

# Optimal Screening Policies for Childhood Obesity

Yan Yang\*, Jeremy D. Goldhaber-Fiebert<sup>†</sup>, Lawrence M. Wein<sup>‡</sup>

September 26, 2011

## Abstract

Due to the health and economic costs of childhood obesity, coupled with studies suggesting the benefits of comprehensive (dietary, physical activity and behavioral counseling) intervention, the United States Preventive Services Task Force recently recommended childhood screening and intervention for obesity beginning at age six. Using a longitudinal data set consisting of the body mass index of 3164 children up to age 18 and another longitudinal data set containing the body mass index at ages 18 and 40 and disease (hypertension and diabetes) at age 40 for 747 people, we formulate and numerically solve – separately for boys and girls – a dynamic programming problem for the optimal biennial (i.e., at ages 2, 4,  $\dots$ , 16) obesity screening thresholds. Unlike most screening problem formulations, we take a societal viewpoint, where the state of the system at each age is the population-wide probability density function of the body mass index. Compared to the biennial version of the task force’s recommendation, the optimal screening thresholds achieve a relative reduction in disease prevalence of 3% at the same screening (and treatment) cost, or – due to the flatness of the disease vs. screening tradeoff curve – achieves the same disease prevalence at a 28% relative reduction in cost. Compared to the task force’s policy, which uses the 95<sup>th</sup> percentile of body mass index (from cross-sectional growth charts tabulated by the Centers for Disease Control and Prevention) as the screening threshold for each age, the optimal policy treats mostly 16 year olds (including many who are not obese) and very few males under 14 years old.

---

\*Institute for Computational and Mathematical Engineering, Stanford University, Stanford, CA 94305, yanyang@stanford.edu

<sup>†</sup>Centers for Health Policy and Primary Outcomes Research, Stanford University School of Medicine, Stanford, CA 94305, jeremygf@stanford.edu

<sup>‡</sup>Graduate School of Business, Stanford University, Stanford, CA 94305, lwein@stanford.edu

# 1 Introduction

The prevalence of obesity in U.S. school-aged children and adolescents tripled during 1980-1999 and has plateaued since 1999 at  $\approx 17\%$ , with 31.7% of U.S. children ages 2-19 overweight or obese (Ogden *et al.* 2010). Obese children tend to have elevated lipid concentrations and blood pressure (Freedman *et al.* 2007), and often become obese adults (Singh *et al.* 2008), with an increased risk of diseases such as hypertension and type 2 diabetes (e.g., Hubert *et al.* 1983, Barrett-Connor 1989). This is not only a public health problem, but an economic (the cost of adult obesity has been estimated at \$147B/yr, Finkelstein *et al.* 2009) and national security ( $> 25\%$  of Americans ages 17-24 are unqualified for the military due to their weight, Mission Readiness 2010) issue as well.

Two complementary approaches have been put forth to address this problem: a universal approach and a targeted approach. The universal approach focuses on better nutrition and more physical activity (as in First Lady Michelle Obama’s “Let’s Move” campaign, White House Task Force on Childhood Obesity 2010), and involves schools, families, communities, and the food industry. Indeed, some school-based initiatives have generated reductions in obesity (Gortmaker *et al.* 1999, Coleman *et al.* 2005), though others have not shown statistically significant effects (The HEALTHY Study Group 2010).

The targeted approach requires measuring children’s body mass index (BMI), which is the weight divided by the square of the height, measured in  $\text{kg}/\text{m}^2$  and rounded to one decimal place. BMI is a practical and reasonably accurate (Reilly *et al.* 2000) measure of a person’s true body fat, and obesity and overweight are defined as being the 95<sup>th</sup> and 85<sup>th</sup> percentile of the gender- and age-based BMI distributions tabulated by the Centers for Disease Control and Prevention (Kuczmarski *et al.* 2000). Many school districts and entire states require measuring a child’s BMI every several years, but for surveillance – not screening – purposes (Nihiser *et al.* 2007).

Although an Institute of Medicine report states that health-care professionals should routinely track children’s BMI and offer evidence-based counseling and guidance (Institute of Medicine 2005), clinical trials for treating childhood obesity have been hampered until recently by small sample sizes, poor generalizability and variable followup, leading to insufficient evidence for effectiveness of intervention in primary care settings (e.g., US Preventive Services Task Force 2005). However, a targeted systematic review to support the updated US Preventive Services Task Force (USPSTF) recommendations concludes that research into obesity treatment has improved in the past several

years, leading to at least short-term benefits of comprehensive medium- to high-intensity ( $> 25$  hr over six months) interventions that include dietary, physical activity and behavioral counseling components (Whitlock *et al.* 2010). Consequently, the USPSTF recommends childhood BMI screening and intervention for obesity starting at age six (USPSTF 2010).

This state of affairs highlights the need to assess possible intervention strategies. To our knowledge, the only study to do this compares screening and treatment at a single age (5, 10 or 15) to universal intervention, and suggests that, should effective universal interventions be available, this approach could lead to greater health benefits than the three single-age screening strategies, and that screening at age 15 is more beneficial than screening at ages 5 or 10 because some children who are obese at these younger ages are not obese at age 18 in the absence of treatment (Goldhaber-Fiebert *et al.* 2010).

In this paper, we derive optimal screening thresholds for children under 18 years old. Using biennial BMI data from the National Longitudinal Survey of Youth (NLSY) samples (Bureau of Labor Statistics 2010), we model the Markovian evolution of the probability density function (pdf) of BMI for all U.S. children at ages  $t = 0, 2, \dots, 18$ , assuming there is no obesity treatment. We optimize over the class of biennial threshold policies, where every child has his BMI measured every two years and receives treatment whenever his BMI exceeds a gender- and age-based threshold. The impact of treatment is estimated using the pooled analysis of three comprehensive moderate- or high-intensity behavioral interventions (Whitlock *et al.* 2010). The objective function of our model is to minimize the disease prevalence of these children when they reach age 40, which is estimated using Panel Study of Income Dynamics (PSID) data (Institute for Social Research 2010). The optimal screening problem is formulated as a deterministic dynamic program with an infinite-dimensional state (the pdf of the BMI in the population at age  $t$ ).

The problem considered here shares features with other screening and treatment problems, such as the screening of cholesterol (Garber *et al.* 1994), HIV viral load (D’Amato, D’Aquila and Wein 2000), CD4 count (Shechter *et al.* 2008), estimated glomerular filtration rate (Lee, Chertow and Zenios 2008), or prostate specific antigen level (Lavieri *et al.* 2009). While the latter four papers consider decisions of when to initiate or switch treatment from an individual’s point of view (i.e., the state of the system measures a person’s health status), our paper is perhaps closest in spirit to Garber *et al.* (1994) because both studies consider the time evolution of the population-wide pdf of a health measure (cholesterol in their case, BMI in our case). While Garber *et al.*’s goal

is forecasting future cholesterol levels in the absence of treatment, we embed the dynamics in the presence of screening and treatment into a dynamic programming framework and derive optimal screening thresholds.

Exploratory data analyses in §2 guides our problem formulation in §3. Parameter estimation is undertaken in §4 and a computational study appears in §5. The results are discussed in §6.

## 2 Exploratory Data Analyses

The NLSY data set has a longitudinal population cohort of children born to women who participated in the NLSY between 1970 and 1988. Each of 3164 children (1566 boys and 1598 girls) has biennial BMI measurements from ages 0 to 18, which are partially self-reported or mother-reported. For children measured at odd-numbered ages, we linearly interpolate to obtain their BMI at even-numbered ages, and the analyses in this section considers BMI at ages  $0, 2, \dots, 18$ . Each child in this data set has an attached weight such that the weighted sample of the 3164 children represents all children born between 1970 and 1988 to women ages 21 to 31 who were present in the U.S. in 1979, and all of our analyses are on the weighted sample.

We denote the BMI of a random child at age  $t$  by  $B_t$  and let  $B$  be the stochastic process  $(B_0, B_2, \dots, B_{18})$ . We explore the Markov property of  $B$  in §2.1, the distribution of  $B_t$  in §2.2, and the distribution of the BMI increments  $(B_{t+2} - B_t)$  in §2.3.

### 2.1 Markov Property of BMI

The mean and the standard deviation of  $B_t$  decrease during ages 0-4, plateau at ages 4-6, and increase during ages 6-18 (Fig. 1 of the Electronic Companion), and hence we dismiss the possibility that  $B$  has stationary transition probabilities. We use the approach in Garber *et al.* (1994), which tests the hypothesis (all tests in this study are performed at the 95% confidence level) that  $B$  is a first-order Markov chain via a direct test of the Chapman-Kolmogorov equations. The significance levels of the chi-square statistics, which measure the difference between the observed and expected two-step (i.e., from  $t$  to  $t + 4$ ) transitions, are derived by a bootstrap technique (see §1 of the Electronic Companion for details). Of the 16 tests (ages  $t = 0, 2, \dots, 14$  for each gender), we reject the null hypothesis that BMI levels follow a first-order Markov chain in four cases: males at age 12 and females at ages 0, 6 and 14. Under a Bonferroni correction, which considers the  $1-0.05/8=0.99625$  fractile of the chi-square distributions because there are eight age groups being

tested, the Markov hypothesis is not rejected for any age-gender combination. However, because we are trying to retain the null hypothesis, this Bonferroni approach is not conservative.

