Taxation of unhealthy food consumption and the intensive versus extensive margin of obesity

Zarko Y. Kalamov[1] · Marco Runkel[2]

Abstract
If an individual's health costs are U-shaped in weight with a minimum at some healthy level and if the individual has both self-control problems and rational motives for over- or underweight, the optimal paternalistic tax on calorie intake mitigates the individual's weight problem (intensive margin), but does not induce the individual to choose healthy weight (extensive margin). Implementing healthy weight by a calorie tax is not only inferior to paternalistic taxation, but may even be worse than not taxing the individual at all. With heterogeneous individuals, the optimal uniform paternalistic tax may have the negative side effect of reducing calorie intake of the under- and normal weights. We confirm these theoretical insights by an empirical calibration to US adults.

Keywords Sin tax · Paternalism · Obesity · Extensive versus intensive margin

JEL Classifications: D03 · D11 · H21 · I18

1 Introduction

Obesity currently represents one of the most pestering health problems worldwide. For example, in the USA more than one third (38.2%) of the population aged 15 years and older is obese, i.e., has a body mass index (BMI) larger than 30 kg/m^2. The average obesity rate in OECD countries is 19.5%. If overweight people with a BMI between 25 and 30 kg/m^2 are included, then even one in two adults and one

✉ Zarko Y. Kalamov
zarko.y.kalamov@tu-berlin.de

Marco Runkel
marco.runkel@tu-berlin.de

1 Faculty of Economics and Management, University of Technology Berlin, CESifo, Straße des 17. Juni 135, H51, 10623 Berlin, Germany

2 Faculty of Economics and Management, University of Technology Berlin, CESifo, NoCeT, Straße des 17. Juni 135, H51, 10623 Berlin, Germany

in six children have weight problems in OECD countries (OECD 2017). It is well known that obesity and overweight are closely related to secondary disorders—like diabetes, cardiovascular diseases or even some cancers—which considerably increase mortality (WHO 2017). As a countermeasure, many countries have implemented taxes on calorie intake like taxes on fat, sugar or soda (WHO 2016a). In practice, the stated aim of such 'sin taxes' is to reduce the prevalence of obesity. For example, the WHO (2016b) argues that "… [e]vidence shows that a tax of 20% on sugary drinks can lead to a reduction in consumption of around 20%, thus *preventing* obesity and diabetes." (italics ours) In theory, sin taxes are often motivated by the paternalistic approach of correcting self-control problems (e.g., O'Donoghue and Rabin 2003, 2006). People choose unhealthy diet since they underestimate their future health costs. In doing so, they inflict an internality on themselves which can be corrected by a sin tax.[1]

Our paper shows that there is a gap between the stated aim of sin taxes on unhealthy food consumption in practice and the paternalistic foundation in the economic literature. We show that the optimal paternalistic tax on calorie intake is effective at the *intensive margin*, i.e., it mitigates weight problems of individuals, but not at the *extensive margin*, i.e., it does not induce individuals to choose a healthy weight. The reason is that self-control problems—which should be corrected by the optimal paternalistic tax—influence only the intensive margin, whereas the extensive margin is purely determined by rational motives of the individuals. Implementing healthy weight by a tax—and thus overcoming the problem at the extensive margin—is also possible. But the associated health-maximizing tax requires a further, welfare-decreasing distortion that may render the tax inferior to not taxing the individuals at all. In addition, with heterogeneous individuals the optimal uniform paternalistic tax may have the negative side effect of reducing calorie intake of individuals with under- and normal weight. We derive these results within a theoretical model and confirm them by an empirical calibration with data on US adults.

In the theoretical model, individuals consume two goods, food and a non-food good. Food consumption is measured by calorie intake and associated with health costs. The model has two decisive features. First, health costs are U-shaped with a minimum at some healthy level of food consumption. This assumption is highly relevant since calorie intake is positively correlated with the BMI (Hall et al. 2011) and the all-cause mortality is U-shaped in the BMI (Global BMI Mortality Collaboration 2016, Figure 3). Second, besides the self-control problem, individuals have a rational motive for choosing an unhealthy diet. Specifically, the net marginal benefit of food, defined as the marginal rate of substitution between food and the non-food good minus the relative food price, may be different from zero. As the individual sets this net marginal benefit equal to the marginal health costs, the chosen calorie intake may deviate from its healthy level. The literature provides evidence for both motives of an unhealthy diet. For instance, Cutler et al. (2003) and Courtemanche

---

[1] Externalities caused in the health system are another justification of sin taxes. Allcott et al. (2019a, 2019b) give an excellent survey on different justifications and design issues of sin taxation. See Griffith et al. (2017) for a recent review on the empirical literature.

et al. (2014) show that obesity can be explained by decreasing food prices and costs, i.e. the rational motive, and that this impact is aggravated by hyperbolic discounting, i.e. the self-control motive.[2]

Our analysis starts with the case of a representative individual in the absence of taxation. We define the tempting calorie intake as the level of food consumption at which the above-mentioned net marginal benefit of food becomes zero. In its diet choice, the individual trades off the incentive to attain this tempting calorie intake against the need to account for health costs. Hence, it realizes a calorie intake between the tempting level and the healthy level. That is to say, food consumption is above (below) the healthy level if tempting calorie intake is larger (lower) than healthy calorie intake, so the individual becomes overweight (underweight). If tempting and healthy calorie intake just coincide, the individual realizes healthy weight. Importantly, whether the individual becomes over- or underweight (the extensive margin) is solely determined by the relation of tempting and healthy calorie intake and not by the degree of self-control. The self-control problem only worsens over- or underweight (the intensive margin), but does not determine the direction of the individual's weight problem.

Next, we determine the paternalistic welfare-maximizing tax on calorie intake. The paternalistic social planner maximizes the individual's true utility taking into account the full health costs. The optimal tax turns out to be positive for an overweight individual and negative (subsidy) for an underweight individual. In both cases, the weight problem is mitigated at the intensive margin, i.e., the overweight (underweight) individual consumes less (more) calories and becomes less overweight (underweight). But the problem at the extensive margin remains. That means, the overweight remains overweight and the underweight remains underweight. Hence, the optimal paternalistic tax on unhealthy food does not give the individual the incentive to implement healthy weight. The reason is that the paternalistic tax corrects only for the self-control problem, but not for the rational motive for an unhealthy food consumption.

In order to implement a diet that minimizes health costs and, thereby, maximizes health, we show that the tax on calorie intake just has to reflect the rational reason for an unhealthy diet, i.e. the net marginal utility of food consumption. However, such a health-maximizing tax leads to a stronger distortion (larger substitution effect) of the individual's consumption decision than the optimal paternalistic tax. Hence, the individual's true utility is lower under health maximization than under paternalistic welfare maximization. In addition, health-maximizing taxation may even be inferior to not taxing the individual at all. This result occurs if the individual's self-control problem is not too pronounced. In our empirical calibration, we confirm these theoretical results and show that for the average US adult the welfare increase caused by the paternalistic tax amounts to a money equivalent of \$2.16 per month, whereas the health-maximizing tax indeed lowers welfare by \$3.35 per month.

---

[2] Evidence in general and for other specific sin goods like smoking can also be found in Gruber and Kőszegi (2001), Ikeda et al. (2010), Andersen et al. (2014) and Loewenstein and Chater (2017).

We finally extend the analysis to heterogeneous individuals and consider three types of consumers that may differ in the preferences for food consumption and the extent of the self-control problem. We show that the basic insights from the analysis of the representative individual carry over to each type of individual, if we focus on type-specific paternalistic taxation. Moreover, we investigate uniform calorie taxation and show that the optimal paternalistic tax rate reflects the weighted average of the individuals' self-control problems. If the three consumer types represent the groups of over-, under- and healthy weight individuals and if overweight is more prevalent than underweight, then our theoretical analysis implies that the optimal paternalistic tax on calorie intake is positive and mitigates the weight problem of the overweights. But this comes at the costs of individuals with underweight and healthy weight, as they are also incentivized to reduce calorie intake and, thus, (further) deviate from a healthy diet. These insights are confirmed by the empirical calibration, where we show that uniform paternalistic calorie taxation increases welfare of the overweight by $3.82 per month, whereas the under- and normal weights lose $4.92 and $2.17 per month, respectively.

Our paper links two important strands of the theoretical literature. First, there is a literature on paternalistic taxation of sin goods, for instance, O'Donoghue and Rabin (2003, 2006), Haavio and Kotakorpi (2011, 2016) and Cremer et al. (2012, 2016). In contrast to our analysis, all these studies assume monotonically increasing instead of U-shaped health costs. Hence, their approaches do not contain the distinction between extensive and intensive margin and cannot reveal the impact of taxation on these margins. Second, there is a literature on rational obesity, e.g., Levy (2002), Dragone (2009) and Dragone and Savorelli (2012). In contrast to our analysis, this literature ignores taxation and, thus, also cannot address the question whether healthy weight is obtained by a tax on unhealthy food consumption.[3] We combine both strands of literature by taking the issue of taxation in the presence of self-control problems from the first strand and the rational motive for weight problems from the second strand. By doing so, we are the first to investigate the implications of food taxation in the presence of extensive and intensive margins. The analysis builds on our earlier analysis in Kalamov and Runkel (2020), where we use a more general, fully dynamic model with different kinds of present-focused preferences in order to identify the distinction between extensive and intensive margins. But there we also ignore taxation issues and, thus, cannot derive the results obtained in the present paper.

In the next section, we introduce the basic framework. Section 3 considers food taxation in the case of a representative individual, before we turn to the case with heterogeneous individuals in Sect. 4. Section 5 concludes.

---

[3] The studies on sin taxation build on the hyperbolic discounting framework of Laibson (1997), while the rational obesity literature is related to rational addiction (Becker and Murphy 1988).

2 Basic framework

Model assumptions Consider the case with a large number of individuals. We first assume that all individuals are identical and focus on the representative individual. This individual consumes food in quantity x and a non-food good in quantity z. Food consumption x may be measured in calorie intake. We therefore refer to x interchangeably as food consumption and calorie intake. The individual's consumption utility is given by $V(z, x)$. The utility function $V(z, x)$ is twice continuously differentiable. It satisfies $V_{kk} < 0 < V_k$ for $k = z, x$ and is assumed to be concave in (z, x).

