NBER WORKING PAPER SERIES

HARM REDUCTION: WHEN DOES IT IMPROVE HEALTH, AND WHEN DOES IT BACKFIRE?

John Cawley
Davide Dragone

Working Paper 30926
http://www.nber.org/papers/w30926

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
February 2023

For helpful comments, we thank Catarina Goulao, Catherine Maclean, Jesse Matheson, Michael Pesko, Luca Piccoli, Pedro Pita Barros, Holger Strulik, and participants in the following seminars and conferences: the University of Duisburg-Essen seminar, the Toulouse School of Economics seminar, the Essen Health Conference 2022, the EuHEA 2022 seminar, and the ASHEcon 2022 conference. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Harm Reduction: When Does It Improve Health, and When Does it Backfire?
John Cawley and Davide Dragone
NBER Working Paper No. 30926
February 2023
JEL No. D11,I12,I18

ABSTRACT

A subset of harm reduction strategies encourages individuals to switch from a harmful addictive good to a less harmful addictive good; examples include e-cigarettes (substitutes for combustible cigarettes) and methadone and buprenorphine (substitutes for opioids). Such harm reduction methods have proven to be controversial. Advocates argue that people struggling with addiction benefit because they can switch to a less harmful substance, but opponents argue that this could encourage abstainers to begin using the harm reduction method or even the original addictive good. This paper builds on theories of addiction to model the introduction of a harm reduction method, and it demonstrates the conditions under which each side is correct. The three key factors determining whether the introduction of a harm reduction method reduces or worsens health harms are: 1) the enjoyableness of the harm reduction method, 2) the addictiveness of the harm reduction method, and 3) the substitutability of the harm reduction method with the original addictive good. Knowledge of these conditions can help inform regulation of harm reduction methods.

John Cawley
2312 MVR Hall
Jeb E. Brooks School of Public Policy
and Department of Economics
Cornell University
Ithaca, NY 14853
and NBER
JHC38@cornell.edu

Davide Dragone
Dipartimento di Scienze Economiche
Università di Bologna
Piazza Scaravilli 2
40126, Bologna
Italy
davide.dragone@unibo.it

1 Introduction

Due to the substantial morbidity and mortality attributable to cigarette smoking, alcohol abuse, and drug abuse, nations worldwide have sought methods of reducing the health consequences of such addictive behaviors.[1] One controversial approach is harm reduction, which may be best understood in contrast to the zero-tolerance approach, which argues that society's goal should be to completely eliminate all addictive behaviors. In contrast, harm reduction de-emphasizes the goal of eliminating addictive behaviors and instead focuses on reducing the health harms associated with such behaviors (Erickson, 1995; Single, 1995; Harm Reduction International, 2022).

Harm reduction methods have proven to be controversial, with advocates touting the potential health benefits, and opponents arguing that harm reduction methods that are themselves addictive could prevent current users from quitting and may lead current abstainers to begin using the harm reduction method, or even the original addictive good.

This paper builds on theories of addiction to model how the introduction of a harm reduction method influences addictive consumption and health, and it derives the conditions under which each side is correct. Specifically, we show the conditions under which the harm reduction method leads people to quit using the original, more harmful, addictive good, and the conditions under which the harm reduction method leads previous abstainers to begin using not just the harm reduction method but also the original, more harmful, addictive good. We also demonstrate the conditions under which the introduction of a harm reduction method does in fact reduce health harms and the conditions under which it backfires and health harms increase.

The term harm reduction has been applied to a wide range of approaches, including syringe exchange programs, supervised injection facilities, legalized prostitution, condom distribution, Naloxone access laws, and Good Samaritan Laws (Stancliff et al., 2012; Rees et al., 2019; Packham, 2022; Doleac and Mukherjee, 2022; SAMHSA, 2023; U.S. Department of Health and Human Services, 2023). For the purposes of this paper, we focus on the subset of harm reduction methods that have the following properties. First, the harm reduction method is a substitute for an addictive good. The mechanism for this may be that the harm reduction method binds to the same receptors in the brain as the original addictive good; in such a case the harm reduction method is known as agonist therapy. Full agonists provide roughly the same euphoric effects as the original addictive good, whereas partial agonists are less euphoric (but still reduce feelings of withdrawal). Methadone (a full agonist) is to some extent a substitute for heroin, and buprenorphine (a partial agonist) is to some extent a substitute for opioid pain

[1] The World Health Organization estimates that, worldwide, there are 8 million deaths annually from smoking, 3.3 million deaths annually from alcohol abuse, and 500,000 deaths annually from drug overdose (WHO, 2022a,b,c).

relievers. In both cases, the harm reduction method binds to similar opioid receptors as the original addictive good, leading to the release of similar neurotransmitters in the brain. In this sense, one can interpret the demand for both the original addictive good and the harm reduction method as a derived demand (Marshall, 1890; Lillard, 2020), i.e. derived from the demand for elevated levels of neurotransmitters associated with feelings of pleasure and reward, such as dopamine. Certain harm reduction methods are substitutes for the original addictive good in the production of the release of those neurotransmitters.