## 2.2 Marginal Distribution of BMI

Initial exploratory analysis shows that  $B_t$  follows neither a normal nor a lognormal distribution, although it is much closer to the latter (data not shown). Consequently, to increase the normality of the data, we employ a Box-Cox transformation of  $\ln B_t$ , which we denote by  $X_t = \frac{(\ln B_t)^\lambda - 1}{\lambda}$  (Box and Cox 1964). The optimal value of the Box-Cox parameter is  $\lambda = -2.15$  for males and  $\lambda = -2.02$  for females. The resulting Q-Q plots in Fig. 2 of the Electronic Companion suggest that  $X_t$  is reasonably approximated by a normal distribution, particularly for the higher age groups. The Kolmogorov-Smirnov goodness-of-fit test for a normal distribution shows that at the 95% confidence level,  $X_t$  passes four of the 20 tests: girls at age 0, 8, 10 and 12. Further analysis shows that the majority of the deviation in the failed tests occurs in the left tail of the distribution (i.e., unusually low BMI levels), which is not important for the purposes of obesity screening. We also estimated the normal parameters from just the right half of  $X_t$  (i.e., only the data larger than the median), and obtained qualitatively similar results to those in Table 1 (data not shown).

## 2.3 BMI Increments

Because the increment  $X_{t+2} - X_t$  depends upon the value of  $X_t$ , we assess the normality of the increments using a 95% Kolmogorov-Smirnov test on the increment from age  $t$  to  $t + 2$  for  $t = 0, 2, \dots, 16$ , and where the (untransformed) BMI values are segregated into different buckets. The results are given in Tables 6-7 in the Electronic Companion when BMI is divided into 25 buckets ( $< 12, 12 - 13, 13 - 14, \dots, 34 - 35, 35+$ ). A cruder system with 6 buckets ( $< 16, 16 - 18, 18 - 20, 20 - 22, 22 - 24, > 24$ ) generates the results in Tables 8-9 in the Electronic Companion. With 25 buckets, only 13 of the 362 tests (88 of the 450 possible tests had no data) failed the normality test, and 36 of the 108 normality tests failed in the six-bucket system.

## 3 Problem Formulation

By our analysis in §2.1, we assume that  $X = (X_0, X_2, \dots, X_{18})$  is a Markov process, where  $X_t = \frac{(\ln B_t)^\lambda - 1}{\lambda}$  is the Box-Cox transformation of  $\ln B_t$ . Because we are considering this screening problem from a societal – rather than individual – point of view, the system state at age  $t$  is given by  $f_t(x)$ ,

which is the pdf of  $X_t$ . If treatment is given at age  $t$ , it is given after the measurement  $X_t = x$ . In the absence of treatment at age  $t$ , the increment  $X_{t+2} - X_t$  given  $X_t = x$  has pdf  $g_{xt}(w)$ , where  $w$  is the dummy variable for the value of the increment. By the analysis in §2.3, we assume that  $g_{xt}(w)$  is normal with mean  $\mu_{gxt}$  and standard deviation  $\sigma_{gxt}$ . If treatment is given at age  $t$ , we assume that  $X_{t+2} - X_t$  given  $X_t = x$  has a normal distribution with pdf  $h_{xt}(w)$ , mean  $\mu_{hxt}$  and standard deviation  $\sigma_{hxt}$ ; note that the impact of treatment is assumed to be independent of the impact of treatment received at previous ages. We consider biennial threshold policies with thresholds  $\theta_t$ , where a person with  $B_t > \theta_t$  (i.e.,  $X_t > \frac{(\ln \theta_t)^{\lambda-1}}{\lambda}$ ) receives treatment at age  $t$ , for  $t = 0, 2, \dots, 16$  (while we define  $\theta_0$  for notational convenience, we do not allow treatment at birth and set  $\theta_0 = \infty$ ). The system state equation is

$$f_{t+2}(x) = \int_0^{\frac{(\ln \theta_t)^{\lambda-1}}{\lambda}} f_t(w)g_{wt}(x-w) dw + \int_{\frac{(\ln \theta_t)^{\lambda-1}}{\lambda}}^{\infty} f_t(w)h_{wt}(x-w) dw. \quad (1)$$

In the absence of any treatment (i.e.,  $\theta_t = \infty \forall t$ ), equation (1) implies that if  $X_0$  is normal then  $X_t$  is normal for  $t = 2, 4, \dots, 18$ , which is consistent with our analysis in §2.2-2.3 that concluded that both  $X_t$  and  $X_{t+2} - X_t$  are roughly normal.

We assume that the initial system state,  $f_0(x)$ , is a skew-normal pdf with location parameter  $\mu_{f_0}$ , scale parameter  $\sigma_{f_0}$  and shape parameter  $\alpha_{f_0}$ , which has pdf  $f_0(x) = \frac{2}{\sigma_{f_0}} \phi\left(\frac{x-\mu_{f_0}}{\sigma_{f_0}}\right) \Phi\left(\frac{\alpha_{f_0}(x-\mu_{f_0})}{\sigma_{f_0}}\right)$ , where  $\phi(z)$  and  $\Phi(z)$  are the pdf and cdf of the standard normal distribution (O'Hagan and Leonard 1976). The shape parameter  $\alpha_{f_0}$  indicates the amount of skewness, and  $f_0(x)$  reduces to a normal pdf when  $\alpha_{f_0} = 0$ .

For  $t = 0, 2, \dots, 18$ , we define the disease function  $c_t(x)$  to be the contribution to the disease prevalence at age 40 due to having  $X_t = x$  at age  $t$ . The objective is to minimize the disease prevalence at age 40, which in §4.4 is computed using a nonseparable function of  $(f_0(x), f_2(x), \dots, f_{18}(x))$ . We use a surrogate objective function,

$$\min_{\theta_0, \theta_2, \dots, \theta_{16}} \sum_{t=0}^9 \int_0^{\infty} c_{2t}(x) f_{2t}(x) dx, \quad (2)$$

which is separable, and hence fits into the dynamic programming framework. This issue is discussed further at the end of §4.4.

In §4.4, we estimate the disease functions  $(c_0(x), c_2(x), \dots, c_{18}(x))$  by assuming that the Markov property of  $X$  continues until age 40 (i.e.,  $X_{40}$  depends on  $(X_0, X_2, \dots, X_{18})$  only via  $X_{18}$ )

and by assuming that the prevalence of disease at age 40 is a function of the cumulative lifetime effects of obesity; this latter assumption is relaxed in the sensitivity analyses in §5.3.

Because the USPSTF found no evidence of adverse effects from treatment on growth, eating-disorder pathology or mental health (USPSTF 2010), we simply impose a constraint on the fraction of children ages 2,4,...,16 that screen positive each year (assuming there are the same number of children of each age in the population):

$$\frac{1}{8} \sum_{t=1}^8 \int_{\frac{(\ln \theta_{2t})^{\lambda-1}}{\lambda}}^{\infty} f_{2t}(x) dx \leq p. \quad (3)$$

We refer to  $p$  as the *treatment prevalence*. Because the screening cost is negligible compared to the treatment cost, the left side of (3) is essentially proportional to the annual treatment cost. By varying  $p$  on the right side of (3), we can generate a range of policies and easily compare derived policies to recommended policies (e.g., USPSTF 2010).

Notice that although our system state is a pdf, the state transition is deterministic, and hence our problem, which is given by (1)-(3), is a deterministic constrained dynamic program.

## 4 Parameter Estimation

The initial BMI distribution, the BMI increments in the absence of treatment, the BMI increments in the presence of treatment, and the disease functions are estimated in §4.1-§4.4.

### 4.1 Initial BMI

Using the BMI data at birth, the MLE parameter values for the skew-normal pdf  $f_0(x)$  are  $\mu_{f_0} = 0.42$ ,  $\sigma_{f_0} = 0.009$  and  $\alpha_{f_0} = -1.00$  for boys, and  $\mu_{f_0} = 0.43$ ,  $\sigma_{f_0} = 0.006$  and  $\alpha_{f_0} = -1.30$  for girls. As expected from §2.2, these two distributions have very little skewness (figures not shown).

### 4.2 BMI Increments

In the absence of treatment, we assume that  $X_{t+2} - X_t$  given  $X_t = x$  is normal with pdf  $g_{xt}$ . We denote its mean and standard deviation by  $\mu_{g_{xt}}$  and  $\sigma_{g_{xt}}$ , and estimate the values of these parameters using the NLSY data set. These parameter values are plotted as functions of  $x$  for  $t = 0, 2, \dots, 16$  for boys and girls in Fig. 3 of the Electronic Companion.

### 4.3 Treatment

If treatment is given at age  $t$ , we assume that  $X_{t+2} - X_t$  given  $X_t = x$  is normal with pdf  $h_{xt}$ , mean  $\mu_{hxt}$  and standard deviation  $\sigma_{hxt}$ , where this increment captures the natural progression in the absence of treatment plus the effect of treatment. Although there are a number of meta-analyses that assess the impact of treatment for childhood obesity, we use data from the three comprehensive moderate- or high-intensity programs considered in Table 1 of Whitlock *et al.* (2010) (Savoie *et al.* 2007, Reinehr *et al.* 2006, and Nemet *et al.* 2005), which is a targeted systematic review to support the updated USPSTF recommendations. These data provide aggregate results on the effect of treatment and do not allow us to reliably estimate the dependence of treatment on the gender, age or pre-treatment BMI level; the latter two factors are interrelated because younger children tend to have smaller BMI values (e.g., the CDC obesity threshold at 16 years old is approximately 50% larger than at 6 years old).