As already mentioned in Introduction, there is a U-shaped relation between the BMI and health costs of an individual (Global BMI Mortality Collaboration 2016). Moreover, the BMI is positively correlated with calorie intake (Hall et al. 2011), which in our approach is modeled by food consumption x. Hence, there will also be a U-shaped relation between food consumption (calorie intake) and health costs. In our approach, this relation is represented by the twice continuously differentiable function $C(x)$. We assume $C(x^H) = 0$ and $C_x(x) \lesseqgtr 0$ if and only if $x \lesseqgtr x^H$, where x^H refers to the healthy level of calorie intake. The second derivative of the health cost function is $C_{xx}(x) > 0$. Hence, health costs are U-shaped with a minimum at healthy food consumption x^H, where health costs become zero.[4] For food consumption above or below the healthy level x^H, the individual does not realize the health-maximizing BMI and, therefore, faces positive health costs. In order to ease later discussion and in line with the usual wording in the definition of obesity, the individual is said to be overweight (underweight) if it chooses $x > x^H$ ($x < x^H$) and has healthy weight if $x = x^H$, being aware that the BMI is determined by both body weight and height.

The true net utility of the individual equals utility from consumption less health costs of calorie intake. Formally, it can be written as

$$u = V(z, x) - C(x). \tag{1}$$

In deciding on consumption, the individual suffers from a self-control problem and takes into account only a fraction β of its health costs. The decision utility reads

$$\hat{u} = V(z, x) - \beta C(x). \tag{2}$$

Throughout our formal analysis, we assume $\beta \in [0, 1[$. Decision utility (2) then differs from true utility (1) since the individual underestimates health costs. For comparison purposes only, we sometimes refer to the case $\beta = 1$ where true utility and decision utility coincide, and the individual does not face a self-control problem.

The market price of food consumption x is denoted by p, while non-food consumption z is the numéraire with a price normalized to one. Food may be taxed at rate τ per calorie consumed. For simplicity, we refer to τ as a fat tax. The individual

---

[4] In practice, often a range of BMI levels (between 18.5 and 25) is considered to be healthy. This is consistent with the above-mentioned U-shape of all-cause mortality estimated by The Global BMI Mortality Collaboration (2016) and our corresponding modeling with only one value x^H for healthy food consumption, if the health costs for BMI levels between 18.5 and 25 remain relatively low.

has given income e and receives a lump-sum transfer ℓ from the government. Since the number of individuals is large, each individual takes this transfer as given. The budget constraint of the representative individual can be written as

$$z + (p + \tau)x = e + \ell, \tag{3}$$

and equates consumption expenditures to income plus the lump-sum transfer.[5]

Consumption decision of the individual The individual chooses consumption x and z in order to maximize its decision utility $\hat{u}$ defined in (2) subject to the budget constraint (3), taking as given income e as well as the lump-sum transfer ℓ and the fat tax rate τ. Solving (3) with respect to z, inserting into (2) and setting the derivative with respect to x equal to zero gives the first-order condition

$$V_x(z^*, x^*) - (p + \tau)V_z(z^*, x^*) = \beta C_x(x^*), \tag{4}$$

where the star indicates the solution to the individual's maximization problem. Equation (4) states that the individual chooses food consumption such that the net marginal utility (LHS) just equals the perceived marginal health costs (RHS). Together with Eq. (3), it determines the individual's calorie demand x^* as a function of the self-control parameter β and the policy variables τ and ℓ.

An important property of the individual's decision can be derived when we consider the impact of the self-control parameter on food consumption. Totally differentiating (4) and taking into account that z^* has to satisfy the budget constraint (3) yields

$$\frac{dx^*}{d\beta} = \frac{C_x}{V_{xx} + (p + \tau)^2 V_{zz} - 2(p + \tau)V_{zx} - \beta C_{xx}} \gtreqless 0 \quad \Leftrightarrow \quad x^* \gtreqless x^H. \tag{5}$$

The denominator in (5) has to be negative due to the second-order condition of utility maximization. Equation (5) states that a more severe self-control problem (a lower value of the self-control parameter β) increases food consumption and calorie intake when the individual is already overweight ($x^* > x^H$) and reduces food consumption and calorie intake when the individual is already underweight ($x^* < x^H$). There is no effect of the self-control parameter on food consumption if the individual has healthy weight ($x^* = x^H$). These insights from (5) have the important implication that the self-control parameter β affects the intensive margin but not the extensive margin. Put differently, the self-control problem influences the extent of the weight problem but does not determine whether the individual becomes over- or underweight.

Solution in the absence of any policy As benchmark, we first characterize the individual's consumption decision in the absence of any policy. Setting $\tau \equiv \ell \equiv 0$,

---

[5] Note that our model is a one-shot framework where the individual makes a consumption decision only once. Intuitively, our basic insights should also be true in the fully dynamic model of Kalamov and Runkel (2020), where an individual with a present bias consumes in every period and the relation of the individual's weight and food consumption (calorie intake) is modeled by an equation of motion.

taking into account (3) and denoting the individual's food consumption in the absence of taxation by $x^* = x^o$, the first-order condition in (4) can be rewritten as

$$N(x^o) = V_x(e - px^o, x^o) - pV_z(e - px^o, x^o) = \beta C_x(x^o), \qquad (6)$$

where $N(x) := V_x(e - px, x) - pV_z(e - px, x)$ denotes the net marginal utility of food consumption in the absence of any policy. If we define x^T by $N(x^T) = 0$, then x^T gives the individual's food consumption in the (hypothetical) situation without any policy intervention and without health costs. In the following, x^T is called the tempting food consumption or, equivalently, the tempting calorie intake. We can then prove the following result (all proofs are relegated to the appendix).

Proposition 1 *Assume $\tau \equiv \ell \equiv 0$. The individual's food demand x^o is then characterized by the following properties*:

(a)   If $x^H < x^T$, then $x^H < x^o < x^T$ (Overweight).
(b)   If $x^T < x^H$, then $x^T < x^o < x^H$ (Underweight).
(c)   If $x^T = x^H$, then $x^T = x^o = x^H$ (Healthy weight).

Proposition 1 shows that in the absence of any policy intervention the individual's food consumption x^o lies between the tempting level x^T and the healthy level x^H. If healthy calorie intake x^H falls short of tempting calorie intake x^T, the individual faces a trade-off between eating more in order to reach tempting calorie intake x^T and eating less in order to ensure healthy calorie intake x^H. The solution to this trade-off is a food level x^o between the two extremes. The individual is then overweight ($x^o > x^H$), as shown in part (a) of Proposition 1. If tempting calorie intake x^T is lower than healthy calorie intake x^H, the individual balances the incentive to eat less for tempting calorie intake x^T and the incentive to eat more for healthy calorie intake x^H. The individual ends up with too low food consumption and underweight ($x^o < x^H$), according to Proposition 1 (b). Part (c) of Proposition 1 contains the knife edge case in which tempting calorie intake x^T and healthy calorie intake x^H just coincide. The individual then does not face a trade-off and attains healthy weight ($x^o = x^H$).

The insights from Proposition 1 confirm the conclusion which we already draw from Eq. (5): The self-control problem does not influence the extensive margin of the individual's food consumption decision. The proposition holds for all values of β, even for the case $\beta = 1$, where the individual does not face a self-control problem. Instead, whether the individual becomes over- or underweight solely depends on the relation between tempting and healthy food consumption, x^T and x^H, respectively. The individual will deviate from the healthy weight as long as both consumption levels are not the same and, thus, the net marginal utility of calorie intake is not zero. Since this holds independently of the parameter β, the individual's decision at the extensive margin is determined solely by a rational motive and not by self-control problems.

3 Optimal policy with a representative consumer

Next, we turn to the optimal tax policy and initially keep the assumption of a representative individual. We first consider a paternalistic social planner who maximizes the individual's true utility. Thereafter, we consider the health-maximizing policy.

Paternalistic welfare maximization While the representative individual takes τ and ℓ as given, in setting the optimal policy the social planner accounts for the public budget constraint $\ell = \tau x^*$. Inserting into the private budget constraint (3), the first-order condition (4) of the individual's maximization problem can be rewritten as

$$V_x(e - px^*, x^*) - (p + \tau)V_z(e - px^*, x^*) = \beta C_x(x^*). \tag{7}$$

This condition determines the individual's food demand x^* as a function of the fat tax rate τ. We denote this function by $x^* = X(\tau)$.

In determining the optimal policy, the social planner takes into account the individual's reaction to tax rate changes determined by the function $X(\tau)$. The paternalistic approach implies that the social planner maximizes true utility u from (1) instead of decision utility $\hat{u}$ from (2). The maximization problem reads

$$\max_{\tau} \ u = V\big[e - pX(\tau), X(\tau)\big] - C[X(\tau)].$$

The first-order condition is

$$\frac{du}{d\tau} = \Big[V_x(\cdot) - pV_z(\cdot) - C_x(\cdot)\Big]X_\tau = 0. \tag{8}$$

with $X_\tau := X'(\tau)$. Using the first-order condition (7) of the individual's maximization problem in order to replace V_x in (8) by $(p + \tau)V_z + \beta C_x$ and solving for τ gives

$$\tau^P = (1 - \beta)\frac{C_x[X(\tau^P)]}{V_z[e - pX(\tau^P), X(\tau^P)]}, \tag{9}$$

which is an implicit equation determining the optimal paternalistic fat tax rate τ^P. Inserting the optimal tax rate (9) into the individual's first-order condition (7) implies

$$N[X(\tau^P)] = V_x[e - pX(\tau^P), X(\tau^P)] - pV_z[e - pX(\tau^P), X(\tau^P)] = C_x[X(\tau^P)]. \tag{10}$$

This equation determines the individual's food consumption $X(\tau^P)$ when the paternalistic fat tax is optimally chosen by the social planner.

With the help of Eqs. (9)–(10), we can prove the following result.

Proposition 2 *The optimal paternalistic fat tax τ^P and the corresponding food consumption $X(\tau^P)$ are given by (9) and (10), respectively. The following statements hold*:

(a)   If $x^H < x^T$, then $\tau^P > 0$ and $x^H < X(\tau^P) < x^o < x^T$ (Overweight).

(b)   If $x^T < x^H$, then $\tau^P < 0$ and $x^T < x^o < X(\tau^P) < x^H$ (Underweight).

(c)   If $x^T = x^H$, then $\tau^P = 0$ and $x^T = x^o = X(\tau^P) = x^H$ (Healthy weight).

