Body Weight, Dieting and Obesity Traps

Barbieri, Paolo Nicola

University of Gothenburg, Centre for Health Economics

October 2015

Body Weight, Dieting and Obesity Traps

Paolo Nicola Barbieri[*][†]

November 6, 2015

Abstract

This paper presents a theoretical investigation into why losing weight is so difficult even in the absence of rational addiction, time-inconsistent preferences or bounded rationality. We add to the existing literature by focusing on the role that individual metabolism has on weight loss. The results from the theoretical model provides multiple steady states and a threshold revealing a situation of "obesity traps" that the individual must surpass in order to successfully lose weight. Given such a threshold we investigate if a short-run incentive scheme for weight loss is able to promote persistent weight-losses and what features an incentive scheme should have in order to sustain permanent weight loss. We find that a lump-sum incentive scheme is not always able to lead the individual to permanent weight loss. On the contrary, a non-decreasing incentive scheme with rewards for weight loss according to levels of body weight (i.e. progressive), is able to sustain a steady state reduction in body weight and food consumption.

JEL-Classification: D91, I12, I18

Keywords: Obesity, Dieting, Optimal Control, Multiple Equilibria.

[*]I am thankful to Davide Dragone, Enrico Bernardi, Francesca Barigozzi, Kristian Bolin, and seminar participants at University of Bologna, University of Gothenburg, the Summer School of Health Economics and Game Theory, the Institute for New Economic Thinking (INET), the Center for International Governance Innovation (CIGI) for insightful comments.

[†]Corresponding author. Centre for Health Economics, University of Gothenburg, Vasagatan 1, S-41124 Goteborg, Sweden E-mail address: `paolonicola.barbieri@economics.gu.se`.

1 Introduction

Over the last decade the incidence of obesity has increased dramatically (Ogden and Carroll, 2010, von Ruesten et al., 2011), leading to greater concerns about the promotion of efficient ways to invert this trend by means of dieting (Baradel et al., 2009, Bish et al., 2012) or physical activity (Mozaffarian et al., 2011). However, although attempts at losing weight may have increased, more than 50% of obese individuals still struggle to lose weight permanently (Kruger et al., 2004). Due to these unsuccessful attempts and in order to facilitate weight-loss efforts, recent insights deriving from behavioral economics (Loewenstein et al., 2007) have been used to promote the reduction of excessive weight by motivating individuals to follow a type of behavior that they would not naturally follow thanks to monetary and non-monetary incentives[1] (Augurzky et al., 2012, Cawley and Price, 2013, John et al., 2011, Volpp et al., 2008). Unfortunately, despite promising short-run results and the fact that weight loss was indeed achieved during experiments, substantial weight regain was found after the incentive was removed (John et al., 2011, Volpp et al., 2008, Cawley and Price, 2013), proving that, despite being sustained by incentives, individuals were still failing to lose weight permanently. This poses the questions of why weight loss is so difficult to achieve, what factors influence weight loss and how they even manage to negate the effects of short-run incentives. Current literature relates difficulties in losing weight to several possible explanations such as: time-inconsistent preferences (Dodd, 2008, Ikeda et al., 2010); willpower depletion (Ozdenoren et al., 2006); the interaction between two conflicting systems driving human behavior (cognitive vs. affective) (Loewenstein and O'Donoghue, 2004, Ruhm, 2012); or, more generally, to the failure of individuals to rationally balance current benefits and future costs related to food consumption and body weight (O'Donoghue and Rabin, 1999).

In this paper we add to existing literature by proposing another, complementary, explanation for why dieting is so difficult, even in the absence of rational addiction, time-inconsistent preferences or bounded rationality. To do so we develop an original assumption regarding an individual's metabolism. Current economic literature considers individual metabolism to be linearly related to body weight, such that the more an individual weighs the higher his or her caloric expenditure will be (Dragone, 2009, Levy, 2002, Harris and Benedict, 1918). The key assumption is that calories expended by the metabolism in order to keep the basic functions of the organism running (i.e. basal metabolic rate), is a linear and increasing function of body weight; meaning that of two individuals of a similar age, the one with a higher body weight (i.e. the one with the greater "weight stock") will burn a higher fraction of it to keep his organism running. This assumption entails that the marginal effect of an additional kg on an individual's caloric expenditure is always positive and that the caloric expenditure of obese/overweight individuals is therefore higher than that of leaner individuals, simply because they weigh more.

This notion of body weight accumulation and energy expenditure contradicts recent insights from medical literature, according to which metabolism actively and non-monotonically reacts to changes in body weight (Catenacci et al., 2011, Ebbeling et al., 2012, Gale et al., 2004, Katan and Ludwig, 2010, Leibel and Hirsch,

[1] Experimental attempts used either lump-sum payments or lotteries, in order to exploit individual fallacies in the estimation of the probability of an unlikely event, to motivate subjects to lose weight (Kahneman and Tversky, 1979).

1984, Leibel et al., 1995), such that obese/overweight individuals do not burn more calories simply because they weigh more (Johnstone et al., 2005, Cunningham, 1991, Astrup et al., 1999). Even given the same level of physical activity and food intake, the higher adipose content in the body composition of obese/overweight individuals contributes in reducing their caloric expenditure rather than increasing it, since body fat negatively affects an individual's caloric expenditure. Mifflin et al. (1990), Cunningham (1991), Wang et al. (2000) report that the main predictor for total energy expenditure is the level of non-adipose mass that the individual possesses and that fat mass is not one of its main predictors. Moreover, Johnstone et al. (2005) quantify that fat mass accounts for less than 6% of basal metabolic rate variability, while fat-free mass accounts for almost 70%. Based on this evidence we can conclude that given the minimal incidence that fat mass has on an individual's caloric expenditure and the fact that obese/overweight individuals have a positive imbalance of fat mass over fat-free mass, the relationship between body weight and caloric expenditure is non-linear (Thomas et al., 2009, 2011). This means that of two individuals of the same age, the one who weighs more (i.e. the one with a higher adipose content) is predicted to burn fewer calories with respect to the leaner individual. In order to capture this dynamic we will assume that there exists an individual specific level of body weight over which any additional weight gain will decrease caloric expenditure; conversely any weight gain lower than this threshold will increase caloric expenditure.

The results from the theoretical model using this novel metabolic assumption are threefold. Firstly, we extend the rational eating literature (Dragone, 2009, Levy, 2002) to show the impact that metabolism can have on the steady state decision of a rational individual. Unlike current literature, in which the individual exhibits only a rational and stable outcome of overweight, we allow for multiple equilibria and report one characterized by a healthier body weight. Second, this multiplicity of steady states provides the evidence of an "obesity trap" with an associated threshold dividing the two possible outcomes. Intuitively, any weight-loss efforts that the individual undertakes have to surpass this threshold in order to result in permanent weight loss, otherwise the individual will gradually regain weight and converge to his or her previous body weight. This threshold provides the rationale for any unsuccessful dieting efforts even in the absence of rational addiction, time-inconsistent preferences or bounded rationality. Therefore the presence of such a threshold does indeed have an influence on the possible persistence of dieting efforts. Lastly, to prove the crucial role that metabolic patter have on dieting effort, we provide an explanation for why previous financial incentive scheme were so unsuccessful and propose a possible design for incentives that could sustain a persistent weight loss. A non-decreasing incentive scheme rewarding greater weight loss at a higher level of body weight (i.e. progressive) is able to sustain steady state decreases in body weight and food consumption while a lump-sum payment for weight loss, the most used incentive scheme in experimental literature, is not. This result provide a possible theoretical explanation for the long-run weight regain reported in experimental evidence.