The second characteristic is that the harm reduction method is less harmful to health than the original addictive good. Obviously, if it was more harmful then it would not be a harm reduction method. Note that the harm reduction method may still be harmful to health, just not as harmful as the original product.

There are numerous examples of harm reduction methods that satisfy these two criteria.[2] When the concern is the smoking of combustible cigarettes, harm reduction methods include electronic nicotine delivery systems or ENDS (commonly called e-cigarettes, the use of which is called vaping), and nicotine replacement therapy or NRT (which includes nicotine gum, patches, and lozenges). ENDS and NRT are harm reduction methods for combustible cigarettes because they are substitutes (they bind with the nicotine receptors in the brain and thus can reduce withdrawal from combustible cigarettes) and are believed to be less carcinogenic and toxic than cigarette smoke (although not likely completely safe). There is more concern about the potential harm of ENDS (such as vaping) than of NRT (such as nicotine gum) because the former but not the latter involves inhaling chemicals (WHO, 2021). An additional harm reduction method to combat smoking that is popular in Norway and Sweden is snus, an oral tobacco product (Clarke et al., 2019).

When the concern is opioid addiction, relevant harm reduction methods include opioid agonist therapy, which uses methadone as a substitute for heroin or buprenorphine as a substitute for opioid pain relievers. These are sometimes called Opioid Agonist Therapy (OAT), and are a subset of Medications for Opioid Use Disorder (MOUD). When the concern is the smoking of combustible marijuana, harm reduction methods include edibles containing THC (a cannabinoid that provides a high), which allows the user to consume marijuana without inhaling toxic and carcinogenic smoke. Table 1 provides examples of harm reduction for cigarettes, heroin, opioid pain relievers, and marijuana, explaining why the harm reduction methods are

[2] There is some debate over the scope of activities and products included under the term "harm reduction." Historically it was sometimes used for activities that were not medication therapies, although that is changing (Stancliff et al., 2012; Krawczyk et al., 2022). The National Harm Reduction Coalition interprets medication for opioid use disorder (MOUD), such as methadone and buprenorphine, as harm reduction (Singer, 2018; National Harm Reduction Coalition, 2023). Likewise, state health departments also classify MOUD as a harm reduction strategy (Washington State Health Care Authority, 2023).

Original Addictive Substance	Harm Reduction Method	Why they are Substitutes	Why Harm (Potentially) Reduced	Moral Hazard would take the form of ...
Combustible Cigarettes	Nicotine replacement therapy or NRT (e.g. nicotine gum, patches, and lozenges); snus;	All bind with and activate nicotine receptors	NRT and snus do not involve smoking or inhaling anything, so are less carcinogenic	Current smokers may switch to e-cigarettes rather than quit altogether
	Electronic nicotine delivery systems or ENDS (aka e-cigarettes)		ENDS are more controversial; still involve inhaling chemicals but are believed to be less carcinogenic than regular smoking	Non-smokers may begin to vape; children a particular concern
	Chantix (varenicline)	Chantix is a partial agonist		Worst case: previous nonsmokers initiate smoking because of gateway effect of vaping
Heroin	Methadone (Opioid agonist therapy or OAT)	Both bind with and activate opioid receptors	Overdose is less likely on methadone than heroin, but still possible	Heroin addicts may switch to methadone rather than quit altogether
		Methadone activates receptors more slowly and less strongly so is less euphoric; reduces withdrawal but provides less high.M ethadone is still considered a full agonist	Consumed orally rather than injected; less risk of HIV transmission	
Prescription opioid pain relievers	Buprenorphine (Opioid agonist therapy or OAT)	Both bind with and activate opioid receptors	Because buprenorphine is a partial agonist, overdose is less likely	Those addicted to opioid pain relievers may switch to buprenorphine rather than quit altogether
		Buprenorphine activates receptors more slowly and less strongly (it is a partial agonist) so is less euphoric; reduces withdrawal but provides less high. Buprenorphine is a partial agonist		Buprenorphine pills may be diverted for illicit use; may attract new users of opioids
Combustible marijuana	Edibles with THC	Both bind with and activate cannabinoid receptors	Consumed orally rather than inhaled, so are less carcinogenic	Current marijuana smokers may switch to edibles rather than quit altogether
			Concern about edibles being accidentally consumed, particularly by children	People who would not smoke marijuana may begin consuming edibles

Table 1: Examples of Harm Reduction Methods

substitutes for the original addictive substance, and how they may reduce harm.