Proposition 2 (a) ((b)) shows that the optimal paternalistic fat tax is positive (negative), if the individual is overweight (underweight). Hence, an overweight individual is taxed for its food consumption while an underweight individual receives a food subsidy. The fat tax (fat subsidy) reduces (increases) the individual's food consumption such that the individual becomes less overweight (underweight), compared to the situation without any policy intervention characterized in Proposition 1. An individual with a healthy weight is not taxed at all and, thus, realizes the same food consumption and the same calorie intake as without taxation, as shown in part (c) of Proposition 2.

Even though the optimal paternalistic policy mitigates the weight problem of the individual, the most important insight from Proposition 2 is that the optimal paternalistic policy does not completely eliminate the individual's weight problem. The overweight individual remains overweight ($x^o > X(\tau^P) > x^H$), and the underweight individual remains underweight ($x^o < X(\tau^P) < x^H$). Put differently, the optimal paternalistic fat tax mitigates the weight problem at the intensive margin, but does not solve the weight problem at the extensive margin. The intuition is that the optimal paternalistic policy solely aims at correcting the self-control problem. As explained above, however, the individual deviates from healthy calorie intake not because of self-control problems but because of rational motives; and the rational motives are not corrected for by the paternalistic approach. This story is also confirmed by the fact that in the absence of self-control problems ($\beta = 1$), the optimal tax in (9) is always zero, independent of whether the individual is over- or underweight.

Health maximization As argued in the Introduction, the motivation behind many taxes on unhealthy food in practice is to induce healthy food consumption. In our framework, healthy behavior means that the individual chooses $x^* = x^H$. To find out which tax rate induces this healthy consumption level, we insert $x^* = x^H$ into (4), take into account the public and private budget constraints $\ell = \tau^H x^H$ and $z^* = e - px^H$, respectively, as well as $C_x(x^H) = 0$ and finally solve for τ. The result is

$$\tau^H = \frac{V_x(e - px^H, x^H) - pV_z(e - px^H, x^H)}{V_z(e - px^H, x^H)}. \tag{11}$$

With this expression, we obtain the following result.

Proposition 3 *The health-maximizing fat tax rate τ^H is given by* (11). *The individual then always attains healthy weight $x^* = x^H$. Moreover, the following statements hold*:

(a) If $x^H < x^T$, then $\tau^H > 0$.
(b) If $x^T < x^H$, then $\tau^H < 0$.
(c) If $x^T = x^H$, then $\tau^H = 0$.

The tax rate in Eq. (11) induces the individual to choose a healthy diet and to realize the healthy weight x^H. Hence, it overcomes the individual's health problem at the extensive margin and, thereby, also avoids the problem at the intensive margin without directly targeting the individual's self-control problem. This view is supported by the observation that the health-maximizing tax rate (11) does not contain the self-control parameter β and that its numerator reflects the net marginal utility $N(\cdot) = V_x(\cdot) - pV_z(\cdot)$. As already argued in Proposition 1, this net marginal utility determines the individual's decision at the extensive margin, and a nonzero sign of $N(\cdot)$ gives a rational incentive for over- or underweight. Therefore, the health-maximizing tax rate 'corrects' for the rational reason for an unhealthy diet choice and, by doing so, eliminates the whole weight problem of the individual. As intuitively plausible, parts (a) and (b) of Proposition 3 show that the health-maximizing tax rate is positive if the individual is overweight and negative if the individual is underweight. Of course, an individual with healthy weight needs not to be taxed, as shown by Proposition 3 (c).

How does the health-maximizing tax rate (11) perform in terms of the individual's true welfare? In order to answer this question, we define

$$u^o := V(e - px^o, x^o) - C(x^o), \tag{12}$$

$$u^P := V[e - pX(\tau^P), X(\tau^P)] - C[X(\tau^P)], \tag{13}$$

$$u^H := V(e - px^H, x^H). \tag{14}$$

Equation (12) gives the individual's true utility in the absence of taxation, while (13) and (14) reflect the individual's true utility under paternalistic welfare maximization and health maximization, respectively. If tempting calorie intake equals its healthy level ($x^T = x^H$), we have $x^o = X(\tau^P) = x^H$ and, thus, $u^o = u^P = u^H$. True utility is then always the same, independently of the tax policy. The rationale is obvious since in this knife edge case the individual always realizes healthy weight. More interesting is the welfare comparison if tempting and healthy calorie intake differ. We then obtain

Proposition 4 *If $x^T \neq x^H$, then the individual's true utility satisfies*

(a) $u^H < u^P$.
(b) $u^H \gtreqqless u^o$ iff $\beta \gtreqqless \bar{\beta}$ with $\bar{\beta} \in [0, 1[$.

As shown in Proposition 4 (a), the individual's true utility is always lower under health maximization than under paternalistic welfare maximization. This is obvious

since the health-maximizing tax (11) targets the extensive margin and the complete elimination of the individual's weight problem, whereas the paternalistic tax (9) only corrects the self-control problem and, thus, the individual's weight problem at the intensive margin. Health maximization therefore needs a larger distortion (substitution effect) and, thus, leaves the individual with a lower well-being than paternalistic welfare maximization.

Part (b) of Proposition 4 worsens the picture of the health-maximizing policy. Trying to completely overcome the weight problem and inducing healthy weight with the help of the health-maximizing tax rate (11) may be inferior to not taxing the individual at all. This result is obtained if the self-control problem is not too severe ($\beta > \bar{\beta}$). The rationale is as follows. The health-maximizing tax solves the weight problem at the extensive margin and, thereby, also at the intensive margin. But only the correction of the self-control problem at the intensive margin increases the individual's true utility, whereas eliminating the rational motive at the extensive margin reduces the individual's true utility. Hence, if the self-control problem is relatively moderate, then the utility gain at the intensive margin is overcompensated by the utility loss at the extensive margin and, overall, the individual's true utility is reduced by health-maximizing taxation. Admittedly, it is in the end an empirical question whether β is larger than $\bar{\beta}$ and health-maximization is really worse than not taxing the individual at all. We therefore now calibrate the model to US data.

Empirical Calibration The focus is on the average US adult. We use data from the representative biannual National Health and Nutrition Examination Survey (NHANES) in its most recent version from 2017-18 (CDC 2021). This survey yields information on daily calorie (kcal) intake and the BMI. While the BMI is measured by trained health technicians, the data on calorie intake are self-reported by the participants of the survey.

Column (1) of Table 1 reports descriptive statistics on all US adults (age ≥ 18), while columns (2), (3) and (4) divide the total sample into underweight adults (BMI < 18.5), adults with normal weight (BMI $\in$ [18.5, 25)) and adults with overweight (BMI ≥ 25), respectively. The mean daily calorie intake of all adults is 2152.42 resulting in a mean BMI of 29.74. The reported consumption levels of all subgroups in columns (2)-(4) are almost equal to the average intake, even though the overweights' BMI (32.61) is ten points larger than the normal weights' BMI (22.28), which is almost five points larger than the underweights' BMI (17.56). The calorie intake of the normal weight adults is

Table 1 Daily calorie intake and BMI of US adults (standard errors in parentheses)

	Age ≥ 18	Age ≥ 18 BMI < 18.5	Age ≥ 18 $18.5 \leq$ BMI < 25	Age ≥ 18 BMI ≥ 25
	(1)	(2)	(3)	(4)
Daily kcal intake	2152.42 (18.31)	2186.47 (124)	2152.47 (56.91)	2151.6 (20.66)
BMI	29.74 (0.29)	17.56 (0.11)	22.28 (0.06)	32.61 (0.25)
Observations	4982	87	1229	3666

Source: NHANES 2017-18 (CDC 2021)

Table 2 Estimated parameter values

Parameter	Value[a]
x^H in kcal/day:	2152.47
x^o in kcal/day:	2620.36
β:	0.6
c in cents/kcal:	0.00043
p in cents/kcal:	0.442
e in \$/day:	132.12
γ:	0.073

[a]Source: x^H and BMI from NHANES 2017-18 (CDC 2021), x^o, p, c from authors' calculations based on Hall et al. (2011), Kalamov (2020), Allcott et al. (2019a)

pretty much in line with the amount usually recommended for a healthy diet (see, e.g., Appendix 2 in US Department of Health and Human Services and US Department of Agriculture, 2015). The underweights' calorie intake is unreliable, as there are only 87 observations and its standard error is very large. This is not the case for overweight adults, indeed, but their reported calorie intake seems much too low and not consistent with their high BMI. A possible explanation is that overweight individuals underreport their consumption due to the fear of stigma.

Hence, we interpret the observed calorie intake in column (3) as the calorie intake consistent with a healthy BMI and recalculate the average calorie intake of all US adults in column (1) such that it is consistent with the associated BMI of 29.74. In terms of our model, this implies $x^H = 2152.47$ (all estimated parameter values are listed in Table 2).