2 The Model

2.1 The Utility Function

Consider an individual whose intertemporal utility depends on food consumption, $c(t) \geq 0$ and body weight, $w(t) > 0$, according to the following utility function

$$V(c(t), w(t)) = U(c(t), w(t)) \tag{1}$$

Where $U(.)$ represents the instantaneous utility from food consumption and the disutility from body weight, while $I(.)$ is utility related to incentives for weight loss. We will assume that the instantaneous utility $U(c, w)$ is twice continuously-differentiable, jointly concave with negative second order derivatives, U_{cc} and U_{ww} and separable, $U_{cw} = 0$.

Following Dragone and Savorelli (2011) we assume that there exists a satiation point for individual food consumption, denoted by c^{sat}, with an associated level of body weight w^{sat}, such that $\partial U(c^{sat})/\partial c = 0$, which is to be interpreted as the (steady state) body weight that an agent would reach if she always ate to satiation[2]. We say the agent is on a diet if $U_c > 0$ (utility would be increased by increasing food consumption) and is binging if $U_c < 0$ (utility would be increased by decreasing food consumption). Regarding body weight we assume that there exists BMI that maximizes each agent's health condition; denote the corresponding body weight as w^H, such that $\partial U(w^H)/\partial w = 0$. We say the agent is overweight ($w > w^H$) if $U_w < 0$ (utility would be increased by decreasing body weight) and conversely underweight ($w < w^H$) if $U_w > 0$ (utility would be increased by increasing body weight). It is possible to distinguish between three possible cases. If $w^{sat} > w^H$, satiation induces being overweight; likewise we say that satiation induces one to be underweight if $w^H > w^{sat}$. Lastly if $w^{sat} = w^H$ satiation corresponds to healthy weight.

2.2 The Dynamics of Individuals Body Weight

The determinants of individual's body weight are given by food consumption, $c(t)$, and current body weight, $w(t)$ which will influence the accumulation of body weight in time as follows

$$\dot{w}(t) = c(t) - g(w(t)) \tag{2}$$

Assumption 1. *The $g(.)$ function is twice continuously-differentiable, positive and concave with a maximum in $g(\bar{w})$*

$g(w(t))$ is a concave function determining the amount of energy expended per unit of body weight, which in medical terms is defined as the basal metabolic rate (Harris and Benedict, 1918), that is, the amount of energy expended by individuals at rest. We assume that the contribution of food consumption to body

[2] c^{sat} corresponds to the solution of a standard constrained optimization problem where the agent must choose between two goods, food and non-food, without suffering from any health consequences.

weight is positive and linear. Additionally, the contribution of body weight to the rate of caloric expenditure is assumed to be positive (i.e. $g_w \geq 0$) for $w \in [0, \bar{w}]$ and decreasing afterwards, with $g_w < 0$. The idea is the following: the more an individual eats the more weight he will gain; the more weight he gains, the more calories he will burn until he reaches an individual specific body weight, $\bar{w}$, where the effect of additional body weight will actually decrease this basal metabolic rate. This explanation is related to the fact that the main predictor of basal metabolic rate is individual's lean mass (Johnstone et al., 2005, Mifflin et al., 1990, Wang et al., 2000), which is supposed to be lower (in proportion) in heavier individual. Therefore $\bar{w}$ can be thought of as the maximum sustainable level of body weight above which the organism will store additional calories in fat tissue and not in lean mass, resulting then in a net decrease in caloric expenditure since fat mass burns at a lower rate with respect to lean mass.

We will assume that as long as any additional weight gain is resulting in an increases in caloric expenditure, such that $g_w \geq 0$, the individual can be considered has having a positive imbalance between fat free mass and fat mass. Conversely as soon as the individual surpasses surpass $\bar{w}$, resulting in a decreases in caloric expenditure associated with an increases in body weight, $g_w < 0$, the individual can most certainly be considered overweight or obese, since the percentage of fat mass will be strictly higher than the percentage of fat-free mass.

3 Optimality conditions and stability

The individual's goal is to maximize his intertemporal utility by choosing the amount of food consumption, which in turn will affect her body weight. Given an infinite time horizon and a positive discount rate, ρ, the individual's problem can be written as follows (See Appendix A.1):

$$\max_{c(t)} \int_0^\infty e^{-\rho t}[U(c(t), w(t))]dt \tag{3}$$

$$\text{subject to } \dot{w}(t) = c(t) - g(w(t)) \tag{4}$$

$$w(t), c(t) \geq 0 \tag{5}$$

$$w(0) = w_0, \text{ given} \tag{6}$$

Here we have introduced the non-negativity constraint on c and w since they represent, respectively, food consumption and body weight which, by definition, cannot become negative. The associated current value Hamiltonian (dropping the time indexes for convenience) is

$$H(c, w, \lambda) = U(c, w) + \lambda(c - g(w)) \tag{7}$$

3.1 Obesity Traps

Given joint concavity (See Appendix A.2), the first-order conditions for this problem are

$$H_c = 0 \qquad U_c + \lambda = 0 \tag{8a}$$

$$\dot{\lambda} = \lambda(\rho + g_w) - U_w \tag{8b}$$

$$\dot{w} = c - g(w) \tag{8c}$$

If we time-differentiate equation (8a) and substitute it in (8b) along with (8a) we derive an expression for $\dot{c}$, which, along with (8c) results in the following system of differential equations characterizing the equilibrium

$$\begin{aligned} \dot{c} &= \frac{U_c}{U_{cc}}(\rho + g_w) + \frac{U_w}{U_{cc}} \\ \dot{w} &= c - g(w) \end{aligned} \tag{9}$$

For an internal steady state the following conditions must be satisfied

$$\dot{c} = 0 \qquad U_w = -U_c(\rho + g_w) \tag{10}$$

$$\dot{w} = 0 \qquad c^{ss} = g(w^{ss}) \tag{11}$$

Plugging equation (11) into (10) we get

$$U_c(g(w), w)\,(g_w + \rho) + U_w(g(w), w) = 0 \tag{12}$$

In order to establish the multiplicity of steady states we need to prove that equation (12) has more than one possible solution. The number of steady states will be given by the number of solutions that equation (12) admits. However given the implicit form of equation (12) explicitly deriving the number of solutions is not straightforward. Therefore we must go further and differentiate it with respect to w in order to get

$$g_{ww}U_c(g(w)) + g_w\,(g_w + \rho)\,U_{cc}(g(w)) + U_{ww} \lesseqgtr 0 \tag{13}$$

Due to the fact that the marginal utility from food consumption (U_c) and the marginal contribution of body weight to caloric expenditure (g_w) are changing sign two times, the conditions created are sufficient for the non-monotonicity of equation (13) resulting in the possibility of multiple steady states.

Proposition 3.1. *The intertemporal problem (2)-(5) is associated with multiplicity of steady states characterized by the sufficient condition for an unstable threshold*

$$-g_w(\rho + g_w) > \frac{H_{ww}}{U_{cc}} \tag{14}$$

Associated with the following trade-offs between body weight and food consumption according to equation (10)

1. *Being overweight and on a diet,* $U_w < 0$ *and* $U_c > 0$

2. Being underweight and on binging, $U_w > 0$ and $U_c < 0$

(See Appendix A.3)

Interestingly the only case in which eating up to satiation is a steady state outcome is associated to the case in which $c^{sat} = c^H$, with c^H being the amount of food consumption associated with w^H, because $U_c = -U_w = 0$. There is, albeit under specific circumstances, therefore a possible steady state in which the individual is perfectly healthy and consuming to satiation, with no loss of utility whatsoever. More generally, the steady state will not be associated with satiation, but either with binging or dieting, since $U_c \neq -U_w$ but either $U_c > 0$ and $U_w < 0$ (binging and being underweight) or $U_c < 0$ and $U_w > 0$ (dieting and being overweight).