Although we focus on a subset of harm reduction methods - those that are themselves somewhat addictive - in many ways these are the most interesting ones to examine, as they may be the most likely to involve unintended consequences and to be the most controversial. Our model, however, can be easily extended to account for harm reduction methods that are non addictive and non harmful.

The best case scenario with such harm reduction methods is that they could lead to the total elimination of one's addictive stock. Specifically, users of the original addictive good could transition to the harm reduction method and then eventually quit even that. In essence, the existence of the harm reduction method may offer an additional option to users of the original addictive good who cannot quit "cold turkey" - they can instead use the harm reduction method to gradually wean themselves from their addiction. Even if those who quit the original addictive good end up staying on the harm reduction method indefinitely, that may still be an improvement because the harm reduction method is both less addictive and less harmful to health than the original addictive good. Organizations such as Harm Reduction International, the National Harm Reduction Coalition and the Drug Policy Alliance advocate in favor of liberalizing access to harm reduction methods.

Opponents have the following concerns: 1) harm reduction methods may decrease quitting of the original addictive good. The rationale is: the very harmfulness of the original addictive good may motivate users to quit. However, if there is a substitute product that is less harmful then users may switch to that substitute rather than quit altogether. 2) Harm reduction methods may encourage new people to become addicted. Some people may be abstaining precisely because the original substance is harmful to health; introducing a product that is less harmful may encourage some of those who previously abstained to begin using the new product. Even worse, some of those previous abstainers who begin using the harm reduction method may eventually transition to the original, more harmful, substance in search of a bigger high.[3] The arguments of both sides relate to moral hazard. The health harms are part of the total (shadow) price of addiction, and if one makes addiction less harmful then it lowers the shadow price of addiction and people may demand more of it. This is another version of the argument that innovations in car safety, by making crashes safer, lead people to drive in riskier ways (Peltzman, 1975). The final column of Table 1 provides examples of moral hazard for each of the examples of harm reduction.

The ambivalence about harm reduction affects many aspects of regulation and policy. For example, consider the case of buprenorphine, a harm reduction method for addiction to pre-

[3]See, e.g., Campbell (2009) for a controversy concerning the UN declaration of intent toward harm reduction policies, Satel (2019) in the context of combustible cigarettes and ENDS, and Vestal (2016) on methadone.

scription opioid pain relievers. Buprenorphine is actually more tightly regulated than the opioid pain relievers that have contributed heavily to the fatal drug overdose epidemic in the U.S. (Powell et al., 2020; Maclean et al., 2021; Alpert et al., 2022). Physicians are able to prescribe opioid pain relievers to any number of patients, but in order to prescribe buprenorphine, physicians must undertake 8 hours of training and obtain a waiver from the U.S. Drug Enforcement Agency prior to ever prescribing, and that only enables them to prescribe it to a limited number of patients (Waters, 2019). As a result, 40% of U.S. counties have no waivered physicians who can prescribe buprenorphine, and another 24% of counties have insufficient prescribing capacity (Grimm, 2020). Another policy that restricts access is that numerous states' Medicaid programs require prior authorization before they will cover the cost of buprenorphine (Weber and Gupta, 2019).

Likewise, electronic nicotine delivery systems (ENDS) are sometimes more tightly regulated than combustible cigarettes. The World Health Organization reports that 32 nations (including Australia, Brazil, India, Japan, and Mexico) have banned e-cigarettes entirely (WHO, 2022d). In countries where ENDS are legal, they range from completely unregulated to regulated as pharmaceutical products (WHO, 2022d). In 2022, the U.S. Food & Drug Administration (FDA) issued a marketing denial order to Juul, which banned them from selling any of their ENDS products in the United States, despite the fact that cigarettes remain legal to sell (U.S. Food and Drug Administration, 2022). There is also debate about the optimal taxation of e-cigarettes. In the U.S., 21 states do not tax e-cigarettes at all. Among those that do tax e-cigarettes, the structure and amount of those taxes vary considerably. Among states that tax e-cigarettes on the basis of their wholesale price, the tax rates range from 8% in New Hampshire to 95% in Minnesota (IGEN, 2022).