To determine the true average calorie intake of all US adults, we determine the necessary increase in x, starting from x^H, that would lead to the corresponding increase in BMI from 22.28 to 29.74. Hall et al. (2011) estimate that for the average overweight individual, a 10 kcal/day change in energy intake leads to a 1 pound ≈ 0.45 kg weight change. Thus, an individual who increases consumption by one kcal per day can expect a rise in weight by 0.045 kg, i.e., 0.045 kg/kcal/day. Dividing the weight change by the square of the US adults' average height of 1.68m (CDC 2021), we get $0.045/1.68^2 \approx 0.01594$ BMI/kcal/day. Starting from the observed healthy BMI of 22.28, reaching a BMI of 29.74 requires an increase in daily calorie intake equal to

$$\Delta x = \frac{\Delta \text{ BMI}}{0.01594 \text{ BMI/kcal/day}} = \frac{29.74 - 22.28}{0.01594} \text{ kcal/day} \approx 467.89 \text{ kcal/day}.$$

(15)

Hence, the intake of the average US adult should be $x^H + \Delta x \approx 2620.36$ kcal/day, instead of 2152.42 kcal/day reported in Table 1. In our model, we interpret this

value as the representative individual's choice in the absence of any policy and set $x^o = 2620.36$.[6]

Next, we assume a quasi-linear and iso-elastic utility function $V(z, x) = z + x^{1-\gamma}/(1 - \gamma)$ with $\gamma > 0$ and a quadratic health cost function $C(x) = c(x - x^H)^2/2$ with $c > 0$. Because of quasi-linearity, the cost function measures the health costs in monetary units.[7] In order to estimate the cost parameter c, it is useful to consider the money-metric of the marginal internality defined as the ignored proportion of the marginal health costs. Evaluated in the case without a tax, we obtain

$$\text{Marginal internality} = (1 - \beta)c(x^o - x^H). \tag{16}$$

Allcott et al. (2019a) estimate the marginal internality of sugar-sweetened beverages (SSB) consumption as 0.93 cents/ounce. One SSB serving (a can) contains 12 ounces and 140 kcal (Allcott et al. 2019b). Hence, the marginal internality in (16) can be set equal to $0.93(12)/140 \approx 0.08$ cents/kcal. Moreover, Allcott et al. (2019a) show that a counterfactual normative SSB consumer, who has perfect nutritional knowledge and full self-control, would consume 31% less than the average SSB consumer. Thus, the representative individual overconsumes by $1/(1 - 0.31) - 1 \approx 45\%$. Therefore, the observed marginal internality is likely a result of consumers not taking into account 30-50% of the health costs. In the benchmark case, we assume that the observed marginal internality is due to $\beta = 0.6$. Inserting $\beta = 0.6$ as well as x^o and x^H from Table 2 in (16) and setting the LHS equal to 0.08, we get $c = 0.00043$. Since the estimates of the marginal internality in Allcott et al. (2019a) range between 0.91 and 2.14 cents/ounce and since it holds only for SSBs, it may be that we under- or over-estimate the marginal health costs of general calorie intake. In the later sensitivity analysis, we therefore consider both a half as large marginal health cost $c = 0.00021$ and a four times as large marginal health cost $c = 0.0017$.

It remains to estimate e, p and γ. For income e, we use the per capita disposable income in 2018 reported by the Bureau of Economic Analysis (Table 2.1 in BEA 2021). In 2018 dollars, it is equal to \$48,223 per year. On a daily basis, we obtain \$132.12. For the price p, we rely on Kalamov (2020), who estimates the price per kcal for food-at-home (FAH) and food-away-from-home (FAFH) in 2009 dollars (see Table 2 in Kalamov 2020). Using the estimated daily calorie intake from FAH and FAFH from Table 1 in that paper, we get an average price of 0.3775 cents/kcal. In 2018 dollars, this price is equal to $p = 0.442$ cents/kcal. Finally, we estimate the preference parameter γ from the first-order condition (6), where we used the utility and health cost functions specified above and where x^o, x^H, c and β are taken from Table 2. We get $\gamma \approx 0.073$. Thus, utility is only slightly concave in daily calorie intake.

Simulation Results With the estimated parameters in Table 2, we now simulate our theoretical model. Results are reported in Table 3.

---

[6] Seven US cities tax sugar-sweetened beverages (Allcott et al. 2019b), but there is no US-wide tax on unhealthy foods. Therefore, these taxes can safely be neglected for the purpose of the calibration.

[7] The money-metric of the health costs is defined as $C(x)/V_z(\cdot)$. Since $V_z = 1$ in the quasi-linear case, we obtain $C(x)$ as a money-metric function.

Table 3 Results

Parameter	Value[a]
τ^P in cents/kcal:	$0.049 \approx 11\%$ of p
τ^H in cents/kcal:	$0.127 \approx 29\%$ of p
$X(\tau^P)$ in kcal/day:	2440.11
$\frac{X(\tau^P)-x^H}{x^o-x^H}$:	0.615
$\frac{x^o-X(\tau^P)}{x^o-x^H}$:	0.385
$EV(\tau^P)$ in \$/month:	2.16
$EV(\tau^H)$ in \$/month:	-3.35
$\bar{\beta}$:	0.48

[a]Source: authors' calculations

The optimal paternalistic tax is $\tau^P = 0.049$ cents/kcal and represents an 11% increase in the price per kcal. On the other hand, healthy consumption can be implemented by a tax $\tau^H = 0.127$ cents/kcal, which corresponds to a much larger relative price increase of 29%. Consistently with this observation, the rational motive turns out to be more important for explaining overweight than the self-control problem. To see this, notice that the optimal paternalistic consumption is estimated as $X(\tau^P) = 2440.11$ kcal/day. The rational component of obesity, $X(\tau^P) - x^H$, is therefore equal to 61.5% of the observed deviation from healthy consumption, $x^o - x^H$, while only 38.5% are due to bias.

How could the tax rates reported in Table 3 be implemented in practice? Since it is hard to tax calories directly, a nutrient tax like the sugar tax is a more promising candidate to achieve the required calorie reduction. Harding and Lovenheim (2017) estimate that a 20% tax rate on the value of the sugar content leads to an 18.54% reduction in total calorie intake and that this effect is close to being linear in the tax rate. To implement $X(\tau^P)$, the government needs to lower consumption by $[x^o - X(\tau^P)]/x^o \approx 6.88\%$. To implement x^H, the required reduction in intake is $(x^o - x^H)/x^o \approx 17.86\%$. Therefore, the paternalistic sugar tax would be $(0.0688/0.1854)20\% \approx 7.42\%$, while the health-maximizing sugar tax amounts to $(0.1786/0.1854)20\% \approx 19.26\%$. This, too reveals the large difference between the paternalistic and the health-maximizing policy.

Finally, we estimate the welfare consequences of taxation. Consider the equivalent variation $EV(\tau^t)$ of the tax τ^t with $t = H, P$, i.e., the income increase in the absence of taxation that leads to the equivalent utility change as the tax. Formally, we have

$$V[e + EV(\tau^t) - x^o, x^o] - C(x^o) = V[e - X(\tau^t), X(\tau^t)] - C[X(\tau^t)], \qquad (17)$$

where $X(\tau^H) \equiv x^H$. Our calculations show that the paternalistic policy raises welfare by $EV(\tau^P) = \$2.16$ per month, while the health-maximizing tax lowers it by $EV(\tau^H) = -\$3.35$ per month. Consistently, we find $\bar{\beta} = 0.48$ which is lower than $\beta = 0.6$ used in our calibration. As indicated by Proposition 4, our simulation therefore provides evidence that, in terms of welfare, health-maximizing taxation is not

only inferior to paternalistic taxation, but may be even worse than not taxing food at all.

Sensitivity Analysis We run a variety of robustness checks. Details can be found in an online appendix. We first vary β for $\beta \in [0, 1]$. Consistently with our findings in the baseline case, the health-maximizing tax improves welfare, compared to the non-tax case, if $\beta < 0.48$. Moreover, the gap between the welfare effects of τ^P and τ^H shrinks if β becomes lower, since the self-control problem becomes more important, relative to the rational motive of overweight. However, for $\beta > 0.48$ an increase in β increases the welfare costs of τ^H by more than it reduces the welfare gain of τ^P. Hence, the gap between the welfare effects of paternalistic and health-maximizing policy rises.

We repeat the baseline calculations as well as the variation of β with alternative values for the estimated parameters p and c. There are only small changes, if the price is cut in half to $p = 0.221$ or doubled to $p = 0.884$. In particular, the welfare effects measured by the EV and the critical value $\bar{\beta}$ remain almost unchanged. Variation of the marginal health costs to $c = 0.00021$ or $c = 0.0017$ have non-negligible welfare effects. However, welfare gains and costs are changed by almost the same factor and therefore, the relative welfare evaluation of τ^P and τ^H remains unchanged. Moreover, the critical value $\bar{\beta}$ is influenced by c only very slightly, so the health-maximizing policy is still likely to reduce welfare. Overall, we can conclude from the sensitivity analysis that $\bar{\beta}$ is very robustly estimated to be close to 0.5. Thus, for realistic values of β (above 0.5), the welfare-reducing effect of health-maximizing taxation is quite robust.

4 Optimal policy with heterogeneous consumers

Next, we turn to the case with heterogeneous individuals. In order to highlight the most interesting implications, we focus on the case with only three types of individuals.

Consumption decision of individuals The type of individuals is denoted by $i, j = 1, 2, 3$. The number of type i individuals is n_i with the normalization that the total number of individuals is $\sum_{i=1}^{3} n_i = 1$. As stated by O'Donoghue and Rabin (2006), in the context of fat taxes differences in the self-control parameter and in the preferences for food consumption seem to be most interesting. We therefore assume that type i individuals have a self-control parameter $\beta_i \in [0, 1[$, where $\beta_i < \beta_j$ implies that type i individuals have less self-control than type j individuals. Type i's consumption preferences are captured by the quasi-linear utility function $V^i(z_i, x_i) = z_i + W(x_i, \theta_i)$ with the preference parameter $\theta_i > 0$. The subutility function W reflects the preference for food and satisfies $W_{xx} < 0 < W_x$ and $W_{x\theta} > 0$. Hence, $\theta_i < \theta_j$ implies that type j individuals have a stronger preference for calorie intake than type i individuals, since the marginal utility from food consumption is increasing in the preference parameter.

The focus in the analysis of heterogeneous individuals is on paternalistic welfare maximization. In general, we account for type-specific tax rates τ_i and lump-sum transfers ℓ_i, but the main results will be derived under the more realistic assumption

that all individuals face the same tax rate $\tau_i = \tau$ and obtain the same lump-sum transfer $\ell_i = \ell$. The decision of a type i individual is described by the problem

$$\max_{x_i} \hat{u}_i = e + \ell_i - (p + \tau_i)x_i + W(x_i, \theta_i) - \beta_i C(x_i),$$

where we have used the budget constraint $(p + \tau_i)x_i + z_i = e + \ell_i$ in order to replace z_i. The first-order condition of type i's utility maximization reads

$$W_x(x_i^*, \theta_i) - p - \tau_i = \beta_i C_x(x_i^*). \tag{18}$$

It determines type i's calorie intake as a function of the tax rate, the preference parameter and the self-control parameter, i.e., $x_i^* = X(\tau_i, \theta_i, \beta_i)$. Differentiating gives $X_\tau(\cdot) = 1/(W_{xx} - \beta_i C_{xx}) < 0$, $X_\theta(\cdot) = -W_{x\theta}/(W_{xx} - \beta_i C_{xx}) > 0$ and $X_\beta(\cdot) = C_x/(W_{xx} - \beta_i C_{xx}) \gtreqqless 0$ iff $x_i^* \gtreqqless x^H$. Thus, calorie intake is decreasing in the fat tax rate τ_i and increasing in the preference parameter θ_i. Moreover, less self-control reflected by a reduction in β_i will increase (reduce, leave unchanged) calorie intake of overweight (underweight, normal weight) individuals.