The approach used to model the effect of treatment in the face of limited data is both important (i.e., it will affect the qualitative conclusions of our study) and nonobvious. Given that we cannot develop detailed functions that depend on age or pre-treatment BMI level, three modeling approaches come to mind: let the treatment-induced change in (i) BMI,  $B_t$ ; (ii) transformed BMI,  $X_t$ , or (iii) z-BMI, be independent of gender, age and pre-treatment BMI level. We denote z-BMI by  $Z_t = \frac{(B_t/M_t)^{L_t} - 1}{L_t S_t}$ , where  $M_t$  is the median,  $S_t$  is the generalized coefficient of variation, and  $L_t$  is the parameter of the Box-Cox transformation (these three parameters also depend on gender and are given in Kuczmarski *et al.* (2000));  $Z_t$  can be converted into percentiles using standard normal distribution tables. We pursue option (ii), and in §3 of the Electronic Companion, we show that our treatment model accurately captures the dependence on pre-treatment BMI (albeit in a somewhat narrow range of pre-treatment levels), and successfully predicts that the change in z-BMI is larger for younger children than for older children (as shown in clinical studies by Epstein *et al.* 2007).

We estimate the impact of treatment in three steps; details are in §2 of the Electronic Companion. First, we transform the BMI data in Table 1 of Whitlock *et al.* (2010) from  $B$  to  $X$  using a Taylor series expansion around the mean of  $B$ , which gives us the mean and standard deviation of the change in  $X$  due to treatment for each of the three studies. In the second step, we compute the mean and standard deviation of the net treatment (i.e., treatment minus control) effect, pooled over the three studies. In the final step, we combine the natural progression and the net treatment effect to obtain  $\mu_{hxt}$  and  $\sigma_{hxt}$ .

## 4.4 Disease Functions

The most difficult task in this section is to estimate the four sets of disease functions ( $c_0(x_0), c_2(x_2), \dots, c_{18}(x_{18})$ ), which allow us to predict the disease (hypertension or diabetes) prevalence at age 40 in terms of  $(B_0, B_2, \dots, B_{18})$  for each gender. We use three different data sets, which measure BMI from three different sets of individuals, to estimate these functions: the NLSY data that contain  $(B_0, \dots, B_{18})$ , the PSID data (Institute for Social Research 2010), which have self-reported BMI at ages 18 and 40 and the presence or absence of hypertension and of type 2 diabetes by age 40, and the National Health and Nutrition Examination Survey (NHANES) data (National Center for Health Statistics 2006), which contain annual BMI data for ages [18,40]. The NHANES data set is cross-sectional, and to use it to infer patterns of BMI changes between ages 18 and 40 in the PSID data requires the assumption that there are no substantial cohort-specific trends.

This subsection begins with a literature review, which motivates our key assumption that determines the independent variable in our logistic regression for the prevalence of disease at age 40. Then we use the NHANES database to estimate the parameter  $\gamma$ , which tells us how to weight  $B_{18}$  and  $B_{40}$  in the PSID data so as to succinctly captures how BMI changes between ages 18 and 40. With  $\gamma$  in hand, we re-express the independent variable in terms of  $X_t$  rather than  $B_t$  to obtain the disease functions, and run four logistic regressions (for each combination of gender and disease) by stringing together sample paths from the NLSY and PSID data sets. Finally, we use the regression results to calculate the disease prevalence at age 40 for a generic set of BMI thresholds.

**Literature Review.** We briefly review the literature on how BMI throughout various stages of life impacts adult disease. A key issue - whether, after controlling for adult BMI, childhood BMI impacts adult health - has not been answered conclusively. A 55-year follow-up study found that, after controlling for adult BMI and other factors such as smoking and socioeconomic status, the risk of heart disease mortality among men increased with adolescent BMI, with many of the associated deaths occurring in their 40s (Must *et al.* 1992). While another study (Mahoney *et al.* 1996) obtained similar results, four studies found no influence on adult disease from childhood BMI after controlling for adult BMI. However, these four studies do not necessarily conflict with the finding in Must *et al.* (1992): Colditz *et al.* (1990) considered diabetes in women, Willett *et al.* (1995) considered heart disease in women, Freedman *et al.* (2001) followed up at ages 18-37, and Raitakari *et al.* (2003) followed up at ages 24-39. On a related issue, some studies have found that weight gain or BMI increase in adulthood is a strong predictor of adult disease (Wilsgaard and Arnesen

2007 and references therein).

Several studies have shown that adult disease is a j-shaped function of childhood weight, where low-weight children were undernourished, leading to adult obesity (Abraham, Collins and Nordsieck 1971, Gunnell *et al.* 1998). Another study showed that death by age 32 is a j-shaped function of  $B_{18}$ , with lung cancer and cardiovascular disease being the main cause of death for low- and high-BMI, respectively.

Obesity at a fixed age in childhood and young adulthood is associated with risk factors for heart disease at that age (e.g., Freedman *et al.* 2007). Similarly, childhood obesity is a strong predictor of young-onset type 2 diabetes (Franks *et al.* 2007), and low BMI up to two years of age followed by an early adiposity rebound is associated with diabetes in young adulthood (Bhargava *et al.* 2004). Moreover, childhood obesity is associated with fibrous plaques (Berenson *et al.* 1998) and coronary artery calcification (Mahoney *et al.* 1996) in young adults, suggesting that “the development of coronary heart disease may depend on the cumulative lifetime effects of obesity” (pg 715, Freedman *et al.* 2001).

**Key Assumption.** Due to the conceptual simplicity of the quote from Freedman *et al.* (2001) and for lack of definitive data, we estimate the disease functions under the assumption that *the prevalence of disease at age 40 depends on the average (or, equivalently, cumulative) amount of obesity throughout the first 40 years of life.* We emphasize that we cannot test this assumption directly, but rather we assume that it is true. To operationalize this assumption, we take as our measure of obesity at age  $t$  the BMI  $B_t$ . Before carrying out the remainder of these calculations, we note that the z-BMI  $Z_t$  may be a more appropriate measure of obesity than  $B_t$  due to the increase with age of the population-wide mean BMI. Consequently, we consider a variant of the key assumption where our measure of obesity is  $Z_t$  instead of  $B_t$  in §5.3, where we find that  $B_t$  is a more powerful predictor of adult disease incidence (as measured by McFadden pseudo- $R^2$ , McFadden 1973) than  $Z_t$ .

Recall that the NLYS data has  $B_t$  for  $t = 0, 2, \dots, 18$ , and the PSID data set has  $B_{18}$ ,  $B_{40}$  and  $D_{40}$  (which equals 1 or 0 in the presence or absence of disease) for a different set of individuals. While it would be natural to use  $\sum_{t=0}^{20} B_{2t}$  as a measure of the cumulative amount of obesity up to age 40, even if we string together the sample paths from the NLSY and PSID data sets, we still only have BMI data for  $(B_0, B_2, \dots, B_{18}, B_{40})$ . Hence, we use the NHANES data to approximate the BMI evolution over ages  $[18, 40]$  by a convex combination of  $B_{18}$  and  $B_{40}$ . If we

let  $p_{40} = P(D_{40} = 1)$ , then – based on our key assumption – we run a logistic regression where  $\ln\left(\frac{p_{40}}{1-p_{40}}\right)$  is the dependent variable and the average BMI throughout ages  $[0,40]$  is the independent variable. This latter quantity is given by

$$\frac{18 \sum_{t=0}^8 B_{2t} + B_{18,N}}{10} + \frac{22}{40}[\gamma B_{18,P} + (1 - \gamma)B_{40}], \quad (4)$$

where the subscripts N and P are used to denote the NLSY and PSID data sets, respectively, because we have  $B_{18}$  values from both data sets (note that in both the dependent and independent variables, we are omitting a subscript for each individual person).

**Estimating  $\gamma$ .** To find  $\gamma$  in (4), we use the NHANES data set (National Center for Health Statistics 2006), which contains annual BMI data for ages  $[18,40]$  for 3721 males and 4245 females. The mean BMI for each age is plotted in Fig. 5 in the Electronic Companion, along with the best-fitting quadratic curve. We compute  $\gamma$  by equating the area under the quadratic curve to  $22[\gamma\hat{\mu}_{f,18} + (1 - \gamma)\hat{\mu}_{f,40}]$ , where  $\hat{\mu}_{f,18}$  and  $\hat{\mu}_{f,40}$  are the endpoints of the quadratic curve. While females tend to gain weight uniformly throughout this age range (we do not exclude pregnant women from this analysis), males tend to increase their BMIs in the earlier years, leading to  $\gamma = 0.49$  for females and  $\gamma = 0.23$  for males.