From condition (14) we can also derive an additional (sufficient) condition for multiple steady states. In addition to the unstable threshold in (14), which makes the determinant positive, we can also see that the equation

$$-g_w(\rho + g_w) - \frac{H_{ww}}{U_{cc}}$$

which is the determinant of the Jacobian associated with the intertemporal problem (3)-(6) (See Appendix A.3) can be negative in two different cases, thus allowing two possible stable steady states. The firs one is when $g_w > 0$, while the second one is for $g_w < 0$. This characterization of steady states based on the sign of the derivative of the $g(.)$ function allows us to say that our model permits two stable steady states, one healthier than the other, since $g_w > 0$ is associated, by definition, to a condition of healthy weight, while $g_w < 0$ is associated with overweight/obesity. More generally we will assume that there will be at most three steady states, divided by the threshold defined in equation (14).

Corollary 3.1 (Obesity Traps). *Of the two possible steady states associated with the intertemporal problem (2)-(5), one is characterized by a healthy body weight (i.e. $g_w > 0$) while the other one is characterized by a condition of overweight/obesity (i.e. $g_w < 0$).*

Proposition 3.1 allows us to state that if there are indeed multiple steady states, and condition (14) is therefore satisfied, instability becomes an issue and there exists a threshold such that the individual is trapped in an excessive weight steady state. This result characterize a situation in which there is body resistance to weight-loss that the individual has to surpass in order to permanently lose weight, which can be reconciled with the concept of ”homeostasis” reported in medical evidence (Ebbeling et al., 2012, Gale et al., 2004, Katan and Ludwig, 2010, Leibel and Hirsch, 1984, Leibel et al., 1995). Such that the organism contrast dieting effort since losing weight results in a decrease in energy available and in order to keep the organism in a stable energy state metabolism contrast weight-losses.

To explicitly presents multiple steady states we will resort to specific functions for the utility function and the dynamics of body weight. In order to do so, we will follow Wirl and Feichtinger (2005) and Wirl (2004) to prove that our results hold under the most common and general functional specification to then be able to use explicit formulations for computational purpose without losing too much generality. As for the utility function we will assume a quadratic formulation:

$$U(c, w) = c\left(a - \frac{c}{2}\right) - \beta\frac{(w - w^H)^2}{2}$$

where $c^{sat} = a$.

For the law of motion of body weight we assume that individual body weight is composed by two main components: (1) fat mass (FM) and (2) fat-free mass (FFM). Each is supposed to represent a fraction μ and $1 - \mu$ with $\mu \in [0, 1]$, respectively, of total body weight w, such that

$$\begin{aligned} FM &= (1 - \mu)w \\ FFM &= \mu w \\ w &= FFM + FM = \mu w + (1 - \mu)w = w \end{aligned}$$

As medical literature suggests (Wang et al., 2000, Thomas et al., 2009, 2011), caloric expenditure for each of these components is different and assumed to be given by the following functions

$$\begin{aligned} CE_{FFM} &= \mu w \\ CE_{FM} &= -(1 - \mu)\frac{w^2}{2} \end{aligned}$$

such that: (1) fat-free mass has a linear, positive and monotonically increasing contribution to caloric expenditure and (2) fat mass has a quadratic and negative effect on caloric expenditure. The intuition is the following: initially there is a positive imbalance of fat-free mass over fat mass when, after $\bar{w}$ the (negative) contribution of fat mass tops the positive of the fat-free mass and caloric expenditure starts to decrease (see Figure 1). Overall caloric expenditure is computed by combining CE_{FFM} with CE_{FM}, such that

$$CE_{TOT} = \frac{1}{2}w(2\mu + (\mu - 1)w)$$

Which, given its quadratic characteristic, can be approximated with a general logistic function

$$CE_{TOT} = \frac{1}{2}w(2\mu + (\mu - 1)w) \approx w\left(\bar{w} - \frac{w}{2}\right) = g(w)$$

where $\bar{w}$ is assumed to represent the level of body weight above which there is a positive imbalance of fat mass over fat-free mass or, more intuitively, the starting point after which the individual starts to become overweight and consequently experiences a reduction in caloric expenditure.

[Figure 1 about here.]

To summarize utility and body weight law of motion are the following

$$U(c,w) = c\left(1-\alpha\frac{c}{2}\right) - \beta\frac{(w-w^H)^2}{2}$$
$$\dot{w} = c - g(w) = c - w(\bar{w} - \frac{w}{2}) \tag{15}$$

From this it is clear that we have deliberately chosen a simple standard specification for both equations in order to stress that the results are present even when using a very simple and compact functional form. Appendix A.5 provides a complete solution to this specific example. Fig.2 summaries the number and the dynamic orientation of the three steady states of the system in equation (15), which is characterized by two saddle point points, w_1^{ss} and w_3^{ss}, divided by an unstable node, w_2^{ss}. The figure also shows the saddle path leading to the two steady states and passing through the unstable node.

[Figure 2 about here.]

The multiplicity of steady states gives us a rationale for why, in general, dieting efforts are so difficult and why weight loss is not always guaranteed after a diet is finished. The reason behind this is the presence of a weight threshold above which any weight loss will be permanent. This means that if dieting does not enable the individual to surpass this threshold he or she will start to regain weight once the diet is finished and will then slowly converge back to the previous weight, nullifying all dieting efforts.

4 An application: Financial Incentives for weight loss

Due to unsuccessful dieting attempts and in order to facilitate weight-loss efforts, recent insights deriving from behavioral economics have been used to promote the reduction of excessive weight by motivating individuals to lose weight thanks to monetary and non-monetary incentives. Unfortunately, despite promising short-run results and the fact that weight loss was indeed achieved during experiments, substantial weight regain was found after the incentive was removed proving that, despite being sustained by incentives, individuals were still failing to lose weight permanently. We now investigate how short-run financial incentives interact in our model and if they might be able to produce permanent weight loss, given the established metabolic pattern introduce in Section 2.2.

4.1 Incentive for weight loss

In addition to the instantaneous utility $U(.)$ we also assume that there exists an incentive function for weight loss, which represents any type of incentive, internal and external, that the individual exploits in order to sustain weight loss. This function, denoted by $\alpha I(w^H, w(t))$, can be either positive or equal to zero, depending whether the individual is motivated or not to lose weight. $I(.)$ is a non-decreasing incentive function which achieves a (unique) minimum when $w = w^H$, where w^H is the target body weight and $\alpha \in [0, 1]$

is a parameter related to the individual's responsiveness to incentive. The idea behind the incentive function is that the farther the individual is from his or her target body weight, the higher the incentive to sustain dieting effort, which will gradually decrease as the individual converges towards w^H. The $I(.)$ function captures the idea that heavier individuals should be rewarded more than leaner individuals for weight loss of equal magnitude. In order to focus on healthy weight-loss incentives and to rule out any type of punishment or weight loss below an individual's healthy weight, w^H, it is necessary to impose the following set of restrictions on $I(.)$:

Assumption 2. *The incentive function has two properties:*

1. *For* $w < w^H$ $I(.) = 0$,
2. *For* $w > w^H$ $I(.) > 0$

The first assumption excludes any incentive (or punishment) for weight lower than w^H the second follows from the fact that the incentive function is an increasing function in body weight. If the first case does not hold we would model cases in which individuals are motivated to lose weight despite already being at their healthy body weight level[3].