Solution in the absence of any policy First, consider again the case without taxation, i.e., $\tau_i \equiv \ell_i \equiv 0$. Let $x_i^* = X(0, \theta_i, \beta_i) =: x_i^o$ be food demand of type i in the absence of taxation. The first-order condition (18) can then be rewritten as

$$M(x_i^o, \theta_i) = W_x(x_i^o, \theta_i) - p = \beta_i C_x(x_i^o), \tag{19}$$

with $M(x_i, \theta_i) := W_x(x_i, \theta_i) - p$. The function M represents the net marginal utility of calorie intake in the absence of taxation and plays the same role as the function N in the case of a representative individual. We obtain $M_x(\cdot) = W_{xx}(\cdot) < 0$ and $M_\theta(\cdot) = W_{x\theta} > 0$, so $M(\cdot)$ is decreasing in x_i and increasing in θ_i. Tempting consumption of type i individuals is denoted by x_i^T and implicitly defined by $M(x_i^T, \theta_i) = 0$. Due to $M_\theta > 0$ an increase in type i's food preference θ_i increases type i's tempting consumption x_i^T. Note that both $M(\cdot)$ and x_i^T are independent of β_i. It is then straightforward to prove

Proposition 5 *Proposition 1 applies to all three types of individuals, if we replace x^T by x_i^T and x^o by x_i^o for $i = 1, 2, 3$. Moreover,*

(a) $\beta_i = \beta_j$ and $\theta_i < \theta_j$ imply $x_i^o < x_j^o$,

(b) $\theta_i = \theta_j$ and $\beta_i < \beta_j$ imply $x_i^o \gtreqqless x_j^o$ if $x_i^T = x_j^T \gtreqqless x^H$.

Whether an individual of type i is overweight or underweight (extensive margin) depends solely on the relation between tempting consumption x_i^T and healthy consumption x^H, whereas the self-control parameter β_i determines the extent of the individual's weight problem (intensive margin). This is expressed in the first sentence of Proposition 5 and the same result as obtained in Proposition 1 for the representative individual. In Proposition 5 (a), we obtain the additional insight that, in the absence of taxation and for equal self-control parameters $\beta_i = \beta_j$, type i individuals consume less calories than type j individuals if they have a lower food preference $\theta_i < \theta_j$.

Moreover, according to Proposition 5 (b), for a zero tax rate and equal food preferences $\theta_i = \theta_j$, the deviation of realized weight from healthy weight is larger for type i than for type j individuals, if type i faces the more severe self-control problem with a lower parameter $\beta_i < \beta_j$; this is true for both underweights ($x_i^T = x_j^T < x^H$) and overweights ($x_i^T = x_j^T > x^H$).

Type-specific policy as a benchmark Consider now the case in which each type faces a specific tax rate τ_i and lump-sum transfer $\ell_i = \tau_i X(\tau_i, \theta_i, \beta_i)$.[8] The paternalistic social planner maximizes the sum of true utilities of all consumers, i.e., $w := \sum_{i=1}^{3} n_i u_i$ where $u_i = e + \ell_i - (p + \tau_i)x_i + W(x_i, \theta_i) - C(x_i)$. Using $x_i = x_i^* = X(\tau_i, \theta_i, \beta_i)$ and $\ell_i = \tau_i X(\tau_i, \theta_i, \beta_i)$, the maximization problem reads

$$\max_{\tau_1, \tau_2, \tau_3} w = \sum_{i=1}^{3} n_i u_i = \sum_{i=1}^{3} n_i \Big\{ e - pX(\tau_i, \theta_i, \beta_i) + W[X(\tau_i, \theta_i, \beta_i), \theta_i] - C[X(\tau_i, \theta_i, \beta_i)] \Big\}.$$

The first-order conditions are

$$\frac{\partial w}{\partial \tau_i} = n_i \Big\{ W_x[X(\tau_i, \theta_i, \beta_i), \theta_i] - p - C_x[X(\tau_i, \theta_i, \beta_i)] \Big\} X_\tau^i = 0, \quad i = 1, 2, 3 \quad (20)$$

with $X_\tau^i := X_\tau(\tau, \theta_i, \beta_i)$. Replacing $W_x(\cdot) - p$ from (18) by $\tau_i + \beta_i C_x(\cdot)$ and solving with respect to τ_i yields the optimal paternalistic fat tax rate for type i as

$$\tau_i^P = (1 - \beta_i)C_x[X(\tau_i^P, \theta_i, \beta_i)], \quad i = 1, 2, 3. \quad (21)$$

Inserting back into the first-order condition (18) gives

$$M[X(\tau_i^P, \theta_i, \beta_i), \theta_i] = W_x[X(\tau_i^P, \theta_i, \beta_i), \theta_i] - p = C_x[X(\tau_i^P, \theta_i, \beta_i)], \quad i = 1, 2, 3. \quad (22)$$

From (21) and (22), we obtain

Proposition 6 *Proposition* 2 *applies to all three types of individuals, if we replace x^T by x_i^T, x^o by x_i^o, τ^P by τ_i^P and $X(\tau^P)$ by $X(\tau_i^P, \theta_i, \beta_i)$ for $i = 1, 2, 3$. Moreover,*

(a) $\beta_i = \beta_j$ *and* $\theta_i < \theta_j$ *imply* $X(\tau_i^P, \theta_i, \beta_i) < X(\tau_j^P, \theta_j, \beta_j)$ *and* $\tau_i^P < \tau_j^P$,

(b) $\theta_i = \theta_j$ *implies* $X(\tau_i^P, \theta_i, \beta_i) = X(\tau_j^P, \theta_j, \beta_j)$ *independently of the relation between β_i and β_j. For $\beta_i < \beta_j$, we additionally obtain* $\tau_i^P \gtreqqless \tau_j^P \gtreqqless 0$ *if* $x_i^T = x_j^T \gtreqqless x^H$.

---

[8] Due to the quasi-linear utility, it makes no difference if we assume that each individual obtains the same lump-sum transfer, as long as all revenues are redistributed back to individuals.

According to the first sentence in Proposition 6, type-specific paternalistic fat taxes mitigate the weight problems of each type at the intensive margin, but do not implement healthy weight at the extensive margin. The additional insight from Proposition 6 (a) is that, for identical self-control problems ($\beta_i = \beta_j$), a lower food preference $\theta_i < \theta_j$ translates into a lower paternalistic tax and a lower calorie intake for type i compared to type j. Proposition 6 (b) shows that, for identical food preferences $\theta_i = \theta_j$, the type-specific paternalistic taxes induce identical calorie intakes for types i and j, independently of the degree of self-control β_i and β_j. This is because paternalistic taxation just aims at correcting the self-control problem. Of course, differences in the self-control parameters have an influence on the size of paternalistic taxation. If type i individuals face the more severe self-control problem ($\beta_i < \beta_j$) and both types are overweight (underweight) due to $x_i^T = x_j^T > x^H$ ($x_i^T = x_j^T < x^H$), type i individuals are taxed (subsidized) more, since they need a larger incentive to correct their bias.

Uniform fat taxes A type-specific policy is difficult to implement due to distributional and informational reasons. We therefore now consider the more realistic case that each individual faces the same tax rate τ and obtains the same lump-sum transfer ℓ. As each individual takes the policy instruments τ and ℓ as given, food consumption of a type i individual is still characterized by the first-order condition (18) and now given by $x_i^* = X(\tau, \theta_i, \beta_i)$. The common lump-sum transfer equally distributes total tax revenues over all individuals, i.e. $\ell = \tau \sum_{i=1}^3 n_i X(\tau, \theta_i, \beta_i)$. Taking into account this expression, the social planner's welfare maximization now reads

$$\max_{\tau} \ w = \sum_{i=1}^3 n_i u_i = \sum_{i=1}^3 n_i \Big\{ e - pX(\tau, \theta_i, \beta_i) + W[X(\tau, \theta_i, \beta_i), \theta_i] - C[X(\tau, \theta_i, \beta_i)] \Big\}.$$

Due to the quasi-linear utility function, this objective is the same as in the case with the type-specific policy, except that now all individuals face the same tax rate τ. The first-order condition of paternalistic welfare maximization can be written as

$$\frac{\partial w}{\partial \tau} = \sum_{i=1}^3 n_i \Big\{ W_x[X(\tau, \theta_i, \beta_i), \theta_i] - p - C_x[X(\tau, \theta_i, \beta_i)] \Big\} X_\tau^i = 0. \tag{23}$$

Using the consumers' first-order condition (18) in order to replace $W_x(\cdot) - p$ by $\tau + \beta_i C_x(\cdot)$ and solving for τ gives the optimal paternalistic fat tax rate

$$\tau^P = \frac{\sum_{i=1}^3 (1 - \beta_i) n_i X_\tau^i C_x[X(\tau^P, \theta_i, \beta_i)]}{\sum_{i=1}^3 n_i X_\tau^i} \tag{24}$$

Inserting this tax rate back into the consumers' first-order condition (18) yields

$$M[X(\tau^P, \theta_i, \beta_i), \theta_i] = W_x[X(\tau^P, \theta_i, \beta_i), \theta_i] - p$$

$$= C_x[X(\tau^P, \theta_i, \beta_i)] - \frac{\sum_{j \neq i} n_j X_\tau^j}{\sum_{j=1}^3 n_j X_\tau^j}(1 - \beta_i)C_x[X(\tau^P, \theta_i, \beta_i)]$$

$$+ \frac{\sum_{j \neq i} n_j X_\tau^j(1 - \beta_j)C_x[X(\tau^P, \theta_j, \beta_j)]}{\sum_{j=1}^3 n_j X_\tau^j}$$

(25)

for $i = 1, 2, 3$. This a system of three equations determining the food consumption levels of the three consumer types with uniform taxation, i.e., $X(\tau^P, \theta_i, \beta_i)$ for $i = 1, 2, 3$.