This paper has three purposes. First, we present a model of consumption of an addictive good, both before and after the introduction of a method of harm reduction. Second, we demonstrate the conditions under which the introduction of a method of harm reduction has the following consequences: a) it increases or decreases health harms; b) it leads previous users to quit the original addictive good; c) it leads previous abstainers to begin using the harm reduction method; and d) it leads previous abstainers to begin using the original addictive good. We show that the conditions for these different outcomes depend on three key factors: 1) the enjoyableness of the harm reduction method; 2) the addictiveness of the harm reduction method; and 3) the substitutability of the harm reduction method with the original addictive good.

This paper relates to several literatures. First, by deriving the conditions under which harm reduction leads to increased consumption by users or increased initiation by abstainers, we contribute to the economic literature on moral hazard in health behaviors (e.g. Bhattacharya et al., 2011; Margolis et al., 2014; Simon et al., 2017; Cotti et al., 2019; Dave et al., 2019; Frio and França, 2021; Doleac and Mukherjee, 2022; Agrawal et al., 2022). Second, we contribute to the economic literature on the specific harm reduction methods of ENDS or e-cigarettes and whether they are a substitute for combustible cigarettes (e.g. Lakdawalla et al., 2006; Friedman, 2015; Abouk and Adams, 2017; Abouk et al., 2019; Marti et al., 2019; Pesko and Currie, 2019; Cotti et al., 2020; Pesko et al., 2020; Abouk et al., 2021; Pesko and Warman, 2022; Allcott and Rafkin, 2022). Third, we contribute to the literature regarding the harm reduction methods methadone and buprenorphine and their regulation (e.g. Bishai et al., 2008; Abouk et al., 2019; Rees et al., 2019; Maclean et al., 2021; Barrette et al., 2021; Doleac and Mukherjee, 2022; Allen et al., 2022).

2 A Model of Harm Reduction and Addictive Consumption

Building on the two-stock model of rational addiction developed by Becker and Murphy (1988) and extended by Dockner and Feichtinger (1993), we first consider the conditions for a person to become a consumer of an addictive good. Then we introduce a harm reduction method, and show how this affects health and the consumption of the original addictive good.

2.1 Addictive Consumption in the Absence of a Harm Reduction Method

Define c as the consumption of an addictive good (e.g. combustible cigarettes). Consumption of this good contributes to two stocks. The first one is an addictive stock A, a measure of the past consumption experiences with the addictive good, which evolves over time according to

$$\dot{A}(t) = c(t) - \delta_A A(t) \tag{1}$$

where $\delta_A > 0$ is the depreciation rate of the addictive stock. The second stock is H, which describes the negative health consequences of addictive consumption, i.e. health harm. Stock H increases with both current and past consumption of the addictive good according to Allcott and Rafkin (2022)

$$\dot{H}(t) = c(t) + \omega A(t) - \delta_H H(t) \tag{2}$$

where $\omega > 0$ is the marginal contribution of addiction to health harms and the depreciation rate of health harms is $\delta_H > 0$. The dependence of health harm on A implies that addiction is bad for one's health, which adds to the health harms associated with current consumption.

Consider the following utility function

$$\mathcal{U}(c, q; A, H) = \left(u_c + u_{cA}A + \frac{u_{cc}}{2}c\right)c + \left(u_A + \frac{u_{AA}}{2}A\right)A + \left(u_H + \frac{u_{HH}}{2}H + u_{AH}A\right)H + q \quad (3)$$

where u_A, u_H and $u_{cc}, u_{AA}, u_{HH}, u_{AH}$ are negative, and variable q represents a numeraire composite good.[4] We call parameter $u_c > 0$ the enjoyableness of the addictive good. It corresponds to the marginal utility of consumption absent any current and previous consumption (i.e. when $c = A = 0$). In other words, it is the marginal utility of consumption confronting someone who has until that period abstained from c – i.e. the marginal utility they would experience from their first use of c. A defining assumption of the rational addiction model is that c is addictive, which is formalized assuming that the larger the stock of A the larger the marginal utility of c, i.e. $u_{cA} > 0$. This is the nature of addiction: the more addictive good one has consumed in the recent past, the greater one's marginal utility of consumption (or, put another way, the greater the withdrawal one experiences from not consuming it. In the rational addiction literature, this is referred to as adjacent complementarity or reinforcement(see,e.g., Becker and Murphy, 1988; Becker et al., 1991). For consistency, we refer to $u_{cA} > 0$ as the degree of addictiveness of the good.

