Source: Same as for Table 2.

Table 12

Spatial FOD comparisons for Mozambique, 2008

	Nat	Rur	Urb	NI	CD	NA	ZA	TE	MA	SO	IN	GA	MP	MC	Avg.
National		1													0.08
Rural															0.00
Urban	1	1		1	1	1	1	1							0.54
Niassa					1										0.08
Cabo Delgado															0.00
Nampula															0.00
Zambézia															0.00
Tete															0.00
Manica		1					1								0.15
Sofala															0.00
Inhambane															0.00
Gaza		1			1		1		1		1				0.38
Maputo Province	1	1		1	1	1	1	1	1		1				0.69
Maputo City	1	1	1	1	1	1	1	1	1	1	1	1			0.92
Average	0.23	0.46	0.08	0.23	0.38	0.23	0.38	0.23	0.23	0.08	0.23	0.08	0.00	0.00	0.22

Note: Same as for Table 4. Totals: See explanation in Table 6. Source: Same as for Table 2.

Table 13

Bootstrap spatial FOD comparisons for Mozambique, 2008

	Nat	Rur	Urb	NI	CD	NA	ZA	TE	MA	SO	IN	GA	MP	MC	Avg.
National		1				0.09	0.07								0.09
Rural															0.00
Urban	1	1		0.72	0.94	1	1.00	0.93	0.33	0.20	0.02				0.55
Niassa					0.27		0.05								0.02
Cabo Delgado							0.00								0.00
Nampula															0.00
Zambézia															0.00
Tete															0.00
Manica		0.42				0.00	0.40	0.01							0.06
Sofala		0.04				0.00	0.13	0.01							0.01
Inhambane		0.30			0.04	0.00	0.31								0.05
Gaza	0.04	0.98		0.01	0.44	0.09	0.90	0.28	0.52	0.02	0.26				0.27
Maputo Province	0.99	1	0.00	0.38	0.95	1.00	1.00	0.97	0.65	0.07	0.59	0.03			0.59
Maputo City	1	1	1	1	1	1	1	1	1	1	1.00	0.55			0.89
Average	0.23	0.44	0.08	0.16	0.28	0.24	0.37	0.25	0.19	0.10	0.14	0.04	0.00	0.00	0.20

Note: Same as for Table 4. Totals: See explanation in Table 6. Source: Same as for Table 2.