By including the incentive for weight loss the system of differential equations characterizing the equilibrium is similar to the one we derived in Section 3.1

$$\begin{aligned} \dot{c} &= \frac{U_c(\rho + g_w) + U_w + \alpha I_w}{U_{cc}} \\ \dot{w} &= c - g(w) \end{aligned} \tag{16}$$

With the only difference being the presence of the incentive function in the $\dot{c}$ equation. For an internal steady state the following conditions must be satisfied

$$\dot{c} = 0 \qquad U_w + \alpha I_w = -U_c(\rho + g_w) \tag{17}$$

$$\dot{w} = 0 \qquad c = g(w) \tag{18}$$

In order to investigate what consequences had the introduction of the incentive on the steady states value of body weight and food consumption we will compute the change in the steady state body weight as a response to a change in the sensitivity to the incentive is given by the following expression (See Appendix A.4):

$$\frac{\partial w^{ss}}{\partial \alpha} = -\frac{|P|}{|J|} = -\frac{I_w/U_{cc}}{|J|} \tag{19}$$

Where $|J|$ is the determinant of the Jacobian associated with the system of differential equations (16) and

[3]Such situations can be linked to eating disorders such as bulimia and anorexia which are not explored in this paper.

$|P|$ is the determinant of the following matrix

$$P = \begin{bmatrix} \dfrac{\partial \dot{c}}{\partial \alpha} & \dfrac{\partial \dot{c}}{\partial c} \\ \\ \dfrac{\partial \dot{w}}{\partial \alpha} & \dfrac{\partial \dot{w}}{\partial c} \end{bmatrix}$$

The sign of equation (19) leads to the following proposition.

Proposition 4.1. *A sufficient condition for an incentive for weight loss to sustain an achievable steady state with a lower body weight is* $I_w > 0$.

Proof. In a steady state the determinant of the Jacobian is negative and thus the denominator of equation (19) is negative, so for $I_w > 0$ and $U_{cc} < 0$ the presence of the incentive ($\alpha > 0$) is associated with a stable decreases in body weight. ■

Corollary 4.1. *Incentive schemes unrelated to body weight (i.e.* $I_w = 0$*) are not able to lead to stable changes in body weight.*

Proof. For $I_w = 0$ the effect of an incentives on the equilibrium level of body weight is nil, since equation (19) is equal to zero. ■

As far as steady state food consumption is concerned we can derive a condition similar to (19) (See Appendix A.4)

$$\frac{\partial c^{ss}}{\partial \alpha} = -\frac{|K|}{|J|} = -\frac{g_w I_w / U_{cc}}{|J|} \tag{20}$$

$|K|$ is the determinant of

$$K = \begin{bmatrix} \dfrac{\partial \dot{c}}{\partial \alpha} & \dfrac{\partial \dot{c}}{\partial w} \\ \\ \dfrac{\partial \dot{w}}{\partial \alpha} & \dfrac{\partial \dot{w}}{\partial w} \end{bmatrix}$$

Such that

Proposition 4.2. *Steady state decreases in food consumption are achieved by* $I_w > 0$*, for overweight/obese individuals.*

Proof. By definition excessive body weight, or a positive imbalance between fat mass and fat free mass, is characterized by the condition $g_w < 0$, which will make equation (20) negative resulting in a decrease in steady state food consumption.. ■

The Jacobian matrix associated with the system (19)

$$J = \begin{pmatrix} \rho + g_w & \dfrac{U_{ww} + g_{ww}U_c + \alpha V_{ww}}{U_{cc}} \\ & \\ 1 & -g_w \end{pmatrix} \tag{21}$$

The trace is still positive and equals to $tr(J) = \rho > 0$, so the steady state can be at most a saddle point[4]. However the determinant is now

$$\det(J) = -g_w\left(\rho + g_w\right) - \frac{(U_{ww} + g_{ww}U_c + \alpha I_{ww})}{U_{cc}} \tag{22}$$

If we take a closer look at the determinant in (22) we arrive at the following proposition

Proposition 4.3. *"Progressivity" in the incentive scheme, $I_{ww} > 0$, or a high level of sensitivity in incentive schemes, is a necessary condition in order to sustain persistent weight loss.*

Proof. Rewriting the determinant equation (22) in terms of the instability condition (14) we have that

$$-g_w(\rho + g_w) > \frac{H_{ww}}{U_{cc}} + \alpha\frac{I_{ww}}{U_{cc}} \tag{23}$$

High level of either I_{ww} or α significantly reduce the conditions under which the inequality for instability is verified, since the second term in equation (23) is negative due to $U_{cc} < 0$. ■

Proposition 4.4. proves that incentive schemes which associate higher rewards for weight loss with a higher body weight are able to promote persistent weight loss. Intuitively, weight loss is possible whenever there is an incentive scheme which is more rewarding for weight loss at a higher level of body weight with respect to weight loss achieved when the body weight is relatively low ($I_{ww} > 0$), or when the incentive is so salient that it is able to provide active motivation for the dieter (high α). This progressivity characteristic that incentive schemes need in order to lead to permanent results might raise the controversy that agents should stay fat in order to gain higher rewards. However, one additional assumption has to be taken into consideration: *incentive schemes should be designed to reward weight loss at a higher level of body weight, not higher body weight per se'*[5].

The results on the canonical system (16) of introducing a quadratic incentive function, $\alpha(w - w^H)^2$ are summarized in Fig.3, from which the downward shift of the $\dot{c} = 0$ isocline reduces the number of steady states from three to just one. Pre-incentive individuals who were converging towards an excessive weight steady state are now attracted by the only steady state reachable, which is represented by a lower body weight and, hence, improved health. However, convergence to the lower steady state is not trivial and it can either be

[4] At least one eigenvalues must be positive

[5] This is usually done in experimental literature by providing the reward if and only if the individual were able to decreases his body weight.

temporary or permanent depending on whether the individual is able, once the incentive is removed, to reach the stable lower steady state, thus surpassing the intermediate unstable steady state.

[Figure 3 about here.]

[Figure 4 about here.]

Following Fig. 4 we can characterize the following outcomes

1) **Temporary Result:** The convergence towards the lowest steady state is such that the individual is unable to surpass the unstable steady state. Once the incentive is removed the individual will reconverge to the initial steady state.

2) **Permanent Result:** The convergence towards the lowest steady state is such that the individual is able to surpass the unstable steady state. Once the incentive is removed the individual will continue converging towards the healthier steady state.

In the first case the individual is not able to escape the "obesity trap" and will therefore be unable to permanently lose weight. In the second case, due to the presence of the incentive, the individual is able to avoid the "obesity trap" and permanently converge towards a lower level of body weight. The results from the model show an incentive scheme design which is able to produce permanent weight loss; however, it has to be pointed out that since the intervention is introduced temporarily its design is just a *necessary* condition for the convergence to an improved weight condition.