The analysis of (24) and (25) is now much more complex than in the case of type-specific taxes. However, we can work out an important insight if we consider the constellation with $x_1^T < x_2^T = x^H < x_3^T$. In this case, type 1 individuals are underweight, type 2 individuals have healthy weight and type 3 individuals are overweight. We then obtain

Proposition 7 *Suppose* $x_1^T < x_2^T = x^H < x_3^T$. *Moreover, assume* $\tau^P > 0$. *Then,*

(a) $X(\tau^P, \theta_1, \beta_1) < x_1^o < X(\tau_1^P, \theta_1, \beta_1) < x^H$,
(b) $X(\tau^P, \theta_2, \beta_2) < X(\tau_2^P, \theta_2, \beta_2) = x_2^o = x^H$,
(c) $x^H < X(\tau_3^P, \theta_3, \beta_3) < X(\tau^P, \theta_3, \beta_3) < x_3^o$.

If the optimal uniform fat tax is strictly positive ($\tau^P > 0$), which seems to be the most relevant case, then none of the three types realizes healthy weight at the extensive margin. According to part (c) of Proposition 7, overweight individuals (type 3) reduce their calorie intake at the intensive margin as reaction to the uniform tax, i.e. $X(\tau^P, \theta_3, \beta_3) < x_3^o$, but realize a higher calorie intake than in case of type-specific taxation, i.e. $X(\tau_3^P, \theta_3, \beta_3) < X(\tau^P, \theta_3, \beta_3)$, which is already higher than healthy calorie intake x^H. The reason is that, in order to realize healthy weight, overweight individuals have to be taxed more heavily than in the case of a type-specific policy. But the uniform tax reflects the weighted average of the self-control problems of all three types and, thus, is lower than the type-specific tax on the overweights. Moreover, parts (a) and (b) of Proposition 7 show that the uniform fat tax renders the underweight individuals of type 1 even more underweight, i.e. $X(\tau^P, \theta_1, \beta_1) < X(\tau_1^P, \theta_1, \beta_1) < x^H$, and additionally incentivizes normal weight individuals of type 2 to become underweight, $X(\tau^P, \theta_2, \beta_2) < X(\tau_2^P, \theta_2, \beta_2) = x^H$. The reason is that underweight individuals actually should be subsidized and normal weight individuals should not be taxed at all. In sum, the positive effect of the optimal uniform tax on the overweights is smaller than in case of type-specific fat taxes and, in addition, at the costs of underweights and normal weights that (further) reduce their calorie intake below the healthy level.

In order to derive these insights in Proposition 7, we assume that the optimal uniform tax is positive. This will be the case if overweight is more prevalent than

Table 4 Estimated parameter values [a]Source: n_i from NHANES 2017-18 (CDC 2021), x_i^o, θ_i from authors' calculations

Parameter	Value[a]
n_1:	0.016
n_2:	0.256
n_3:	0.732
$\beta_i, i = 1, 2, 3$:	0.6
x_1^o:	1856.43
x_2^o:	2152.47
x_3^o:	2800.36
θ_1:	0.636
θ_2:	0.776
θ_3:	1.087

underweight, such that averaging in τ^P leads to a positive sign. From the theoretical analysis, however, we cannot exclude that $\tau^P \leq 0$. In order to prove $\tau^P > 0$ and, in addition, to estimate the extent of the adverse effects on underweight and normal weight individuals, we now again calibrate the theoretical model to data on US adults.

Empirical calibration and simulation results We first use the NHANES data to determine the proportion of US adults with under-, healthy- and overweight (the groups from columns (2), (3) and (4) in Table 1). We estimate that 1.6% of US adults are underweight, 25.2% have a normal BMI $\in [18.5, 25)$ and 73.2% have a BMI above 25. Therefore, we simulate the theoretical model by setting $n_1 = 0.016, n_2 = 0.252$, and $n_3 = 0.732$ (we report the estimated parameter values in Table 4).

In the benchmark case, we assume that all types have the same present-bias $\beta_i = 0.6$ for $i = 1, 2, 3$. By keeping β_i equal to the value of β in the section with a representative consumer, we can isolate the impact of preference heterogeneity on the optimal tax rate. In the sensitivity analysis, we vary β_i to look at the effects of two heterogeneity sources. Moreover, we set $x_2^o = x^H = 2152.47$ from Table 2. The values x_i^o for $i = 1, 3$ are derived similarly to x^o from the previous section using Hall et al. (2011)'s weight gain estimate of 0.01594 BMI/kcal/day. Starting from $x^H = 2152.47$ and a healthy BMI of 22.28, an overweight adult should consume $x_3^o = 2800.36$ kcal/day to attain a BMI equal to 32.61, while a BMI of an underweight individual equal to 17.56 can be maintained by $x_1^o = 1856.43$ kcal/day. Furthermore, consumption utility of type i is specified as $V(z_i, x_i; \theta_i) = z_i + \theta_i x_i^{1-\gamma} / (1 - \gamma)$. The health cost function remains unchanged. We assume that c, γ, p and e are identical for all types and take the same values as in Table 2. Thus, we can find θ_i as the preference parameters that satisfy the first-order conditions (19). We find $\theta_1 = 0.636 < \theta_2 = 0.776 < \theta_3 = 1.087$.

Our simulation results are summarized in Table 5. First, the type-specific tax rates are given by $\tau_1^P = -0.03$ cents/kcal, $\tau_2^P = 0$ and $\tau_3^P = 0.068$ cents/kcal. Underweight individuals should be subsidized, while the optimal tax on overweights is

Table 5 Results

Parameter	Value[a]
τ_1^P in cents/kcal:	$-0.03 \approx 6.8\%$ of p
τ_2^P in cents/kcal:	0
τ_3^P in cents/kcal:	$0.068 \approx 15.4\%$ of p
τ^P in cents/kcal:	$0.049 \approx 11\%$ of p
$X_1(\tau_1^P, \theta_1, \beta_1)$ in kcal/day:	1971.08
$X_3(\tau_3^P, \theta_3, \beta_3)$ in kcal/day:	2550.99
$X_1(\tau^P, \theta_1, \beta_1)$ in kcal/day:	1674.98
$X_2(\tau^P, \theta_2, \beta_2)$ in kcal/day:	1971.35
$X_3(\tau^P, \theta_3, \beta_3)$ in kcal/day:	2619.71
$EV_1(\tau^P)$ in \$/month:	-4.92
$EV_2(\tau^P)$ in \$/month:	-2.17
$EV_3(\tau^P)$ in \$/month:	3.82
$\overline{EV}(\tau^P)$ in \$/month:	2.17

[a]Source: authors' calculations

positive and larger than the optimal tax on the representative agent in the homogeneous model. Trivially, the healthy weight adults are not taxed. The optimal uniform paternalistic tax is $\tau^P = 0.049$ and is almost equal to the one from the representative agent model (they differ slightly only after the third decimal point). As a consequence, the optimal individual tax τ_3^P lowers the calorie intake of type 3 individuals to 2550.99 kcal/day, while the optimal uniform tax lowers it only to 2619.71 kcal/day. An overweight's welfare gain from the uniform tax measured by the equivalent variation is $EV_3(\tau^P) = \$3.82$ per month. However, the average underweight individual further lowers the calorie intake under uniform taxation, while they increase the calorie intake under type-specific taxation. Thus, an underweight consumer has a welfare loss under uniform taxation equal to $EV_1(\tau^P) = -\$4.92$. The average healthy weight individual also lowers the energy intake under uniform taxation which leads to a utility loss of $EV_2(\tau^P) = -\$2.17$ per month. Owing to the large prevalence of overweight, the average equivalent variation is positive and equals $\overline{EV}(\tau^P) = \$2.17$ per month. Interestingly, this value is almost equal to the equivalent variation in the representative agent model of \$2.16 per month. Hence, the simulation with heterogenous individuals supports the simulation in the case of a representative individual and additionally reveals the distributional impact of calorie taxation. Most importantly, while overweights gain even from uniform taxation, underweights realize a substantial welfare loss and the normal weights' utility decreases, too. Note that this insight confirms the conclusions derived from Proposition 7.

Sensitivity analysis In the sensitivity analysis, we additionally allow for heterogeneity in the present bias parameter β_i (details on the simulation can be found in an online appendix). In a first step, we allow the overweight individuals of type 3 to be more present-biased and set $\beta_3 = 0.2$, while keeping $\beta_1 = \beta_2 = 0.6$. This change slightly lowers the optimal uniform tax to $\tau^P = 0.047$. The welfare loss of

underweights and normal weights decreases slightly, while the welfare gain of the overweights increases substantially to $EV_3(\tau^P) = \$16.58$ per month, so the average welfare gain increases to $\overline{EV}(\tau^P) = \$11.55$ per month. In the second modification, we assume that underweights are also more present-biased and suppose $\beta_1 = \beta_3 = 0.2$, while the normal weights still have $\beta_2 = 0.6$. The optimal uniform tax now further decreases to $\tau^P = 0.046$. The welfare loss of the underweights increases considerably to $EV_1(\tau^P) = -\$26.63$ per month, whereas the welfare changes of normal weights and overweights are almost the same as in the first robustness check. Moreover, due to the low number of underweights, the average welfare change remains almost unchanged at $\overline{EV}(\tau^P) = \$11.19$ per month.

5 Conclusion

Our theoretical analysis shows that the optimal paternalistic tax on unhealthy food consumption internalizes the self-control internality that the representative individual inflicts on itself at the intensive margin, but it leaves uncorrected rational motives for an unhealthy diet choice at the extensive margin. Targeting also the extensive margin and, thus, inducing the individual to choose healthy weight, requires a further distortion that may overcompensate the beneficial effects and, thereby, render food taxation inferior to non-taxation. With heterogeneous individuals, the optimal paternalistic tax is to the advantage of overweight individuals, but reduces welfare of underweight and normal weight individuals, since they obtain incentives to reduce their calorie intake.

Our empirical calibration to data on US adults confirms these insights from the theoretical model and, in addition, reveals that the welfare gains from paternalistic taxation are modest (on average around $2 per month). However, the simulation results in case of consumer heterogeneity also allow a more differentiated view due to the results on the distributional impact of unhealthy food taxation. Underweight and normal weight individuals may suffer a welfare loss, indeed, but overweight individuals may realize substantial welfare gains. This result may be used to motivate (tax) policies that are more targeted at overweights, leaving the other types of individuals more or less unaffected. Sugar or soda taxes may be good examples for such a targeted policy, as such goods are usually consumed to a larger proportion by overweights. Similarly, subsidies to calorie intake from healthier foods, such as, e.g., fruits and vegetables, may compensate the under- and healthy weight individuals for their utility losses.