**Conversion from  $B_t$  to  $X_t$ .** Note that for the purposes of solving the dynamic program, it suffices to minimize the independent variable (i.e., the independent variable in (4) can be taken as the cost function); the regression coefficients are needed to estimate  $D_{40}$  under the optimal policy, but are not needed to derive the optimal screening and treatment policy. Now that we have the independent variable in terms of  $B_t$ , we need to convert it into functions of  $X_t = \frac{(\ln B_t)^\lambda - 1}{\lambda}$  for  $t = 0, 2, \dots, 18$ . By (4), the cost function can be given by

$$\begin{aligned} & \frac{18 \sum_{t=0}^8 \exp\left[(\lambda X_{2t} + 1)^{1/\lambda}\right] + \exp\left[(\lambda X_{18,N} + 1)^{1/\lambda}\right]}{40} \\ & + \frac{22}{40} \left[ \gamma \exp\left[(\lambda X_{18,P} + 1)^{1/\lambda}\right] + (1 - \gamma) \exp\left[(\lambda X_{40} + 1)^{1/\lambda}\right] \right]. \end{aligned} \quad (5)$$

To express (5) solely in terms of  $(X_0, X_2, \dots, X_{18})$ , we use the PSID data set to write  $B_{40}$  as a quadratic function of  $B_{18}$ , which yields  $B_{40} = a_2 B_{18,N}^2 + a_1 B_{18,N} + a_0$ , where  $(a_2, a_1, a_0) = (-0.0198, 1.94, -7.27)$  for males and  $(-0.0210, 2.19, -11.19)$  for females.

By (5) and by factoring out  $\frac{1}{40}$  and dropping the data set subscript from  $X_{18}$ , the disease

functions are

$$c_t(x_t) = \frac{9}{5} \exp\left[(\lambda X_t + 1)^{1/\lambda}\right] \quad \text{for } t = 0, 2, \dots, 16, \quad (6)$$

$$\begin{aligned} c_{18}(x_{18}) &= \left(\frac{9}{5} + 22\gamma\right) \exp\left[(\lambda X_{18} + 1)^{1/\lambda}\right] \\ &+ 22(1 - \gamma) \left(a_2 \exp\left[2(\lambda X_{18} + 1)^{1/\lambda}\right] + a_1 \exp\left[(\lambda X_{18} + 1)^{1/\lambda}\right] + a_0\right), \end{aligned} \quad (7)$$

where  $(\lambda, \gamma, a_2, a_1, a_0) = (-2.15, 0.23, -0.0198, 1.94, -7.27)$  for males and  $(-2.02, 0.49, -0.0210, 2.19, -11.19)$  for females.

**Logistic Regression.** We now run the logistic regression, where the dependent variable is  $\ln\left(\frac{p_{40}}{1-p_{40}}\right)$  and the independent variable is (4). To review, the BMI values in the independent variable are  $(B_0, B_2, \dots, B_{16}, B_{18,N}, B_{18,P}, B_{40})$ , where we have  $(B_0, B_2, \dots, B_{16}, B_{18,N})$  for 1566 boys and 1598 girls from the NLSY data set, and we have  $(B_{18,P}, B_{40})$  for 279 men and 468 women from the PSID data set. Using the Markov assumption that  $B_{40}$  depends on  $(B_0, B_2, \dots, B_{18})$  only via  $B_{18}$ , we string the data sets together by randomly assigning sample paths  $(B_0, B_2, \dots, B_{18,N})$  from the NLSY data based on the value of  $B_{18,P}$  in the PSID data, although we employ bootstrapping to overcome the the limited sample size of the latter data set. More specifically, the PSID data for 747 people are sampled with replacement with sampling probability proportional to the sample weights (which describe how representative each of the 747 people are in the general population), and create a bootstrapped data set of 747,000 people. For each bootstrap sample  $(B_{18,P}, B_{40}, D_{40})$ , one NLSY sample path  $(B_0, B_2, \dots, B_{16}, B_{18,P})$  of the same gender is selected to link with it with a probability that is proportional to  $\phi\left(\frac{B_{18,P} - B_{18,N}}{0.03}\right)$ , where  $\phi(x)$  is the standard normal pdf and the standard deviation of 0.03 induces a strong distance effect while maintaining a reasonable number of NLSY samples in each selection. The results of the four logistic regressions are given in Table 12 of the Electronic Companion.

**Computing the Disease Prevalence.** Using equation (1), we simulate  $10^6$  sample paths of BMI values  $(x_0, x_2, \dots, x_{18})$  under a generic threshold policy  $(\theta_0, \theta_2, \dots, \theta_{16})$ . By construction, note that the independent variable in (4) can also be expressed as  $\frac{1}{40} \sum_{t=0}^9 c_{2t}(x_{2t})$ . For each individual, we calculate his value of the independent variable using (6)-(7), and use the regression coefficients in

Table 11 of the Electronic Companion to determine his value of  $p_{40}$ , which is

$$\frac{\exp\left(\beta_0 + \frac{\beta_1}{40} \sum_{t=0}^9 c_{2t}(x_{2t})\right)}{1 + \exp\left(\beta_0 + \frac{\beta_1}{40} \sum_{t=0}^9 c_{2t}(x_{2t})\right)}. \quad (8)$$

We compute the disease prevalence by taking the average of  $p_{40}$  over all  $10^6$  individuals (where each individual has a BMI sample path). From a computational viewpoint, this approach is much easier than integrating (8) over the ten-dimensional joint pdf of  $f_t(x)$ .

**The Surrogate Objective Function.** Note that our disease prevalence is the expected value of the quantity in (8), which is a nonseparable function of  $(f_0(x), f_2(x), \dots, f_{18}(x))$ , and hence cannot be used in the dynamic programming formulation. Instead, we minimize the separable cost function in equation (2). Although minimizing (2) does not necessarily minimize the simulated disease prevalence in (8), the simulated disease prevalence is indeed monotonic in  $\sum_{t=0}^9 \int_0^\infty c_{2t}(x) f_{2t}(x) dx$  in our numerical computations in §5.2.

## 5 Computational Results

We briefly discuss the computational aspects of the dynamic program in §5.1. The base-case results and the sensitivity analyses are presented in §5.2 and §5.3, respectively.

### 5.1 The Dynamic Program

To deal with constraint (3), we use a “finite-fuel” approach (e.g., Benes, Shepp and Witsenhausen 1980), where  $y_t = p - (\sum_{s=1}^{t/2-1} \int_{\frac{(\ln \theta_{2s})^{\lambda-1}}{\lambda}}^\infty f_{2s}(x) dx)/8$  is the slack in the treatment prevalence from age  $t$  to age 16, and augment the system state to be  $(f_t(x), y_t)$ . If we define  $V_t(f_t(x), y_t)$  to be the minimum remaining cost until age 18 when the state at age  $t$  is  $(f_t(x), y_t)$ , then the optimality equation for  $t = 2, 4, \dots, 16$  is

$$V_t(f_t(x), y_t) = \min_{\theta_t} \left\{ \int_0^\infty c_t(x) f_t(x) dx + V_{t+2} \left( f_{t+2}(x), y_t - \frac{1}{8} \int_{\frac{(\ln \theta_t)^{\lambda-1}}{\lambda}}^\infty f_t(x) dx \right) \right\}, \quad (9)$$

and

$$V_{18}(f_{18}(x), y_{18}) = \int_0^\infty c_{18}(x) f_{18}(x) dx, \quad (10)$$

where the minimization is subject to the finite-fuel constraint  $y_{18} \geq 0$ , and the state transitions for  $f_t(x)$  are defined in (1). Note that the optimal thresholds are independent of disease (i.e., hypertension vs. diabetes) because the disease prevalence for both diseases is an increasing function of the objective function in equation (2).

Although our dynamic program is deterministic, its continuous state (i.e.,  $f_t(x)$ ) makes for unwieldy calculations, and we have been unable to solve for the optimal thresholds for more than three ages (i.e., we can only solve a three time-period dynamic program). Consequently, we also use an approximate dynamic programming algorithm, more specifically an extension of the rollout algorithm that is specifically tailored to our setting of discrete-time constrained deterministic dynamic programming problems (Bertsekas 2005). The algorithm, specified in §2 of Bertsekas (2005), relies on a suboptimal policy called the base heuristic, upon which it improves, and we use the optimal (14,16) policy as the base heuristic.

## 5.2 Base-case Results

In this subsection, we solve the dynamic program for a wide variety of treatment prevalences, and compare these policies to a variant of the USPSTF recommendation. Recall that the USPSTF recommends that children be screened and treated for obesity starting at age six (USPSTF 2010). Although the USPSTF does not explicitly recommend a screening frequency, they suggest that screening be done during routine pediatric assessments, which typically occur every year or two. We adapt their recommendation to biennial screening, and refer to the biennial USPSTF policy as the one that sets the BMI thresholds  $\theta_t$  equal to the 95<sup>th</sup> percentile of the BMI distributions in the CDC growth charts (Kuczmarski *et al.* 2000) for ages  $t = 6, 8, \dots, 16$  (and sets  $\theta_t = \infty$  for  $t = 0, 2, 4$ ). It is important to remember that the CDC BMI distributions do not reflect the BMI distributions in the NLSY data set: obesity has tripled in recent decades and approximately 15% of children in the NLSY data set exceed the 95<sup>th</sup> percentile of the CDC BMI distributions.

Because we have been able to solve the dynamic program for up to only three time periods, we present results that progressively restrict the number of ages that are tested to one, two and three, respectively. The main results are presented as tradeoff curves in Fig. 1 as well as Figs. 6 and 7 in the Electronic Companion, respectively, with selected numerical results appearing in Table 1.

Fig. 6 of the Electronic Companion has eight tradeoff curves that test at a single age, 2, 4,  $\dots$ , 16. For each gender and disease, performance improves with increasing age (except for males

at age 8). Fig. 7 of the Electronic Companion displays tradeoff curves for three two-age screening policies, at ages (6,16), (10,16) and (14,16). Screening at ages (14,16) is the best among these three policies. Fig. 1 reveals that the three-age policy using ages (12,14,16) performs only marginally better than the optimal two-age policy.