Given a discount rate ρ, the intertemporal rational addiction problem of the agent is

$$\max \int_0^\infty e^{-\rho t}\mathcal{U}(c(t), q(t), A(t), H(t))\,\mathrm{d}t \quad (4)$$

subject to the law of motions 1 and 2, and the budget constraint

$$M(t) = q(t) + p_c c(t) \quad (5)$$

where M is income and p_c describes the monetary price, possibly including taxes, of the addictive good.[5]

To determine the optimal quantity of addictive consumption, we apply the Pontryagin's maximum principle and obtain the optimal trajectory of consumption, addiction and health harm toward the steady state. As shown in Appendix A.1, the steady state level of consumption

[4]The linear-quadratic specification, which can be considered to be a second-order approximation of a more general utility function, is common in the rational addiction literature because it allows for a closed-form solution of the optimal trajectory and the steady state (see, e.g., Becker and Murphy, 1988; Chaloupka, 1991; Gruber and Köszegi, 2001; Dragone and Raggi, 2018, 2021; Piccoli and Tiezzi, 2021). The quasi-linear specification rules out income effects to better focus on substitution patterns (Dragone and Vanin, 2022), as we do on this paper.

[5]To ease the exposition, here we abstract from considerations about the time cost of obtaining the good due to, e.g. the time spent on consumption, the need for medical prescriptions (if consumption is regulated), or expected sanctions and the risk of accessing the black market (if consumption of the good is illegal).

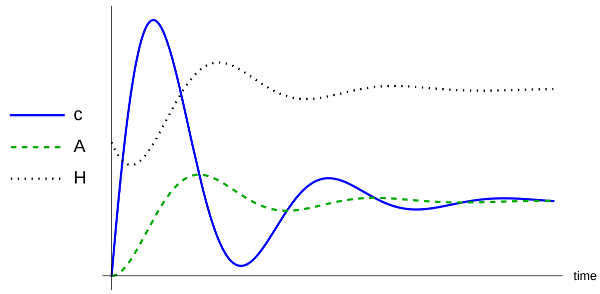

Figure 1: Initiation and addictive consumption when one addictive good is available. Illustration of possible time-paths for c, A and H.

of the addictive good is

$$c^{ss} = \alpha \left(u_c - p_c - \pi_c\right) \tag{6}$$

with $\alpha > 0$. Hence, steady-state consumption of the addictive good will be positive if $u_c > p_c + \pi_c$. The right-hand side represents the (shadow) price of the addictive good as the sum of its monetary price p_c and of $\pi_c > 0$, a term that describes the decrease in future utility resulting from consumption raising addiction and health harms. According to eq. 6, the addictive good is consumed only if the benefits exceed the costs – i.e. if the instantaneous marginal utility of consumption from first time use exceeds the sum of the monetary costs of purchasing the addictive good and the future consequences of consumption.

Along the path toward the steady state, consumption depends on current addiction and health harm according to

$$\hat{c}\left(A, H\right) = a_c c^{ss} + a_A A + a_H H \tag{7}$$

Since $a_c > 0$, $\hat{c}$ will be higher, for a given stock of addiction A and health harm H, the higher the steady-state consumption c^{ss}. Note that, despite being a linear equation in addiction and health harm, the time trajectory for consumption allows for oscillations as a possible consequence of the underlying dynamics of the state variables (see Figure 1 for an illustration). This implies that, even if consumption is zero in the long-run, it does not preclude the possibility that the person had earlier experimented with the addictive good.

2.2 Addictive Consumption in the Presence of a Harm Reduction Method

We now introduce a harm reduction method v, which is an addictive and less-harmful substitute for the addictive good c. For example, if c describes cigarettes then v can describe vaping of e-cigarettes. The harm reduction method v adds to both the addictive stock A and the stock of health harms H

$$\dot{A}(t) = c(t) + \varepsilon_A v(t) - \delta_A A(t) \tag{8}$$

$$\dot{H}(t) = c(t) + \varepsilon_H v(t) + \omega A(t) - \delta_H H(t) \tag{9}$$

The harm reduction method v is assumed to contribute less to the addictive stock than the original good. Given that the contribution of the original addictive good c to the stocks A and H is normalized to one, then $\varepsilon_A \in [0,1)$. This seems particularly likely if the harm reduction method is a partial rather than full agonist. Analogously, the debate on the possible introduction of harm reduction methods is typically focused on methods that are considered to be less harmful to health than the original addictive good, hence $\varepsilon_H \in [0,1)$.[6]