5 Conclusion

Given the increasing incidence of obesity, the promotion of efficient ways to invert this trend is a major health concern. A reduction in caloric intake or an increase in caloric expenditure via physical activity, are the two main strategies normally prescribed. Although there has been an increase in weight-loss attempts using both of these recommended strategies, their beneficial effects are yet to be seen. Due to unsuccessful attempts and in order to facilitate weight-loss efforts, recent insights from behavioral economics have been used to promote the reduction of excessive weight by motivating individuals to follow a type of behavior that they would not naturally follow by means of monetary and non-monetary incentives. Although the results seem to support the ability of the scheme to promote weight loss during the time of the experimentation, few of the experiments were able to sustain a long-lasting effect and almost all weight was regained post-experiment. This poses the question of understanding why weight loss is so difficult to achieve and what factors influence weight loss and how they can negate the effects of short-run incentives. Thus, in order to design more effective interventions for future implementation, a theoretical investigation into the effects that these incentives have on individuals' behavior and why weight loss maintenance after the cessation of the incentive is not sustained is needed. In order to do so we add to existing literature by proposing another, complementary, explanation for why dieting efforts are so difficult, even in the absence of rational addiction, time-inconsistent preferences or bounded rationality.

In this paper we have presented a theoretical model with an original assumption regarding individual metabolism, in order to investigate why weight-loss attempts are so difficult to achieve. The novelty of the model lies in the presence of a non-monotonic relationship between caloric expenditure and body weight, resulting in lower caloric expenditure for overweight/obese individuals. The theoretical analysis helps us to understand why losing weight is so difficult without relying on rational addiction, time-inconsistent preferences or bounded rationality and, most importantly, the conditions under which an incentive for weight loss should be designed such that the individual is able to permanently lose weight. Multiple steady states and a threshold characterizing a situation of "obesity traps", which the individual must try to surpass by means of dieting, are derived. Moreover, the presence of incentives is not always sufficient to lead to a permanent effect; such an effect is related to the design of the incentive scheme. A non-decreasing incentive scheme rewarding greater weight loss at a higher level of body weight (i.e. progressive) is able to sustain steady state decreases in body weight and food consumption, while a lump-sum incentive scheme is not.

References

Astrup, A., P. C. Gøtzsche, K. van de Werken, C. Ranneries, S. Toubro, A. Raben, and B. Buemann (1999). Meta-analysis of resting metabolic rate in formerly obese subjects. *The American journal of clinical nutrition 69*(6), 1117–1122.

Augurzky, B., T. K. Bauer, A. Reichert, C. Schmidt, and H. Tauchmann (2012). Does money burn fat? evidence from a randomized experiment.

Baradel, L., C. Gillespie, J. Kicklighter, M. Doucette, M. Penumetcha, H. Blanck, et al. (2009). Temporal changes in trying to lose weight and recommended weight-loss strategies among overweight and obese americans, 1996-2003. *Preventive medicine 49*(2-3), 158.

Bish, C., H. Blanck, M. Serdula, M. Marcus, H. Kohl, and L. Khan (2012). Diet and physical activity behaviors among americans trying to lose weight: 2000 behavioral risk factor surveillance system. *Obesity Research 13*(3), 596–607.

Catenacci, V., P. Maclean, L. Ogden, S. Polsky, H. Wyatt, and J. Hill (2011). *Chapter 46 in: John Cawley (editor), The Oxford Handbook of the Social Science of Obesity,*. Oxford University Press, USA.

Cawley, J. and J. A. Price (2013). A case study of a workplace wellness program that offers financial incentives for weight loss. *Journal of health economics 32*(5), 794–803.

Cunningham, J. J. (1991). Body composition as a determinant of energy expenditure: a synthetic review and a proposed general prediction equation. *The American journal of clinical nutrition 54*(6), 963–969.

Dodd, M. (2008). Obesity and time-inconsistent preferences. *Obesity Research & Clinical Practice 2*(2), 83–89.

Dragone, D. (2009). A rational eating model of binges, diets and obesity. *Journal of health economics 28*(4), 799–804.

Dragone, D., F. Manaresi, and L. Savorelli (2013). Obesity and smoking: can we catch two birds with one tax?

Dragone, D. and L. Savorelli (2011). Thinness and obesity: A model of food consumption, health concerns, and social pressure. *Journal of Health Economics*.

Ebbeling, C., J. Swain, H. Feldman, W. Wong, D. Hachey, E. Garcia-Lago, and D. Ludwig (2012). Effects of dietary composition on energy expenditure during weight-loss maintenance. *JAMA: The Journal of the American Medical Association 307*(24), 2627–2634.

Gale, S., V. Castracane, and C. Mantzoros (2004). Energy homeostasis, obesity and eating disorders: recent advances in endocrinology. *The Journal of nutrition 134*(2), 295–298.

Harris, J. A. and F. G. Benedict (1918). A biometric study of human basal metabolism. *Proceedings of the National Academy of Sciences of the United States of America 4*(12), 370.

Ikeda, S., M. Kang, and F. Ohtake (2010). Hyperbolic discounting, the sign effect, and the body mass index. *Journal of health economics 29*(2), 268–284.

John, L., G. Loewenstein, A. Troxel, L. Norton, J. Fassbender, and K. Volpp (2011). Financial incentives for extended weight loss: a randomized, controlled trial. *Journal of general internal medicine 26*(6), 621–626.

Johnstone, A. M., S. D. Murison, J. S. Duncan, K. A. Rance, and J. R. Speakman (2005). Factors influencing variation in basal metabolic rate include fat-free mass, fat mass, age, and circulating thyroxine but not sex, circulating leptin, or triiodothyronine. *The American journal of clinical nutrition 82*(5), 941–948.

Kahneman, D. and A. Tversky (1979). Prospect theory: An analysis of decision under risk. *Econometrica: Journal of the Econometric Society*, 263–291.

Katan, M. and D. Ludwig (2010). Extra calories cause weight gain - but how much? *JAMA: the journal of the American Medical Association 303*(1), 65–66.

Kruger, J., D. Galuska, M. Serdula, D. Jones, et al. (2004). Attempting to lose weight: specific practices among us adults. *American journal of preventive medicine 26*(5), 402.

Leibel, R. and J. Hirsch (1984). Diminished energy requirements in reduced-obese patients. *Metabolism 33*(2), 164–170.

Leibel, R., M. Rosenbaum, and J. Hirsch (1995). Changes in energy expenditure resulting from altered body weight. *New England Journal of Medicine 332*(10), 621–628.

Levy, A. (2002). Rational eating: can it lead to overweightness or underweightness? *Journal of Health Economics 21*(5), 887–899.

Loewenstein, G., T. Brennan, and K. Volpp (2007). Asymmetric paternalism to improve health behaviors. *JAMA: The Journal of the American Medical Association 298*(20), 2415–2417.

Loewenstein, G. and T. O'Donoghue (2004). Animal spirits: Affective and deliberative processes in economic behavior. *Available at SSRN 539843*.

Mifflin, M. D., S. St Jeor, L. A. Hill, B. J. Scott, S. A. Daugherty, and Y. Koh (1990). A new predictive equation for resting energy expenditure in healthy individuals. *The American journal of clinical nutrition 51*(2), 241–247.

Mozaffarian, D., T. Hao, E. B. Rimm, W. C. Willett, and F. B. Hu (2011). Changes in diet and lifestyle and long-term weight gain in women and men. *New England Journal of Medicine 364*(25), 2392–2404.

O'Donoghue, T. and M. Rabin (1999). Doing it now or later. *American Economic Review*, 103–124.

Ogden, C. and M. Carroll (2010). Prevalence of overweight, obesity, and extreme obesity among adults: United states, trends 1960–1962 through 2007–2008. *NCHS Health and Stats*.

Ozdenoren, E., S. Salant, and D. Silverman (2006). Willpower and the optimal control of visceral urges. Technical report, National Bureau of Economic Research.