The optimal (12,14,16) policy in Table 1 has two potential sources of suboptimality: it uses the sum of BMIs as its separable objective function, rather than the disease prevalence computed by simulation via equation (8), and the dynamic programming solution only screens at three ages. We perform two other calculations to assess these suboptimalities. First, an exhaustive search of the (14,16) policy using the simulation model yields exactly the same solution as the optimal (14,16) policy in Table 1 (however, an exhaustive search for the optimal three-stage policy is extremely onerous from a computational standpoint). Second, while the approximate dynamic programming procedure based on Bertsekas (2005) generates a policy that treats at more than three ages, it generates nearly the same disease prevalences as the optimal (12,14,16) policy (Table 1). Taken together, all of these results strongly suggest that the optimal (12,14,16) policy is nearly optimal using the simulation in equation (8) among the class of all threshold policies (i.e., allowing screening at ages 2,4,...,16) for our set of parameter values.

Overall, the optimal tradeoff curves are only slightly convex and are rather flat. The flatness of the tradeoff curves leads to different qualitative conclusions regarding the magnitude of the performance gap between the optimal three-age policy and the biennial USPSTF policy (Fig. 1 and Table 1): the optimal three-age policy provides a 3% mean relative reduction in disease prevalence (2.0% for male hypertension, 3.2% for male diabetes, 4.4% for female hypertension, and 4.3% for female diabetes) at the same screening prevalence, or a 28% mean relative reduction in screening prevalence at the same disease prevalence.

Relative to no obesity screening, the optimal three-age policy with a screening prevalence equal to that of the biennial USPSTF policy achieves a disease prevalence relative reduction of 8.1% for male hypertension, 13.8% for male diabetes, 17.8% for female hypertension, and 21.3% for female diabetes. To calibrate the screening prevalence, we note that if the thresholds were set at the overweight level (i.e., 85<sup>th</sup> percentile of the CDC distributions) for ages 6, 8, ..., 16, then the treatment prevalence is 0.2158 for boys and 0.1930 for girls. As an upper bound, the optimal three-age policy with a screening prevalence equal to 0.25 reduces disease prevalence (compared to no screening) by 15.6% for male hypertension, 24.5% for male diabetes, 35.0% for female hypertension,

and 37.0% for female diabetes.

The optimal multi-age policies achieve their improvements over the biennial USPSTF policy by being less selective – as measured by the threshold’s corresponding percentile in the CDC BMI distributions – with increasing age; the same is not always true for the actual thresholds, e.g., the optimal BMI threshold is smaller at age 14 than age 16 in the optimal (14,16) female policy. The majority ( $\approx 51\%$  for females and  $\approx 70\%$  for males) of treatment occurs at 16 years old (including both males and females who fall under the 85<sup>th</sup> CDC percentile, which is the overweight threshold), and very few ( $\approx 1.7\%$ ) males under 14 years old are treated (Fig. 2).

### 5.3 Sensitivity Analyses

In this subsection, we carry out four sensitivity analyses: we allow treatment efficacy to be dependent on age and pre-treatment BMI, and we consider two variants of the key assumption: disease prevalence depends on average z-BMI rather than on average BMI, and disease prevalence at age 40 depends on childhood BMI only via  $B_{18}$ .

**Age-dependent Treatment Efficacy.** To assess the robustness of our results with respect to age-dependent treatment efficacy, we assume that treatment is more effective for children ages 6,8,10 than for children ages 12,14,16, and we consider the ratio of the change in untransformed BMI due to treatment for children ages 6,8,10 divided by the change in untransformed BMI due to treatment for children ages 12,14,16. This ratio is computed by replacing the mean net treatment effect,  $\mu_T$  (-0.00183 for males and -0.00214 for females in §4.3), by  $\mu_T - d$  for ages 6,8,10 and  $\mu_T + d$  for ages 12,14,16, and choosing  $d$  to achieve the desired ratio, after converting from  $X_t$  to  $B_t$  and averaging over the top 15% of NLSY BMI values for each age, which is the subset of children who are apt to receive treatment. Under this construction, the mean and variance of the treatment effect for all children is independent of  $d$ . Note that when  $d = 0$ , which corresponds to our base case, this ratio is 0.61 for females and 0.62 for males (Fig. 4 of the Electronic Companion).

We now consider two policies, the optimal two-stage (10,16) policy and the optimal two-stage (14,16) policy, and compute the breakeven value of the ratio in the previous paragraph that equates the performance of these two policies. The breakeven ratio is 0.92 for males and 0.85 for females. If instead we compare the biennial USPSTF policy and the optimal (14,16) policy, the breakeven values are 1.03 for males and 0.94 for females, which are approximately 60% higher than the base-case values.

**BMI-dependent Treatment Efficacy.** We perform a similar sensitivity analysis as above, but assume bigger treatment effects occur for children with higher pre-treatment BMI levels rather than for younger children. We let the net treatment effect be  $\mu_T - 0.441d$  for males ( $\mu_T - 0.856d$  for females) with pre-treatment BMI levels above the 97<sup>th</sup> CDC percentile, and be  $\mu_T + d$  for those below the 97<sup>th</sup> percentile (0.441 and 0.856 maintain the mean treatment effect at 0 for the top 15% of NLSY BMI values), where there is a one-to-one correspondence between  $d$  and the ratio of the change in untransformed BMI due to treatment for children above the 97<sup>th</sup> percentile divided by the change in untransformed BMI due to treatment for children between the 85<sup>th</sup> NLSY percentile and the the 97<sup>th</sup> CDC percentile. When  $d = 0$ , this ratio is 1.54 for males and females (Fig. 4 of the Electronic Companion). We find that the breakeven ratio that equates the performance of the biennial USPSTF policy and the optimal (14,16) policy is 5.86 for males and 5.14 for females.

**Obesity Measured as z-BMI Rather Than BMI.** Here, we re-estimate the disease functions in §4.4, but assume that obesity at age  $t$  is measured by the z-BMI  $Z_t$  rather than  $B_t$ . We set  $\gamma = 0.5$  (which generates higher pseudo- $R^2$  values than the  $\gamma$  values used earlier) and replace  $B_t$  by  $Z_t$  in equation (4). Because the CDC charts are not tabulated through age 40, to compute the values of  $M_{40}, S_{40}$  and  $L_{40}$ , we estimate these parameters from the PSID age-40 data, and then modify them by the ratios of the age-18 parameters derived from the CDC charts divided by the age-18 parameters derived from the PSID data. The new McFadden pseudo- $R^2$  values are (0.020,0.032,0.100,0.080), which are smaller than the values in Table 11 of the Electronic Companion. Overall, the results are qualitatively similar to the base-case results, in that the optimal (12,14,16) policy provides a 3% reduction in disease prevalence relative to the biennial USPSTF policy. The main difference is that the optimal single-age policy (i.e., test only at age 16) is nearly optimal compared to the policy that screens at ages 12,14,16.

**Adult Disease Independent of BMI Before Age 18.** Finally, we forego the key assumption that disease prevalence at age 40 depends on the cumulative lifetime obesity, and consider the extreme case in which the BMI at age 18 is the sole childhood factor (aside from gender) that affects the disease prevalence at age 40. More specifically, the objective function of our dynamic program is changed from (2) to minimizing  $\int_0^\infty x f_{18}(x) dx$ . Not surprisingly, this change causes a shift towards more treatment at age 16 and less treatment at earlier ages. For example, in Table 1 the CDC threshold percentiles for the optimal (14,16) policy for males are 90.8 for age 14 and 74.1 for age 16, and they change to 96.1 for age 14 and 62.1 for age 16 under the new objective function.

For females, the (14,16) percentiles change from (85.7,81.6) to (93.9,72.8).

## 6 Discussion

Our primary result is that – relative to a biennial version of the USPSTF policy, the optimal policy is predicted to achieve a 3% reduction in hypertension or diabetes prevalence at age 40 at the same cost, or achieve the same disease prevalence at a 28% reduction in cost. While the USPSTF biennial policy sets thresholds at the 95<sup>th</sup> percentile of the CDC BMI distributions for ages 6,8,...,16, the optimal policy treats mostly 16 year olds and few children under age 14. This is largely due to the high false positive rate associated with treating younger children (Goldhaber-Fiebert *et al.* 2010); i.e., in the absence of treatment, many children who are obese at age 10 or younger are no longer obese at age 18.

Childhood obesity screening and treatment is a prime candidate for mathematical modeling and analysis because of the impracticality of running 50-year clinical trials that link childhood BMI to adult disease. The USPSTF policy has not been previously analyzed with respect to suboptimality or supply-side feasibility (over 5M children per year would be referred for treatment under the USPSTF policy), and our study is intended to inform the debate on the appropriate screening policy for childhood obesity. Because our model captures only some of the aspects of this very complex issue, our results do not imply that the USPSTF policy should be overturned. Rather, they suggest that other arguments, e.g., the social and psychological effects of childhood obesity, younger children are predicted to have higher treatment compliance, or that treatment also instills other good lifestyle choices (e.g., no smoking) that provide health benefits later in life (Magnussen *et al.* 2011), or that screening earlier would prevent type 2 diabetes in childhood (the prevalence of type 2 diabetes in childhood is on the order of 0.1%, and is concentrated in Native American populations and in minority children with a family history of diabetes, Fagot-Campagna *et al.* 2000), would be needed to justify the policy recommendations of the USPSTF. Indeed, our sensitivity analysis in §5.3 provides a crude assessment for what these arguments would have to look like: the ratio of the absolute value of BMI reduction due to treatment for ages 6,8,10 divided by the absolute value of BMI reduction due to treatment for ages 12,14,16 needs to increase by 60% relative to our base-case assumptions in order for the optimal biennial policy to achieve the same performance as the optimal (14,16) policy.