When the harm reduction method is available, the agent's instantaneous utility function is

$$\mathcal{V}(c,v,q;A,H) = \mathcal{U}(c,q;A,H) + \left(u_v + u_{vc}c + u_{vA}A + \frac{u_{vv}}{2}v\right)v \tag{10}$$

where $\mathcal{U}(\cdot)$ is defined in 3 and $u_{vv} < 0$. Analogously to the benchmark case, the positive term u_v describes the enjoyableness of the harm reduction method, absent previous and current consumption, and $u_{vA} \geq 0$ describes the addictiveness of the harm reduction method (i.e. the effect of past use on the marginal utility of current use). The term u_{vc} describes the substitutability (in preferences) between simultaneous consumption of the two addictive goods. It is negative if the harm reduction method is a substitute for the original addictive good (which is what we assume). In contrast, that term would be positive if the harm reduction method is a complement with the original addictive good.[7]

Overall, use of the harm reduction method v can affect consumption of the original addictive good directly through preferences, and indirectly through the accumulation of addiction experiences and health harms. The direct channel involves individual preferences, in that v is enjoyable ($u_v > 0$), addictive ($u_{vA} > 0$) and affects the marginal utility of the original addictive

[6] This is a reasonable assumption to make, but there have been times when, although the perceived harm was lower than the original addictive good, the actual harm was greater than the substance they were intended to replace. For example, heroin was originally marketed as a safe and non-addictive alternative to morphine, and OxyContin was originally marketed as a safer and less addictive alternative to older opioid pain relievers.

[7] There are harm reduction methods that decrease the marginal utility of the original addictive good, but are non-addictive. These include antagonists, which block rather than activate the receptors used by the original addictive good. Examples include the opioid antagonists naloxone and naltrexone, and the alcohol antagonist Antabuse (disulfiram). Here, for clarity and focus we assume it is a full or partial agonist rather than an antagonist. However, as shown in Section 4, our model is flexible and allows the harm reduction method to be an antagonist. Note that, by referring to substitutability in preferences, we consider a property of the utility function, as described by the cross-derivative u_{vc}. We are not referring to the definition of gross complements and substitutes, used to describe how the demand for one good responds to changes in the price of another good.

good (through u_{vc}). The indirect channel passes through the accumulation of addictive stock and health harms, as described by ε_A and ε_H in the law of motion of addiction and health harms.

The steady state consumption of the harm reduction method can be described as follows

$$v_d^{ss} = \theta_v \left(u_v - p_v - \pi_v\right) + \gamma c^{ss} \tag{11}$$

where θ_v is positive, and $\pi_v > 0$ describes the decrease in future utility resulting from the harm reduction method raising the addictive stock A and health harms H (see Appendix A.2 for details).[8]

The logic of equation 11 is that people will consume the harm reduction method only if the benefits exceed the costs. In this case the benefits include the enjoyability of the harm reduction method (u_v), while the costs include the monetary costs of purchasing the harm reduction method (p_v) and the future harmful consequences of consuming it (π_v). Moreover, due to the interdependence between the two addictive goods, use of the harm reduction method also depends on the consumption of the original addictive good.

Specifically, the term γ, which is multiplied by the steady-state level of consumption of the original addictive good (c^{ss}), is higher the greater u_{vA}, u_{cA} and u_{vc}, i.e. the greater the degree of complementarity between the goods and the addictive stock (see eq. 45). This occurs because both goods contribute to the accumulation of the addictive stock A, so that consumption of one good reinforces consumption of the other one. On the contrary, the higher the substitutability between the original addictive good and the harm reduction method (i.e. low values of u_{vc}), the lower the mutual reinforcing effect of consuming c and v, and γ.

The new steady-state levels of consumption of the addictive good and health harm are

$$c_d^{ss} = c^{ss} + \theta_c \left(u_{vA} - r_H\left(u_{vc}\right)\right) v_d^{ss} \tag{12}$$

$$H_d^{ss} = H^{ss} + \theta_H \left(u_{vA} - r_L\left(u_{vc}\right)\right) v_d^{ss} \tag{13}$$

where $\theta_c, \theta_H > 0$, and $r_L\left(u_{vc}\right) < r_L\left(u_{vc}\right)$ are threshold levels that depend on the degree of substitutability between v and c.[9]

Equations 12 and 13 will be used in the next Section to illustrate the conditions under which the introduction of a harm reduction method increases or decreases the consumption of

[8] Subscript d is mnemonic of dual consumption (i.e., after the introduction of the harm reduction method). Recall that the benchmark steady-state values have no subscripts.