Ruhm, C. (2012). Understanding overeating and obesity. *Journal of Health Economics*.

Thomas, D. M., A. Ciesla, J. A. Levine, J. G. Stevens, and C. K. Martin (2009). A mathematical model of weight change with adaptation. *Mathematical biosciences and engineering: MBE 6*(4), 873.

Thomas, D. M., C. K. Martin, S. Heymsfield, L. M. Redman, D. A. Schoeller, and J. A. Levine (2011). A simple model predicting individual weight change in humans. *Journal of biological dynamics 5*(6), 579–599.

Volpp, K., L. John, A. Troxel, L. Norton, J. Fassbender, and G. Loewenstein (2008). Financial incentive–based approaches for weight loss. *JAMA: the journal of the American Medical Association 300*(22), 2631–2637.

von Ruesten, A., A. Steffen, A. Floegel, G. Masala, A. Tjønneland, J. Halkjaer, D. Palli, N. Wareham, R. Loos, T. Sørensen, et al. (2011). Trend in obesity prevalence in european adult cohort populations during follow-up since 1996 and their predictions to 2015. *PLoS One 6*(11), e27455.

Wang, Z., S. Heshka, D. Gallagher, C. N. Boozer, D. P. Kotler, and S. B. Heymsfield (2000). Resting energy expenditure-fat-free mass relationship: new insights provided by body composition modeling. *American Journal of Physiology-Endocrinology And Metabolism 279*(3), E539–E545.

Wirl, F. (2004). Thresholds in concave renewable resource models. *Ecological economics 48*(2), 259–267.

Wirl, F. and G. Feichtinger (2005). History dependence in concave economies. *Journal of Economic Behavior & Organization 57*(4), 390–407.

Yaari, M. E. (1965). Uncertain lifetime, life insurance, and the theory of the consumer. *The Review of Economic Studies 32*(2), 137–150.

A Appendix A.

A.1 The intertemporal problem (3) - (6)

In this appendix we will prove that the intertemporal problem (3) - (6), with an infinite time horizon and constant discount rate, is an approximation of a similar model with stochastic time of death.

A.1.1 The Model with stochastic time of death

In order to take into account that individual's longevity is stochastic and impossible to known before time of death, we assume that the agent's life is finite but with an uncertain terminal date, T, (Yaari, 1965). $F(T)$ represents the probability o dying at time T and $f(T)$ is the associated density function. Given initial body weight w_0, the agent must choose the path of food consumption satisfying the following intertemporal problem

$$\max_{c(t)} \; E\Big[\int_0^T e^{-\tilde{\rho}t}[U(c(t),w(t))] + \alpha I(w^H, w(t))dt\Big] \tag{24}$$

$$\text{subject to } \dot{w}(t) = c(t) - g(w(t)) \tag{25}$$

$$w(t), c(t) \geq 0 \tag{26}$$

$$w(0) = w_0, \text{ given} \tag{27}$$

where $\tilde{\rho} > 0$ is the intertemporal discount rate representing the agent's impatience. The objective function (24) , differently from (3) represents the expected lifetime utility function of an agent with stochastic terminal time. Following Dragone et al. (2013), Yaari (1965) we can exploit a very useful result in order to prove the equivalency of the two. Equation (24) can be equivalently represented in terms of the following objective function

$$\int_0^T [1 - F(t)]e^{-\tilde{\rho}t}U(.) + \alpha I(.)dt$$

where $1-F(t)$ is the probability of living beyond t (Yaari, 1965, Levy, 2002). In order to prove the equivalency of (24) and (3) we will focus on the special case where $f(T) = \hat{\rho}e^{-\hat{\rho}t}$ so that the density function is exponential.

Under this assumption the expected intertemporal utility of the agent can be equivalently written as follows

$$E\left[\int_0^T e^{-\tilde{\rho}t}U(.) + \alpha I(.)dt\right] = \int_0^T e^{-\hat{\rho}t}e^{-\tilde{\rho}t}U(.) + \alpha I(.)dt = \int_0^T e^{-\rho t}U(.) + \alpha I(.)dt \tag{28}$$

where $\rho = \hat{\rho} + \tilde{\rho}$ and it is the overall discount rate depending on impatience ($\tilde{\rho}$) and on the hazard rate ($\hat{\rho}$). Equation (28) represents the discounted stream of utility of an infinitely-lived agent, however, an alternative interpretation is possible, whereby the objective function (28) represents the expected intertemporal utility of an agent with stochastic life and whose hazard rate is constant. Thus, providing a bridge between finite and infinite horizon models.

A.2 Concavity of the Hamiltonian function

In this appendix a formal analysis on the concavity of the Hamiltonian functions of the paper is provided. Assuming that the Utility function is separable in c and w such that $U_{cw} = 0$ we can write the Hessian of the Hamiltonian as

$$\tilde{H} = \begin{bmatrix} H(.)_{\lambda\lambda} & H(.)_{\lambda c} & H(.)_{\lambda w} \\ H(.)_{c\lambda} & H_{cc} & H(.)_{cw} \\ H(.)_{w\lambda} & H(.)_{wc} & H(.)_{ww} \end{bmatrix} = \begin{bmatrix} 0 & 1 & -g_w \\ 1 & H_{cc} & 0 \\ -g_w & 0 & H_{ww} \end{bmatrix}$$

with $H_{cc} = U_{cc}$ and $H_{ww} = U_{ww} - \lambda g_{ww}$.

If the last n minors alternate in sign provided that the det $\tilde{H}$ itself has the sign of $(-1)^n$, where n is the number of variable in the utility function, then the point is a local maximum i.e. the Hessian is concave. On the other hand if the largest n minors all have positive sign and that sign is $(-1)^m$ where m is the number of multipliers in the problem the point is a local minimum. Thus, concavity requires that the principal largest leading minor of $\tilde{H}$ to be mixed in signs:

$$|\tilde{H}_2| = \begin{vmatrix} 0 & 1 \\ 1 & H_{cc} \end{vmatrix} = -1$$

$$|\tilde{H}_3| = \begin{vmatrix} 0 & 1 & -g_w \\ 1 & H_{cc} & 0 \\ -g_w & 0 & H_{ww} \end{vmatrix} = -g_w^2 H_{cc} - H_{ww} > 0$$

Given the assumption on $U(.)$ and $g(.)$ joint concavity is satisfied as long as $g_w^2 U_{cc} + U_{ww} + \alpha I_{ww} > \lambda g_{ww}$.