In our view, the total amount of screening to perform (i.e., setting the treatment prevalence

$p$  in equation (3)) is not clear cut because the tradeoff curve is only modestly convex: the absolute value of the slope of the optimal tradeoff curve in Fig. 1, which is a measure of cost-effectiveness, decreases by an average of 72% as the treatment prevalence increases from 0 to 0.25. At the treatment prevalence of the biennial USPSTF policy, the optimal (12,14,16) policy treats 51.1% of 16-year-old males in the NLSY data set. If this magnitude of intervention was desired, universal intervention may make more practical sense than screening over half of 16-year-old males. Goldhaber-Fiebert *et al.* (2010) predicts that universal intervention provides greater health benefits than screening-guided interventions, although cost-effectiveness is not considered. See Cecchini *et al.* (2010) for a study that compares the cost-effectiveness of various obesity interventions.

**Limitations of the Study.** The stochastic model for the evolution of the population-wide BMI pdf captures the salient features of the NLSY data and appears to be at the appropriate level of complexity for the strategic issues addressed here. Perhaps the most unrealistic aspect of the model formulation is that it implicitly assumes 100% treatment compliance for all children who screen positive. In practice, many families will refuse treatment for their children, even if the treatment is free. Moreover, compliance among those who initiate treatment may be lower than the compliance achieved in clinical trials. However, if compliance is independent of pre-treatment BMI level and age, then the model (in particular, equation (1)) is easily modified to incorporate imperfect compliance. Extensive numerical results in Wein, Yang and Goldhaber-Fiebert (2011) suggest that the relative reduction in disease prevalence (at the same treatment prevalence) under imperfect compliance is slightly less than the compliance rate times the relative reduction in the 100% compliance case, and the relative reduction in treatment prevalence (at the same disease prevalence) is 15-30% less than it is under 100% compliance.

While the NLSY BMI data by and large follow the Markov property in the absence of treatment, our model also assumes that the Markov property holds after treatment. That is, if a child receives treatment at age  $t$  and consequently has BMI  $b_{t+2}$  (a realization of the random variable  $B_{t+2}$ ) two years later, then his future BMI trajectory is statistically identical to a child who has never received treatment but who also has BMI  $b_{t+2}$  at age  $t+2$ . Data suggest that some treatment effects are transient (Oude Luttikhuis *et al.* 2009), while others are not (Epstein *et al.* 2007), and this assumption deserves additional attention in the future. In addition, we may be overestimating the impact of multi-age policies if an individual's efficacies across different treatments are positively correlated: we treat them as iid, and people with lower efficacy are more apt to exceed the

threshold in subsequent years. In this regard, it is worth noting that the optimal single-age policy outperforms the biennial USPSTF policy (except for diabetes in women) and is nearly optimal if obesity is measured by z-BMI.

Most of the limitations of the study are due to the paucity of data rather than the structure of the model. The data include BMI evolution, the effect of treatment on BMI, and the effect of BMI on disease. The BMI evolution data are plentiful, but somewhat dated (and also partially self-reported, which may understate BMI levels): the NLSY data cover children born between 1970 and 1988, and the PSID data used for this analysis cover adults born in the late 1960s and early 1970s. Childhood obesity has tripled since 1980, but has leveled off in the last decade with the exception of the far right tail of the BMI distribution for boys (Ogden *et al.* 2010), suggesting that our results underestimate the fraction of children who would screen positive. In addition, although our model looks at each gender separately, the NLSY data set is not sufficiently large to look at other (e.g., race or socioeconomic) subgroups.

Given that it is only very recently that treatment has been shown to be moderately effective at reducing BMI (Whitlock *et al.* 2010), it is not surprising that the data on the effect of treatment on BMI are sparse. Indeed, while our model can accommodate efficacy that varies by gender, age and pre-treatment BMI, the existing data do not allow us to take full advantage of this. In the near term, this – along with compliance data – probably represents the most pressing data requirement. Despite this lack of data, we assume that the effect of treatment on the transformed BMI,  $X_t$ , is independent of age and pre-treatment BMI level. This assumption forces the reduction in untransformed BMI due to treatment to increase with pre-treatment BMI, which fits the data from the high-intensity programs in Table 1 of Whitlock *et al.* (2010) (Fig. 4 of the Electronic Companion). This assumption also leads to a larger z-BMI reduction for ages 6,8,10 than for ages 12,14,16, which is consistent with the studies summarized in Epstein *et al.* (2007). In addition, our sensitivity analysis in §5.3 shows that our qualitative conclusions are insensitive to the possibility that children above the 97<sup>th</sup> percentile of the CDC BMI distributions achieve significantly larger BMI reductions than those below the 97<sup>th</sup> percentile, or if, in our key assumption, obesity is measured by z-BMI rather than BMI.

The data on the lifetime effects of BMI on adult disease are by far the most difficult to obtain, perhaps requiring a longitudinal study of at least five decades. In the absence of additional studies on this topic, we believe that it is prudent to employ our key assumption that adult disease depends

on the average lifetime obesity. Our sensitivity analysis in §5.3 on this issue shows that our key assumption is conservative with respect to the argument that the optimal policy outperforms the biennial USPSTF policy.

**Conclusion.** While additional treatment efficacy and compliance data would allow us to refine our recommendations, and the lifetime effects of obesity on adult disease are not well understood, our results suggest that the most effective way to reduce adult disease via childhood obesity screening and treatment is by focusing on older adolescents. Although our policy recommendations are at odds with the USPSTF recommendations, we feel that our main contribution may not be in the policy recommendations, but in the problem framing and modeling: viewing the issue as a constrained optimization problem of minimizing adult disease prevalence subject to a treatment budget, modeling the evolution of the population-wide distribution of BMI, investigating the Markov behavior of BMI and the normality of BMI and its increments, the modeling of treatment with a discussion of the three modeling options, and the key assumption underlying the calculation of the adult disease prevalence. More specifically, we find that – for our set of parameter values – the solution to the dynamic program, which assumes a Markov BMI process, Box-Cox normality of the BMI increments, and a separable objective function, appears to be very similar to the optimal solution for the more complex simulation model. While our modeling choices may not all stand the test of time, they do provide a basis for discussion and a framework on which to build.

## References

- [1] Abraham, S., Collins, G., Nordsieck, M. Relationship of childhood weight status to morbidity in adults. *HSMHA Health Reports*, **86**, 273-284, 1971.
- [2] Barrett-Connor, E. Epidemiology, obesity, and non-insulin-dependent diabetes mellitus. *Epidemiologic Reviews* **11**, 172-181, 1989.
- [3] Benes, V. E., Shepp, L. A., Witsenhausen, H. S. Some solvable stochastic control problems. *Stochastics* **4**, 39-83, 1980.
- [4] Berenson, G. E., Srinivasan, S. R., Bao, W., Newman III, W. P., Tracy, R. E., Wattigney, W. A. Association between multiple cardiovascular risk factors and atherosclerosis in children and young adults. *NEJM* **338**, 1650-1656, 1998.

- [5] Bertsekas, D. P. Rollout algorithms for constrained dynamic programming. Technical Report LIDS-2646, MIT, Cambridge, MA. Accessed at [http://web.mit.edu/dimitrib/www/Rollout\\_Constrained.pdf](http://web.mit.edu/dimitrib/www/Rollout_Constrained.pdf) on February 15, 2010.
- [6] Bhargava, S. K., Sachdev, H. S., Fall, C. H. D., Osmond, C., Lakshmy, R., Barker, D. J. P. *et al.* Relation of serial changes in childhood body-mass index to impaired glucose tolerance in young adulthood. *NEJM* **350**, 865-875, 2004.
- [7] Box, G. E. P., Cox, D. R. An analysis of transformations. *JRSS B* **26**, 211-252, 1964.
- [8] Bureau of Labor Statistics. National longitudinal surveys: NLSY79 children and young adults. U.S. Department of Labor, Washington D.C., Accessed at [www.bls.gov/nls/nlsy79ch.htm](http://www.bls.gov/nls/nlsy79ch.htm) on August 12, 2010.
- [9] Cecchini, M., Sassi, F., Lauer, J. A., Lee, Y. Y., Guajardo-Barron, V., Chisholm, D. Tackling of unhealthy diets, physical inactivity, and obesity: health effects and cost-effectiveness. *Lancet* **376**, 1775- 1784, 2010.
- [10] Colditz, G. A., Willett, W. C., Stampfer, M. J., Manson, J. E., Hennekens, C. H., Arky, R. A., Speizer, F. E. Weight as a risk factor for clinical diabetes in women. *Am. J. Epidemiology* **132**, 501-513, 1990.
- [11] Coleman, K. J., Tiller, C. L., Sanchez, J., Heath, E. M., Sy, O., Milliken, G., Dzewaltowski, D. A. Prevention of the epidemic increase in child risk of overweight in low-income schools. *Arch. Pediatric Adolescent Medicine* **159**, 217-224, 2005.
- [12] D'Amato, R., D'Aquila, R., Wein, L. M. Management of antiretroviral therapy for HIV infection: analyzing when to change therapy. *Management Science* **50**, 1200-1213, 2000.
- [13] Epstein, L. H., Paluch, R. A., Roemmich, J. N., Beecher, M. D. Family-based obesity treatment, then and now: twenty-five years of pediatric obesity treatment. *Health Psychology* **26**, 381-391, 2007.
- [14] Fagot-Campagna, A., Pettitt, D. J., Engelgau, M. M., Burrows, N. R., Geiss, L. S., Valdez, R. *et al.* Type 2 diabetes among North American children and adolescents: an epidemiologic review and a public health perspective. *J. Pediatrics* **136**, 664-672, 2000.