[9] The steady-state conditions hold even if parameters are allowed to depend on age-class. In such a case, the steady-state would depend on the parameters that hold at old age, while the policy function would still be linear in addiction and health harm, although with age-specific parameters.

the original addictive good and the magnitude of health harms.

3 Consequences of the introduction of a harm reduction method

Just because a harm reduction method becomes available does not mean that people will use it. The model implies that people will use the harm reduction method in the long run when

$$u_v > p_v + \pi_v - \frac{\gamma}{\theta_v} c^{ss} \tag{14}$$

Based on the previous discussion about the determinants of γ, the following holds:

Remark 1 *Use of the harm reduction method is more likely the greater its enjoyableness and the lower its full price (which includes monetary price and the future health harms). If c and v are sufficiently addictive and complements, previous consumers of the original addictive good are more likely to eventually use the harm reduction method then previous abstainers.*

In what follows, we focus on scenarios in which the individual decides to use the harm reduction method (i.e. condition 14 is satisfied). We show that the consequences of using the harm reduction method depend on three key factors: the enjoyableness of the harm reduction method (u_v), the addictiveness of the harm reduction method (u_{vA}), and the substitutability of the harm reduction method for the original addictive good (u_{vc}). For later reference, consider the following terminology:

Definition 1 *A harm reduction method is defined as*

- *Mildly addictive if* $u_{vA} < r_L(u_{vc})$;
- *Moderately addictive if* $u_{vA} \in (r_L(u_{vc}), r_H(u_{vc}))$;
- *Highly addictive if* $u_{vA} > r_H(u_{vc})$.

According to the above definitions, the special case of a harm reduction method that is completely non-addictive and non-harmful ($u_{vA} = \varepsilon_A = \varepsilon_H = 0$) and a substitute for the original addictive good ($u_{vc} < 0$) would be categorized as mildly addictive.

We will discuss implications for two types of individuals. First, we consider impacts on the consumers of the original addictive good. This category of consumers (e.g. smokers or heroin users) is a main target for policies that aim at reducing health harms and addiction. Second, we consider impacts on individuals who were abstaining from the addictive good before the harm reduction method was introduced. This group is of interest because their use of the harm

Harm reduction method used?	Addictiveness of harm reduction method: Mild	Moderate	High
Yes	**Harm reduction** $c^{ss}\downarrow$, $H^{ss}\downarrow$	**Substitution** $c^{ss}\downarrow$, $H^{ss}\uparrow$	**Harm reduction backfires** $c^{ss}\uparrow$, $H^{ss}\uparrow$
No	**Harm reduction method irrelevant** $v^{ss}=0$; c^{ss} and H^{ss} remain unchanged		

Table 2: Long-run changes in consumption and health harm for an individual **already consuming** the addictive good prior to the introduction of a harm reduction method ($c^{ss} > 0$). If the harm reduction method v is enjoyable enough, harm reduction results when it is mildly addictive ($u_{vA} < r_L$), substitution results when v is moderately addictive ($u_{vA} \in (r_L, r_H)$), and harm reduction backfires – the worst case scenario – when v is highly addictive ($u_{vA} > r_H$).

reduction method does not entail any benefits in terms of reduced use of the original addictive good, and because of concerns that the harm reduction method could turn out to be a gateway, leading them to consume even the original addictive good from which they originally abstained.

3.1 Effects on consumption of the original addictive good: gateway effects, substitution effects, and quitting

To assess how use of the harm reduction method affects consumption of the original addictive good, consider expression 12, which can be conveniently rewritten as

$$c_d^{ss} - c^{ss} = \theta_c \left(u_{vA} - r_H \left(u_{vc}\right)\right) v_d^{ss} \tag{15}$$

Consistent with the intuition, equation 15 shows that, if the individual is not using the harm reduction method in the long run, the steady-state consumption of the original addictive good is unaffected and the introduction of a harm reduction method is irrelevant in the long run: consumption, addiction and health remain unchanged.

Moreover, since θ_c is a positive parameter the following result holds:

Proposition 1 (Consumption of the original addictive good) *Conditional on using the harm reduction method, consumption of the original addictive good c increases in the long run if the harm reduction method v is highly addictive ($u_{vA} > r_H$), and it decreases otherwise.*

Table 2 organizes the results for an individual already consuming the original addictive good prior to the introduction of a harm reduction method. The first row concerns the case where the individual uses the harm reduction method, and the second row concerns the case where the individual does not use it, after it has been introduced. Columns 1, 2, and 3 describe the cases where the addictiveness of the harm reduction method is mild, moderate, and high.