A.3 The Derivation of Condition (14)

The Jacobian associated with the intertemporal problem (3)-(6) is

$$J = \begin{pmatrix} \rho + g_w & \dfrac{U_{ww} + g_{ww} U_c}{U_{cc}} \\ \\ 1 & -g_w \end{pmatrix} \tag{29}$$

Given that the trace is positive and equals to $tr(J) = \rho > 0$, the steady state can be at most a saddle point[6]. The determinant is harder to sign

$$\det(J) = -g_w\,(\rho + g_w) - \frac{(U_{ww} + g_{ww} U_c)}{U_{cc}} \tag{30}$$

A negative determinant, $\det(J) < 0$ implies that one of the eigenvalues is negative, characterizing a saddle point[7] A positive determinant, $\det(J) > 0$, is then a necessary and sufficient conditions to exclude a saddle (i.e. conditional) stability, both eigenvalues are positive so that the associated steady state is an unstable node. To characterize the cases in which such instability is achieved we can rewrite the determinant as

$$\det(J) = -g_w\,(\rho + g_w) - \frac{H_{ww}}{U_{cc}} \tag{31}$$

given the assumptions on $U(.)$ and $g(.)$ H is concave and the last term of equation (26) is negative meaning that a positive determinant is only possible when, $g_w < 0$, providing the following sufficient condition for instability

$$-g_w(\rho + g_w) > \frac{H_{ww}}{U_{cc}} \tag{32}$$

A.4 Comparative Statics

In order to compute the effect of a change in a parameter x on steady-state levels of body weight (w^{ss}) and food consumption (c^{ss}), we apply the Cramer's rule to the following system (all functions are evaluated

[6] At least one of the two eigenvalues of J

$$e_{1,2} = \frac{1}{2}\left[tr(J) \pm \sqrt{tr(J)^2 - 4\det(J)}\right]$$

is positive (or has positive real parts)

[7] Which even if unstable has a stable manifold (of dimension one) determining optimal policy. Given the positive trace with a little but of abuse of notation I will define *'stable'* such outcomes, since it is the maximum achievable stability given by the canonical equation in (8).

at the steady state)

$$\begin{pmatrix} \dfrac{\partial \dot{c}}{\partial c} & \dfrac{\partial \dot{c}}{\partial w} \\ \\ \dfrac{\partial \dot{w}}{\partial c} & \dfrac{\partial \dot{w}}{\partial w} \end{pmatrix} \begin{pmatrix} \dfrac{\partial c^{ss}}{\partial x} \\ \\ \dfrac{\partial w^{ss}}{\partial x} \end{pmatrix} = \begin{pmatrix} -\dfrac{\partial \dot{c}}{\partial x} \\ \\ -\dfrac{\partial \dot{w}}{\partial x} \end{pmatrix}$$

where the first term is the Jacobian matrix ($|J|$) of the system of differential equations (16) evaluated at the steady state. Applying Cramer's rule for a change in α we have

$$\begin{pmatrix} \dfrac{\partial c^{ss}}{\partial x} \\ \\ \dfrac{\partial w^{ss}}{\partial x} \end{pmatrix} = \begin{pmatrix} -\dfrac{|K|}{|J|} \\ \\ -\dfrac{|P|}{|J|} \end{pmatrix}$$

where $|P|$ and $|K|$ are the determinant of the following matrices

$$K = \begin{bmatrix} \dfrac{\partial \dot{c}}{\partial \alpha} & \dfrac{\partial \dot{c}}{\partial w} \\ \\ \dfrac{\partial \dot{w}}{\partial \alpha} & \dfrac{\partial \dot{w}}{\partial w} \end{bmatrix} \qquad P = \begin{bmatrix} \dfrac{\partial \dot{c}}{\partial \alpha} & \dfrac{\partial \dot{c}}{\partial c} \\ \\ \dfrac{\partial \dot{w}}{\partial \alpha} & \dfrac{\partial \dot{w}}{\partial c} \end{bmatrix}$$

A.5 The example with explicit functions

Threshold and history dependence are possible in a perfectly governed economy (i.e. strictly concave). The first term of the utility function specifies that $c^{sat} = 1/\alpha$ where α captures individual characteristics (e.g. income., gender, age) which affects food satiation; the second term represents the consequence of begin either overweight ($w > w^H$) or underweight ($w < w^H$) with $\beta > 0$ representing the scaling parameter of the (quadratic) appreciation of an healthy body weight. (12) is determining the following canonical equations

$$\begin{aligned} \dot{c} &= \frac{(\alpha c - 1)(\rho - (w - \bar{w}))}{\alpha} + \beta \frac{(w - w^H)}{\alpha} \\ \dot{w} &= c - w\left(\bar{w} - \frac{w}{2}\right) \end{aligned} \tag{33}$$

From which we can derive the $\dot{c} = 0$-isocline and the $\dot{w} = 0$-isocline, which are respectively

$$\begin{aligned} \dot{c} = 0 \qquad & c = -\beta \frac{(w - w^H)}{a[\rho - (w - \bar{w})]} + \frac{1}{a} \\ \dot{w} = 0 \qquad & c = w\left(\bar{w} - \frac{w}{2}\right) = g(w) \end{aligned} \tag{34}$$

From which we can verify the arrows of motion of Fig.2. For the $\dot{w} = 0$-isocline we have $\dot{w} < 0$ for $c < g(w)$ and $\dot{w} > 0$ otherwise; while for the $\dot{c} = 0$-isocline we have $\dot{c} < 0$ for $c > -\beta \frac{(w - w^H)}{a[\rho - (w - \bar{w})]} + \frac{1}{a}$ and $\dot{c} > 0$ otherwise.

The system (34) has up to three steady states since substituting the stationary solution for $\dot{c} = 0$ into $\dot{w} = 0$ results in a cubic equation in w. The Jacobian as well as the determinant simplify to

$$j = \begin{pmatrix} \rho + g_w & \dfrac{\beta - (\alpha c - 1)}{\alpha} \\ & \\ 1 & -g_w \end{pmatrix} \tag{35}$$

$$\det(j) = -g_w(\rho + g_w) + \frac{(\alpha c - 1) - \beta}{\alpha} \frac{\beta}{\alpha} - cg''(w) + \frac{g''(w)}{\alpha} - \rho g'(w) - g'(w)^2 \tag{36}$$

The $\dot{w} = 0$, $c = w(\bar{w} - w/2)$ is determined by the metabolic level of caloric consumption, $c = g(w)$ with a maximum in $\bar{w}$. The $\dot{c} = 0$ hits this function in three points corresponding to three steady state characterized by an incremental level of steady state body weight.

Figure 1: The law of motion of body weight for $\bar{w} = 0,82$.