- [15] Finkelstein, E., Trogon, J., Cohen, J., Dietz, W. Annual medical spending attributable to obesity: payer- and server-specific estimates. *Health Affairs* **28**, w822-w831, 2009.
- [16] Franks, P. W., Hanson, R. L., Knowler, W. C., Moffett, C., Enos, G., Infante, A. M., Krakoff, J., Looker, H. C. Childhood predictors of young-onset type 2 diabetes. *Diabetes* **56**, 2964-2972, 2007.
- [17] Freedman, D. S., Khan, L. K., Dietz, W. H., Srinivasan, S. R., Berenson, G. S. Relationship of childhood obesity to coronary heart disease risk factors in adulthood: the Bogalusa Heart Study. *Pediatrics* **108**, 712-718, 2001.
- [18] Freedman, D. S., Mei, Z., Srinivasan, S. R., Berenson, G. S., Dietz, W. H. Cardiovascular risk factors and excess adiposity among overweight children and adolescents: the Bogalusa Heart Study. *J. Pediatrics* **150**, 12-17, 2007.
- [19] Garber, A. M., Olshen, R. A., Zhang, H., Venkatraman, E. S. Predicting high-risk cholesterol levels. *International Statistical Review* **62**, 203-228, 1994.
- [20] Goldhaber-Fiebert, J. D., Wise, P., Bhattacharya, J., Rubinfeld, R. U.S. childhood obesity policies and their effect on adult health through 2040. Centers for Health Policy and Primary Outcomes Research, Stanford University School of Medicine, Stanford, CA 2010.
- [21] Gortmaker, S. L., Peterson, K., Wiecha, J., Sobol, A. M., Dixit, S., Fox, M. K., Laird, N. Reducing obesity via a school-based interdisciplinary intervention among youth. *Arch. Pediatric Adolescent Medicine* **153**, 409-418, 1999.
- [22] Gunnell, D. J., Frankel, S. J., Nanchahal, K., Peters, T. J., Smith, G. D. Childhood obesity and adult cardiovascular mortality: a 57-y follow-up study based on the Boyd Orr cohort. *Am. J. Clinical Nutrition* **67**, 1111-1118, 1998.
- [23] The HEALTHY Study Group. A school-based intervention for diabetes risk reduction. *NEJM* **363**, 443-453, 2010.
- [24] Hubert, H. B., Feinleib, M., McNamara, P. M., Castelli, W. P. Obesity is an independent risk factor for cardiovascular disease: a 26-year followup of participants in the Framingham Heart Study. *Circulation* **67**, 968-977, 1983.

- [25] Institute for Social Research. Panel study of income dynamics. University of Michigan, accessed at [psidonline.irs.umich.edu](http://psidonline.irs.umich.edu) on August 12, 2010.
- [26] Institute of Medicine. *Preventing childhood obesity: health in the balance*. National Academies Press, Washington D.C. 2005.
- [27] Kuczmarski, R. J., Ogden, C. L., Grummer-Strawn, L. M., Flegal, K. M., Guo, S. S., Wei, R., Mei, Z., Curtin, L. R., Roche, A. F., Johnson, C. L. CDC growth charts: United States. *Advance Data* **314**, 1-27, December 4, 2000.
- [28] Lavieri, M., Puterman, M., Tyldesley, S., Morris, W. When to treat prostate cancer patients based on their PSA dynamics. Working paper, Sauder School of Business, University of British Columbia, Vancouver, BC, Canada, 2009.
- [29] Lee, C. P., Chertow, G. M., Zenios, S. A. Optimal initiation and management of dialysis therapy. *Operations Research* **56**, 1428-1449, 2008.
- [30] Magnussen, C. G., Thomson, R., Cleland, V. J., Ukoumunne, O. C., Dwyer, T., Venn, A. Factors affecting the stability of blood lipid and lipoprotein levels from youth to adulthood. *Arch. Pediatr. Adolesc. Med.* **165**, 68-76, 2011.
- [31] Mahoney, L. T., Burns, T. L., Stanford, W., Thompson, B. H., Witt, J. D., Rost, C. A., Lauer, R. M. Coronary risk factors measured in childhood and young adult life are associated with coronary artery calcification in young adults: the Muscatine Study. *J. Am. College Cardiology* **27**, 277-284, 1996.
- [32] McFadden, D. Conditional logit analysis of qualitative choice behavior. In *Frontiers in Econometrics*, Zarembka, P. ed., pp. 105-142, Academic Press, New York, 1973.
- [33] Mission: Readiness - Military Leaders for Kids. *Too fat to fight: retired military leaders want junk food out of America's schools*. Mission: Readiness, New York, 2010.
- [34] Must, A., Jacques, P. F., Dallal, G. E., Bajema, C. J., Dietz, W. H. Long-term morbidity and mortality of overweight adolescents: a follow-up of the Harvard Growth Study of 1922 to 1935. *NEJM* **327**, 1350-1355, 1992.

- [35] National Center for Health Statistics. Analytic and reporting guidelines: The National Health and Nutrition Examination Survey (NHANES). Centers for Disease Control and Prevention, Hyattsville, MD, 2006.
- [36] Nemet, D., Barkan, S., Epstein, Y., Friedland, O., Kowen, G., Eliakim, A. Short- and long-term beneficial effects of a combined dietary-behavior-physical activity intervention for the treatment of childhood obesity. *Pediatrics* **115**, e443-e449, 2005.
- [37] Nihiser, A. J., Lee, S. M., Wechsler, H., McKenna, M., Odom, E., Reinold, C., Thompson, D., Grummer-Strawn, L. Body mass index measurement in schools. *J. School Health* **77**, 651-671, 2007.
- [38] Ogden, C. L., Carroll, M. D., Curtin, L. R., Lamb, M. M., Flegal, K. M. Prevalence of high body mass index in US children and adolescents, 2007-2008. *JAMA* **303**, 242-249, 2010.
- [39] O'Hagan, A., Leonard, T. Bayes estimation subject to uncertainty about parameter constraints. *Biometrika* **63**, 201-202, 1976.
- [40] Oude Luttikhuis, H., Baur, L., Jansen, H., Shrewsbury, V. A., O'Malley, C., Stolk, R. P., Summerbell, C. D. Interventions for treating obesity in children (Review). *The Cochrane Library*, Issue 1, John Wiley & Sons, New York, 2009.
- [41] Raitakari, O. T., Juonala, M., Kahonen, M., Taittonen, L., Laitinen, T., Maki-Torkko, N., Jarvisalo, M. J., Uhari, M., Jokinen, E., Ronnema, T., Akerblom, H. K., Viikari, J. S. a. Cardiovascular risk factors in childhood and carotid artery intima-media thickness in adulthood: the Cardiovascular Risk in Young Finns Study. *JAMA* **290**, 2277-2283, 2003.
- [42] Reilly, J. J., Dorosty, A. R., Emmett, P. M., The ALSPAC Study Team. Identification of the obese child: adequacy of the body mass index for clinical practice and epidemiology. *International Journal Obesity* **24**, 1623-1627, 2000.
- [43] Reinehr, T., de Sousa, G., Toschke, A. M., Andler, W. Long-term follow-up of cardiovascular disease risk factors in children after an obesity intervention. *American J. Clinical Nutrition* **84**, 490-496, 2006.

- [44] Reinehr, T., Temmesfeld, M., Kersting, M., de Sousa, G., Toschke, A. M. Four-year follow-up of children and adolescents participating in an obesity intervention program. *International J. Obesity* **31**, 1074-1077, 2007.
- [45] Savoye, M., Shaw, M., Dziura, J., Tamborlane, W. V., Rose, P., Guandalini, C., Goldberg-Gell, R., Burgert, T. S., Cali, A. M. G., Weiss, R., Caprio, S. Effects of a weight management program on body composition and metabolic parameters in overweight children. *JAMA* **297**, 2697-2704, 2007.
- [46] Shechter, S. M., Bailey, M. D., Schaefer, A. J., Roberts, M. S. The optimal time to initiate HIV therapy under ordered health states. *Operations Research* **56**, 20-33, 2008.
- [47] Singh, A. S., Mulder, C., Twisk, J. W., van Mechelen, W., Cinapaw, M. J. Tracking of childhood overweight into adulthood: a systematic review of the literature. *Obesity Reviews* **9**, 474-488, 2008.
- [48] US Preventive Services Task Force. Screening and interventions for overweight in children and adolescents: recommendation statement. *Pediatrics* **116**, 205-209, 2005.
- [49] US Preventive Services Task Force. Screening for obesity in children and adolescents: US Preventive Services Task Force recommendation statement. *Pediatrics* **125**, 361-367, 2010.
- [50] Wein, L. M., Yang, Y., Goldhaber-Fieber, J. D. Assessing screening policies for childhood obesity. Graduate School of Business, Stanford University, Stanford, CA, 2011.
- [51] White House Task Force on Childhood Obesity. Solving the problem of childhood obesity within a generation. Domestic Policy Council, Washington D.C., 2010.
- [52] Whitlock, E. P., O'Connor, E. A., Williams, S. B., Beil, T. L., Lutz, K. W. Effectiveness of weight management interventions in children: a targeted systematic review for the USPSTF. *Pediatrics* **125**, e396-e418, 2010.
- [53] Willett, W. C., Manson, J. E., Stampfer, M. J., Colditz, G. A., Rosner, B., Speizer, F. E., Hennekens, C. H. Weight, weight change, and coronary heart disease in women: risk within the "normal" weight range. *JAMA* **273**, 461-465, 1995.
- [54] Wilsgaard, T., Arnesen, E. Body mass index and coronary heart disease risk score: the Tromso Study, 1979 to 2001. *Annals Epidemiology* **17**, 100-105, 2007.