Of course, this argument also reveals a drawback of our analysis. We only take into account a single food good, whereas a more sophisticated analysis would distinguish between several food goods that impact the individuals' health to different degrees. The idea would then be that overweight individuals consume a larger portion of unhealthy food goods and, therefore, will benefit more from the taxation of these goods, while underweight and normal weight individuals are affected to a lesser extent. Assuming monotone health costs, Kalamov (2020) analyzes the taxation of two different food goods that differ in their calorie content, and Arnabal (2021) studies the optimal taxation of one sin good when complementary/

substitute sin goods are nontaxable. However, to the best of our knowledge, no paper has considered the optimal taxation of multiple unhealthy goods under non-monotone health costs, so that individuals may have over-, under- or healthy weight. Because different food goods likely differ in other nutritional components than simply calorie content, a corresponding theoretical and empirical analysis is much more complex and, therefore, left for future research.

Appendix

Proof of Proposition 1 Concavity of V implies $V_{xx}(\cdot)V_{zz}(\cdot) - [V_{zx}(\cdot)]^2 > 0$ or, equivalently, $V_{xx}(\cdot) < [V_{zx}(\cdot)]^2/V_{zz}(\cdot)$. From the definition of $N(x)$, we then obtain $N_x(\cdot) = V_{xx}(\cdot) + p^2 V_{zz}(\cdot) - 2p V_{zx}(\cdot) < [V_{zx}(\cdot) - p V_{zz}(\cdot)]^2/V_{zz}(\cdot) < 0$. Hence, the function $N(x)$ is decreasing with a zero at $x = x^T$. Due to $C_{xx}(x) > 0$, the functions $C_x(x)$ and $\beta C_x(x)$ are both increasing with a zero at $x = x^H$. Hence, if $x^H < x^T$, then the intersection between $N(x)$ and $\beta C_x(x)$, as required for the first-order condition (6), lies in the interval $]x^H, x^T[$ where $C_x(x) > 0$. It follows $x^H < x^o < x^T$, as stated in (a) of Proposition 1. The proofs of parts (b) and (c) are perfectly analogous. □

Proof of Proposition 2 Remember that $N(x)$ is decreasing with a zero at $x = x^T$, whereas $C_x(x)$ and $\beta C_x(x)$ are both increasing with a zero at $x = x^H$. Using (10), the property that $X(\tau^P)$ always lies between x^T and x^H is proven by the same steps as we show in Proposition 1 that x^o lies between x^T and x^H. In order to show the relation between x^o and $X(\tau^P)$, note that the former is determined by (6) whereas the latter is determined by (10). While the RHS of (6) equals $\beta C_x(\cdot)$, the RHS of (10) only contains $C_x(\cdot)$. For $x > x^H$ ($x < x^H$) the function $\beta C_x(x)$ is below (above) the function $C_x(x)$ since $\beta < 1$. Hence, if $x^H < x^T$, the intersection of $N(x)$ with $C_x(x)$, as required for the determination of $X(\tau^P)$, lies to the left of the intersection of $N(x)$ with $\beta C_x(x)$, as required for the determination of x^o. It follows $X(\tau^P) < x^o$, as stated in part (a) of Proposition 2. This argument is reversed if $x^T < x^H$, proving the statement in part (b). For $x^T = x^H$, all three functions $N(x)$, $\beta C_x(x)$ and $C_x(x)$ intersect at the same point on the x-axis implying $x^o = X(\tau^P)$, as stated in part (c) of Proposition 2.

It remains to prove the sign of the optimal tax rate τ^P. If $x^H < x^T$, we know $X(\tau^P) > x^H$ and, thus, $C_x[X(\tau^P)] > 0$. Inserting into (9) yields $\tau^P > 0$, as stated in part (a) of Proposition 2. Conversely, for $x^T < x^H$ we obtain $X(\tau^P) < x^H$ and therefore $C_x[X(\tau^P)] < 0$ and $\tau^P < 0$, while $x^T = x^H$ implies $C_x[X(\tau^P)] = 0$ and $\tau^P = 0$. This shows part (b) and part (c) and completes the proof of Proposition 2. □

Proof of Proposition 3 It only remains to prove parts (a)–(c). We can rewrite the health-maximi3zing tax rate (11) as $\tau^H = N(x^H)/V_z(\cdot)$. The term $V_z(\cdot)$ is positive, so sign $\{\tau^H\} = $ sign $\{N(x^H)\}$. Since $N(x)$ is decreasing in x and $N(x^T) = 0$, the relation $x^H < x^T$ implies $N(x^H) > 0$ and $\tau^H > 0$, as stated in part (a) of Proposition 3. Analogously, $x^T < x^H$ yields $N(x^H) < 0$ and $\tau^H < 0$, whereas for $x^T = x^H$ we obtain $N(x^H) = 0$ and $\tau^H = 0$. This completes the proof of Proposition 3. □

Proof of Proposition 4 *Part (a) is obvious. Under welfare maximization, we maximize true utility u with respect to the tax rate τ. The solution is τ^P given by (9) and maximized true utility u^P defined in (13). Since τ^H from (11) is not equal to τ^P, it does not maximize true utility and it follows $u^H < u^P$, as stated in Proposition 4 (a).*

In order to prove part (b), let us first determine the impact of β on u^o and u^H. From (14), we see that u^H is independent of β. The first-order condition (6) implies

$$\frac{dx^o}{d\beta} = \frac{C_x}{V_{xx} + p^2 V_{zz} - 2p V_{xz}}. \tag{26}$$

Differentiating (12) then yields

$$\frac{du^o}{d\beta} = (V_x - pV_z - C_x)\frac{dx^o}{d\beta} = (\beta - 1)\frac{(C_x)^2}{V_{xx} + p^2 V_{zz} - 2p V_{xz}} > 0, \tag{27}$$

where we have used (6) and (26). Note that $V_{xx} + p^2 V_{zz} - 2p V_{xz} < 0$ equals the slope of $N(\cdot)$ which we have proven to be negative for all x, so also for $x = x^o$. According to (27), a reduction in β reduces u^o, while u^H remains unchanged.

Now start with the largest possible value of β, which is $\beta = 1$. From (9), we then obtain $\tau^P = 0$. Equations (6) and (10) imply $X(\tau^P) = X(0) = x^o$, so we obtain $u^P = u^o$ from (12) and (13). Hence, for $\beta = 1$ the utility without taxation (u^o) is equal to the maximum utility under paternalistic welfare maximization (u^P) which, for $x^T \neq x^H$, is always larger than utility under health maximization (u^H), as proven by part (a) of Proposition (4). We therefore have proven $u^H < u^o$ for the limiting case of $\beta = 1$. Reducing β below one, leaves unchanged u^H and decreases u^o according to (27), so u^o may fall below u^H. Thus, there exists a $\bar{\beta} \in [0, 1[$ such that $u^H \gtreqless u^o$ iff $\beta \gtreqless \bar{\beta}$, which completes the proof of part (b) of Proposition 4. Note that this statement includes the case in which $\bar{\beta} = 0$, so that $u^H < u^o$ for all $\beta \in [0, 1]$: at the lowest possible value $\beta = 0$, we obtain $x^o = x^T$ from (6), so (12) can be rewritten as $u^o = V(e - px^T, x^T) - C(x^T)$. Using (14), it follows $u^H < u^o$ if and only if $V(e - px^H, x^H) - V(e - px^T, x^T) + C(x^T) < 0$. Hence, if the later condition is satisfied, then we have $u^H < u^o$ even at the lowest possible value of β and, thus, $\bar{\beta} = 0$. □

Proof of Proposition 5 The proof of Proposition 1 can be applied to each type of individuals, except that we have to replace N by M. This proves the first sentence in Proposition 5. In order to show part (a), note that for $\beta_i = \beta_j$ and $\theta_i < \theta_j$ we obtain $x_i^o := X(0, \theta_i, \beta_i) < X(0, \theta_j, \beta_j) = x_j^o$ due to $X_\theta(\cdot) > 0$. For the proof of part (b), we have to take into account that $\theta_i = \theta_j$ implies $x_i^T = x_j^T$ and $M(x, \theta_i) = M(x, \theta_j)$ for all x. Moreover, for $\beta_i < \beta_j$, the function $\beta_j C_x(x)$ is steeper than the function $\beta_i C_x(x)$, but both have the same zero at $x = x^H$. Hence, if $x^H < x_i^T = x_j^T$, then the intersection of $\beta_j C_x(x)$ with $M(x, \theta_i) = M(x, \theta_j)$ lies to the left of the intersection of $\beta_i C_x(x)$ with $M(x, \theta_i) = M(x, \theta_j)$. It follows $x_j^o < x_i^o$ from (19). This argument is reversed for $x_i^T = x_j^T < x^H$, implying $x_j^o > x_i^o$. Finally, if $x_i^T = x_j^T = x^H$, then $\beta_i C_x(x)$, $\beta_j C_x(x)$ and $M(x, \theta_i) = M(x, \theta_j)$ all intersect at $x = x^H$, implying $x_i^o = x_j^o$. □