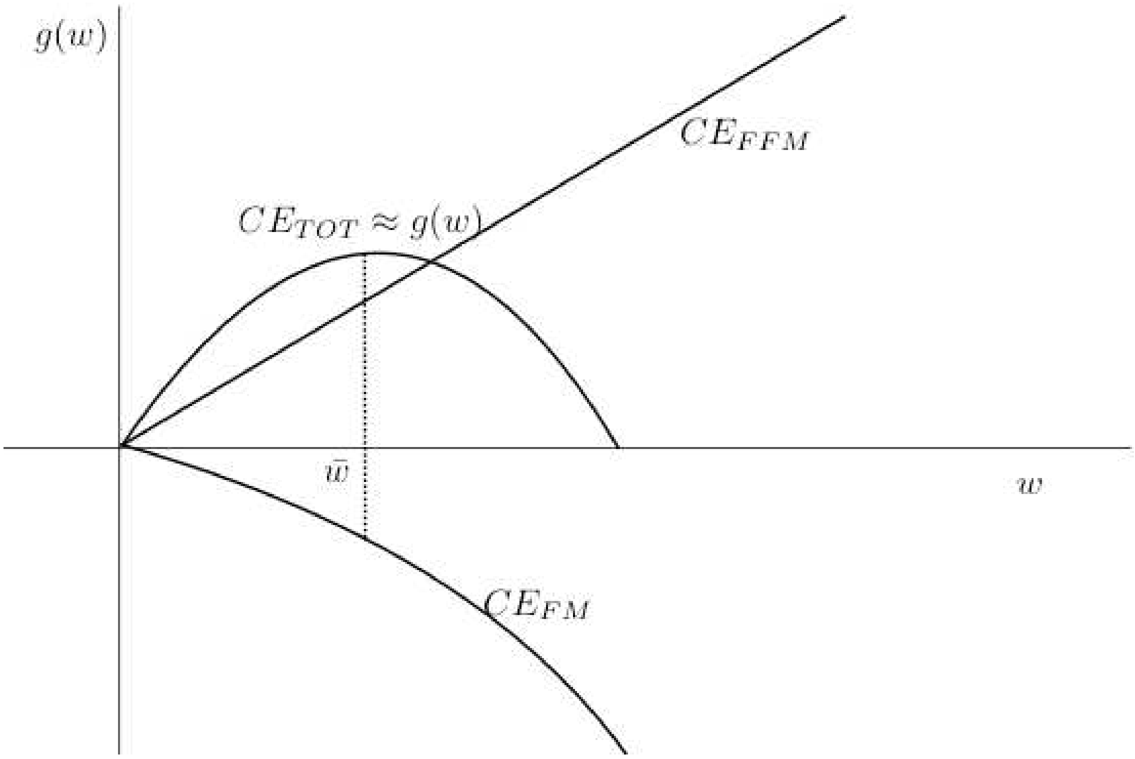

Figure 2: Phase diagram of (15) for the parameters values: $a = 3.9$, $\beta = 0.05$, $\bar{w} = 0.82$, $w^H = 0.44$ and $\rho = 0.8$

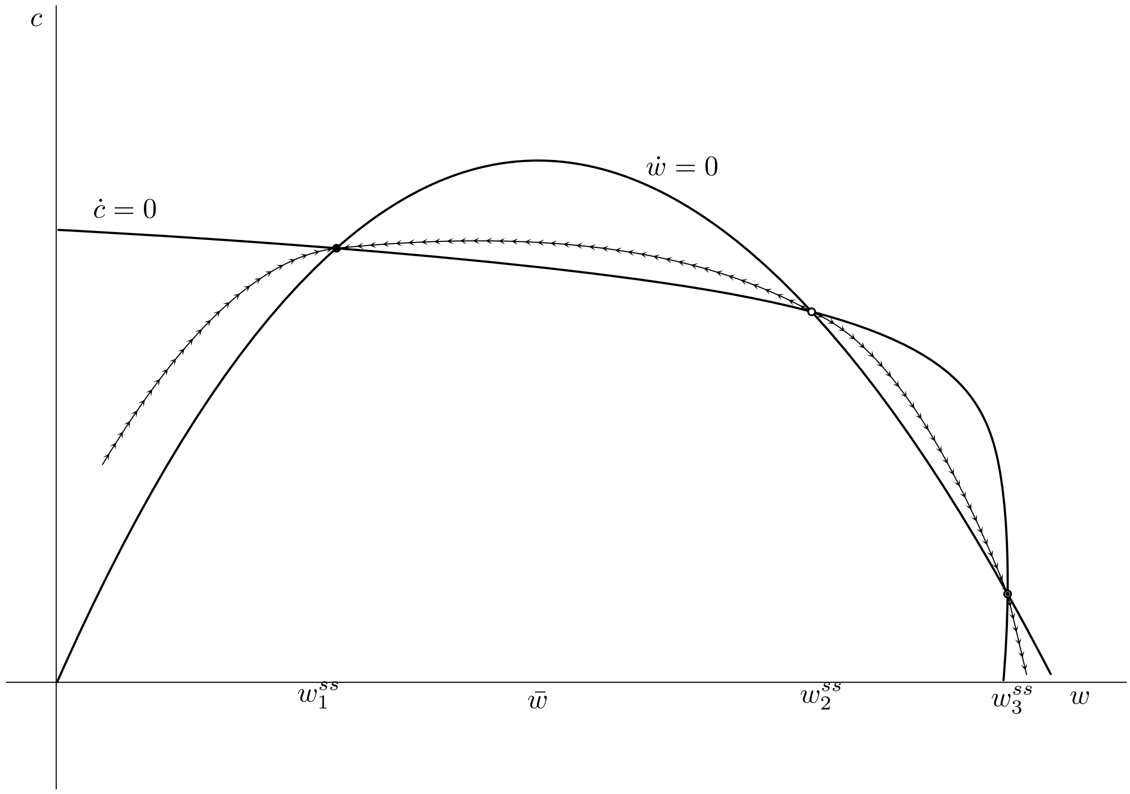

Figure 3: Phase diagram of (16) after the introduction of the incentive for the parameters values: $a = 3.9$, $\beta = 0.05$, $\bar{w} = 0.82$, $w^H = 0.44$, $\alpha = 0.4$ and $\rho = 0.8$

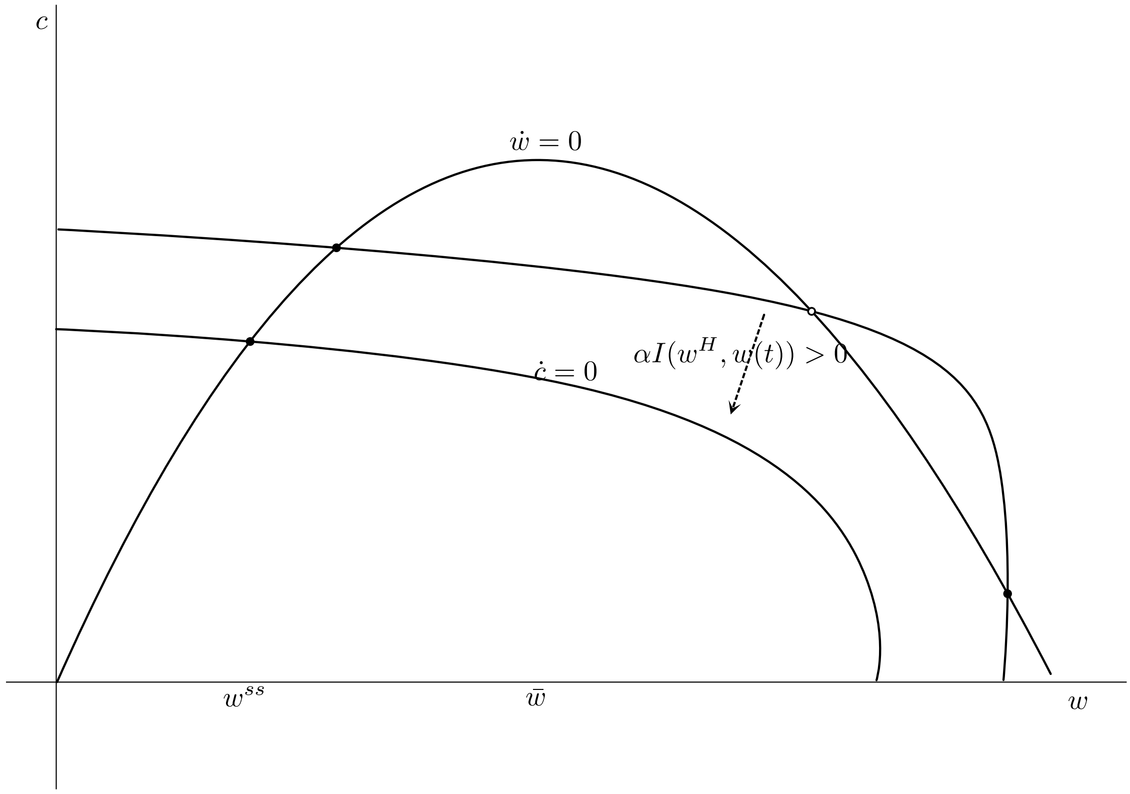

Figure 4: Comparative Statics, pre- and post-incentive.