Policy	Prevalences for Males			Prevalences for Females		
	Treatment	Hypertension	Diabetes	Treatment	Hypertension	Diabetes
No screening	0	0.2214	0.0636	0	0.2350	0.0568
Biennial USPSTF	0.1027	0.2076	0.0566	0.0850	0.2020	0.0467
Optimal: age 16	0.1027	0.2065	0.0566	0.0850	0.2003	0.0474
Optimal: ages (6,16)	0.1027	0.2050	0.0559	0.0850	0.2003	0.0474
Optimal: ages (10,16)	0.1027	0.2049	0.0558	0.0850	0.1987	0.0467
Optimal: ages (14,16)	0.1027	0.2035	0.0548	0.0850	0.1941	0.0451
Optimal: ages (12,14,16)	0.1027	0.2035	0.0548	0.0850	0.1931	0.0447
Approximate DP: all ages	0.1027	0.2035	0.0548	0.0850	0.1930	0.0446
Optimal: ages (12,14,16)	0.2500	0.1869	0.0480	0.2500	0.1527	0.0358

Table 1: For a variety of policies, the treatment prevalence from (3) and the hypertension and diabetes prevalences at age 40. The optimal policies are derived from the dynamic program using the listed ages, the exhaustive search policy uses only the simulation model, and the approximate DP policy uses the algorithm in Bertsekas (2005).

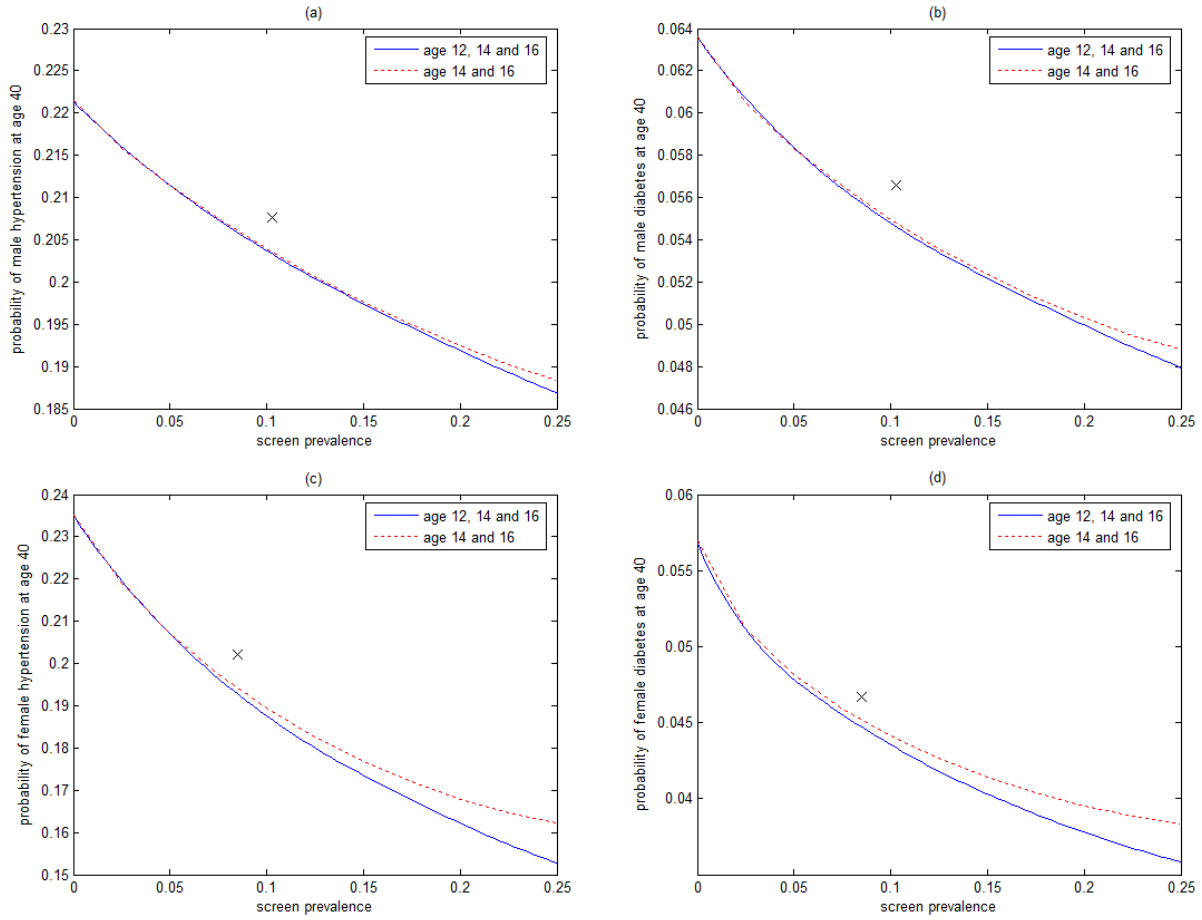


Figure 1: Tradeoff curves of disease prevalence at age 40 vs. treatment prevalence (equation (3)) for the optimal three-age screening policy for ages (12,14,16) and the optimal two-age screening policy for ages (14,16) from Fig. 7 of the Electronic Companion. The x denotes the biennial version of the USPSTF recommendation. **(a)** hypertension among males; **(b)** diabetes among males; **(c)** hypertension among females; and **(d)** diabetes among females.

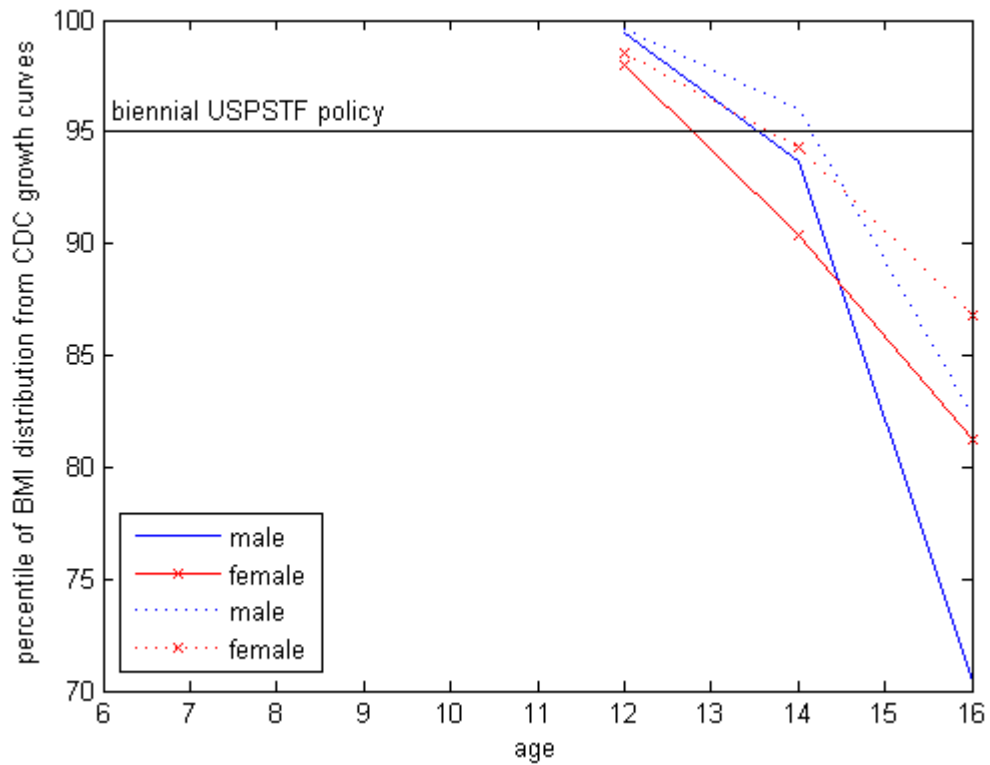


Figure 2: Optimal screening thresholds for the three-age screening policy at ages (12,14,16), expressed as the percentile of the gender- and age-based BMI distributions tabulated by the CDC (Kuczmarski *et al.* 2000), for two cases: the screening prevalence equals that of the biennial USPSTF policy (— for males and -x- for females) and the disease prevalence equals that of the biennial USPSTF policy ( $\cdots$  for males and  $\cdot x \cdot$  for females). The biennial USPSTF policy uses the 95<sup>th</sup> percentile for ages 6,8, $\dots$ ,16.