Proof of Proposition 6 *The proof of Proposition 2 can be applied to each type of individuals, except that we have to replace N by M. This proves the first sentence in Proposition 6. The relation $X(\tau_i^P, \theta_i, \beta_i) < X(\tau_j^P, \theta_j, \beta_j)$ for $\beta_i = \beta_j$ and $\theta_i < \theta_j$ follows from $x_i^T < x_j^T$, $M_x < 0$ and $M_\theta > 0$, since the intersection between $M(\cdot)$ and $C_x(\cdot)$ for type i individuals lies to the left of the intersection between $M(\cdot)$ and $C_x(\cdot)$ for type j individuals. The inequality $X(\tau_i^P, \theta_i, \beta_i) < X(\tau_j^P, \theta_j, \beta_j)$ implies $C_x[X(\tau_i^P, \theta_i, \beta_i)] < C_x[X(\tau_j^P, \theta_j, \beta_j)]$ and, thus, $\tau_i^P < \tau_j^P$ by (21) and $\beta_i = \beta_j$, which completes the proof of part (a). In order to prove part (b), note that $\theta_i = \theta_j$ implies $x_i^T = x_j^T$ and the same function $M(\cdot)$ for i and j. Condition (22) then implies $X(\tau_i^P, \theta_i, \beta_i) = X(\tau_j^P, \theta_j, \beta_j)$, independently of β_i and β_j since these parameters are not contained in (22). Note that $X(\tau_i^P, \theta_i, \beta_i) = X(\tau_j^P, \theta_j, \beta_j)$ implies $C_x[X(\tau_i^P, \theta_i, \beta_i)] = C_x[X(\tau_j^P, \theta_j, \beta_j)]$. Hence, for $x_i^T = x_j^T > x^H$, the intersection between $M(\cdot)$ and $C_x(\cdot)$ is above the x-axis, so $C_x[X(\tau_i^P, \theta_i, \beta_i)] = C_x[X(\tau_j^P, \theta_j, \beta_j)] > 0$, which together with $\beta_i < \beta_j$ gives $\tau_i^P = (1 - \beta_i)C_x[X(\tau_i^P, \theta_i, \beta_i)] > (1 - \beta_j)C_x[X(\tau_j^P, \theta_j, \beta_j)] = \tau_j^P > 0$. This argument is reversed for $x_i^T = x_j^T < x^H$, yielding $\tau_i^P = (1 - \beta_i)C_x[X(\tau_i^P, \theta_i, \beta_i)] < (1 - \beta_j)C_x[X(\tau_j^P, \theta_j, \beta_j)] = \tau_j^P < 0$. Finally, $x_i^T = x_j^T = x^H$ gives $C_x[X(\tau_i^P, \theta_i, \beta_i)] = C_x[X(\tau_i^P, \theta_i, \beta_i)] = 0$ and $\tau_i^P = (1 - \beta_i)C_x[X(\tau_i^P, \theta_i, \beta_i)] = (1 - \beta_j)C_x[X(\tau_j^P, \theta_j, \beta_j)] = \tau_j^P = 0$.* □

Proof of Proposition 7 *The relationships between x^H, x_i^o and $X(\tau_i^P, \theta_i, \beta_i)$ for $i = 1, 2, 3$ can immediately be derived from Proposition 5 and Proposition 6. It remains to show the relation of $X(\tau^P, \theta_i, \beta_i)$ to x^H, x_i^o and $X(\tau_i^P, \theta_i, \beta_i)$ for $i = 1, 2, 3$. In parts (a) and (b), the relations $X(\tau^P, \theta_1, \beta_1) < x_1^o = X(0, \theta_1, \beta_1)$ and $X(\tau^P, \theta_2, \beta_2) < X(\tau_2^P, \theta_2, \beta_2) = x_2^o = X(0, \theta_2, \beta_2)$ follow from X_τ^1, $X_2^\tau < 0$, $\tau_2^P = 0$ (since $x_2^T = x^H$) and $\tau^P > 0$. The same argument applies to the statement $X(\tau^P, \theta_3, \beta_3) < x_3^o = X(0, \theta_3, \beta_3)$ in part (c). In order to prove $X(\tau_3^P, \theta_3, \beta_3) < X(\tau^P, \theta_3, \beta_3)$, note that $X(\tau_3^P, \theta_3, \beta_3)$ is determined by (22) for $i = 3$, which reads*

$$M[X(\tau_3^P, \theta_3, \beta_3), \theta_3] = C_x[X(\tau_3^P, \theta_3, \beta_3)], \tag{28}$$

whereas $X(\tau^P, \theta_3, \beta_3)$ is determined by (25) for $i = 3$, which can be rewritten as

$$\begin{aligned}
M[X(\tau^P, \theta_3, \beta_3), \theta_3] = &[1 - (1 - \beta_3)\chi]C_x[X(\tau^P, \theta_3, \beta_3)] \\
&+ \frac{\sum_{j \neq 3} n_j X_\tau^j (1 - \beta_j)C_x[X(\tau^P, \theta_j, \beta_j)]}{\sum_{j=1}^3 n_j X_\tau^j},
\end{aligned} \tag{29}$$

with $\chi := (\sum_{j \neq 3} n_j X_\tau^j)/(\sum_{j=1}^3 n_j X_\tau^j) \in]0, 1[$ due to $X_\tau^i < 0$ for $i = 1, 2, 3$. Note that $1 - (1 - \beta_3)\chi \in]0, 1[$ since $\beta_3 \in]0, 1[$ and $\chi \in]0, 1[$. Moreover, from parts (a) and (b) we know $X(\tau^P, \theta_j, \beta_j) < x^H$ and, thus, $C_x[X(\tau^P, \theta_j, \beta_j)] < 0$ for $j = 1, 2$. Hence, the term in the second row of (29) is negative and we obtain

$$[1 - (1 - \beta_3)\chi]C_x(x) + \frac{\sum_{j \neq 3} n_j X_\tau^j (1 - \beta_j)C_x[X(\tau^P, \theta_j, \beta_j)]}{\sum_{j=1}^3 n_j X_\tau^j} < C_x(x),$$

for all $x > x^H$. *The RHS of ([29](#)) therefore is always below the RHS of ([28](#)) for $x > x^H$. From $M_x < 0$ it then follows $X(\tau_3^P, \theta_3, \beta_3) < X(\tau^P, \theta_3, \beta_3)$, which completes the proof.* □

References

Allcott, H., Lockwood, B. B., & Taubinsky, D. (2019a). Regressive Sin taxes, with an application to the optimal soda tax. *Quarterly Journal of Economics, 134*(3), 1557–1626.

Allcott, H., Lockwood, B. B., & Taubinsky, D. (2019b). Should we tax sugar-sweetened beverages? An overview of theory and evidence. *Journal of Economic Perspectives, 33*(3), 202–227.

Andersen, S., Harrison, G. W., Lau, M. I., & Rutström, E. E. (2014). Discounting behavior: A reconsideration. *European Economics Review, 71*, 15–33.

Arnabal, L. R. (2021). Optimal design of sin taxes in the presence of nontaxable sin goods. *Health Economics, 30*, 1580–1599.

BEA (2021). Table 2.1 Personal income and its disposition. U.S. Bureau of economic analysis. Retrieved 12 May, 2021, from https://www.bea.gov/ national/index.htm.

Becker, G. S., & Murphy, K. M. (1988). A theory of rational addiction. *Journal of Political Economy, 96*, 675–700.

CDC (2021). National Health and Nutrition Examination Survey 2017-2018. Centers for Disease Control and Prevention, National Center for Health Statistics. Retrieved 12 May, 2021, from wwwn.cdc.gov/nchs/nhanes/continuousnhanes/default.aspx?BeginYear=2017.

Courtemanche, C., Heutel, G., & McAlvanah, P. (2014). Impatience, incentives and obesity. *Economic Journal, 125*, 1–31.

Cremer, H., De Donder, P., Maldonado, D., & Pestieau, P. (2012). Taxing Sin goods and subsidizing health care. *Scandinavian Journal of Economics, 114*, 101–123.

Cremer, H., Goulao, C., & Roeder, K. (2016). Earmarking and the political support of fat taxes. *Journal of Health Economics, 50*, 258–267.

Cutler, D., Glaeser, E., & Shapiro, J. (2003). Why have Americans become more obese? *Journal of Economic Perspectives, 17*, 93–118.

Dragone, D. (2009). A rational eating model of binges, diets and obesity. *Journal of Health Economics, 28*, 799–804.

Dragone, D., & Savorelli, L. (2012). Thinness and obesity: A model of food consumption, health concerns, and social pressure. *Journal of Health Economics, 31*, 243–256.

Global BMI Mortality Collaboration. (2016). Body-mass index and all-cause mortality: Individual-participant-data meta-analysis of 239 prospective studies in four continents. *Lancet, 388*, 776–786.

Griffith, R., Oconnell, M., & Smith, K. (2017). Corrective taxation and internalities from food consumption, *CESifo Economic Studies*, 1–14.

Gruber, J., & Kőszegi, B. (2001). Is addiction rational? *Theory and Evidence, Quarterly Journal of Economics, 116*, 1261–1303.

Haavio, M., & Kotakorpi, K. (2011). The political economy of Sin taxes. *European Economic Review, 55*, 575–594.

Haavio, M., & Kotakorpi, K. (2016). Sin licenses revisited. *Journal of Public Economics, 44*, 40–51.

Hall, K. D., Sacks, G., Chandramohan, D., Chow, C. C., Wang, Y. C., Gortmaker, S. L., & Swinburn, B. A. (2011). Quantification of the effect of energy imbalance on bodyweight. *Lancet, 378*(9793), 826–837.

Harding, M., & Lovenheim, M. (2017). The effect of prices on nutrition: Comparing the impact of product- and nutrient-specific taxes. *Journal of Health Economics, 53*, 53–71.

Ikeda, S., Kang, M. I., & Ohtake, F. (2010). Hyperbolic discounting, the sign effect, and the body mass index. *Journal of Health Economics, 29*, 268–284.

Kalamov, Z. (2020). A sales tax is better at promoting healthy diets than the fat tax and the thin subsidy. *Health Economics, 29*, 353–366.

Kalamov, Z., & Runkel, M. (2020). Present-focused preferences and Sin good consumption at the extensive and intensive margins, CESifo Working Paper 8237.

Laibson, D. (1997). Hyperbolic discounting and golden eggs. *Quarterly Journal of Economics, 112*, 443–477.

Levy, A. (2002). Rational eating: Can it lead to overweightness or underweightness? *Journal of Health Economics, 21*, 887–899.

Loewenstein, G., & Chater, N. (2017). Putting nudges in perspective. *Behavioural Public Policy, 1*, 26–53.

O'Donoghue, T., & Rabin, M. (2003). Studying optimal paternalism, illustrated by a model of sin taxes. *American Economic Review, 93*, 186–191.

O'Donoghue, T., & Rabin, M. (2006). Optimal sin taxes. *Journal of Public Economics, 90*, 1825–1849.

OECD (2017). *Obesity Update 2017*, https://www.oecd.org/els/health-systems/Obesity-Update-2017.pdf (download November 2017).

U.S. Department of Health and Human Services and U.S. Department of Agriculture (2015). 2015–2020 Dietary Guidelines for Americans, 8th Edition.

WHO (2016a). Putting taxes into the diet equation. *Bulletin World Health Organization, 94*, 239–240.

WHO (2016b). Taxes on sugary drinks: Why do it?', http://apps.who.int/iris/ bitstream/10665/250303/1/ WHO-NMH-PND-16.5-eng.pdf (download November 2017).

WHO (2017). *Obesity and overweight: Fact sheet*. http://www.who.int/mediacentre/ factsheets/fs311/en/ (download November 2017).