When the harm reduction method is highly addictive, steady-state consumption of the addictive good increases. Intuitively, by using the harm reduction method, the addictive stock A increases, which increases the marginal utility not just of the harm reduction method but also of the original addictive good. Incentivized by the higher marginal utility of consumption, the individual increases their consumption of the original addictive good. Hence, the harm reduction policy fails: use of v induces increased consumption of the original addictive good c. Overall, dual consumption and higher health harm result, as illustrated in the right-most panel of Figure 2.[10]

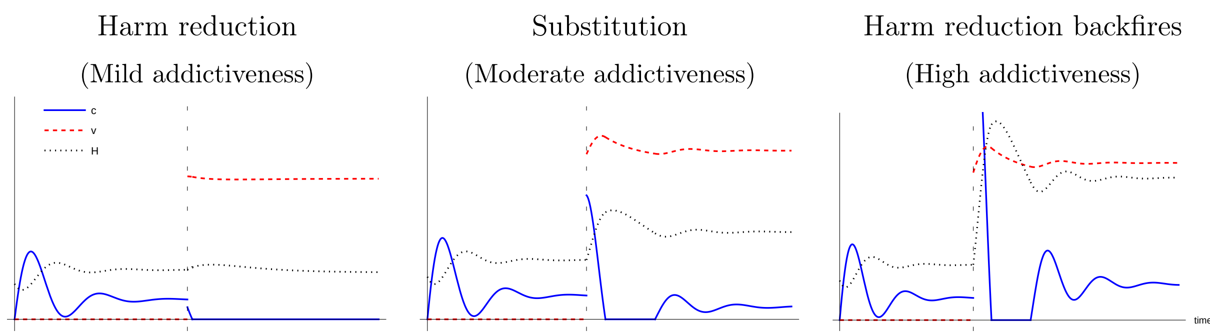

Figure 2: Illustration of some possible trajectories of consumption and health harm for an individual already consuming the addictive good c before the harm reduction method v is introduced. The vertical dashed line denotes when the method becomes available. Left panel: the consumer quits the original addictive good and health harm is reduced. Center panel: the harm reduction method substitutes for the original addictive good, health harm increases. Right panel: addictive consumption and health harm increase. In all panels, the individual condition 14 is satisfied and the individual uses the harm reduction method.

Remark 2 *When the harm reduction method is highly addictive, the availability and use of the harm reduction method backfires, because it induces higher consumption of the original addictive good.*

A particularly undesirable outcome related to Remark 2 arises when the introduction of the harm reduction method leads previous abstainers to initiate consumption of the original addictive good (see Table 3). In this case the harm reduction method has become a gateway drug. It is the consequence of the harm reduction method being enjoyable and highly addictive which produces dual addictive consumption. This outcome is more likely, the lower the threshold r_H defining the harm reduction method as highly addictive.

[10]When the harm reduction method is not highly addictive, its use effectively replaces consumption of the original addictive good. Observationally, at each point in time the two goods behave like complements. If, instead, the harm reduction method is not highly addictive, they behave like substitutes.

Harm reduction method used?	Addictiveness of harm reduction method: Mild or Moderate	High
Yes	**Initiate only the harm reduction method** $c^{ss} = 0$, $H^{ss} \uparrow$	**Gateway effect: Initiate both the harm reduction method and addictive good** $c^{ss} \uparrow$, $H^{ss} \uparrow$
No	**Harm reduction method is irrelevant** $v^{ss} = c^{ss} = H^{ss} = 0$	

Table 3: Long-run changes in consumption and health harm for an individual **previously abstaining** from the addictive good c prior to the introduction of a harm reduction method. If the harm reduction method is used, health harm always increases. Gateway effects and initiation with both addictive goods result if v is highly addictive ($u_{vA} > r_H$): the harm reduction policy backfires.

Consider now the possibility that consumption of the original addictive good decreases. Based on the fact that $\partial r_H / \partial u_{vc} < 0$ (see eq. equation 52), the following holds:

Remark 3 *Conditional on using the harm reduction method, the long run consumption of the original addictive good is more likely to decrease:*

- *The lower the addictiveness of the harm reduction method;*
- *The greater the substitutability between the harm reduction method and the original addictive good.*

The two left-most panels of Figure 2 illustrate this case when v is only mildly or moderately addictive, so that consumption of the original addictive good (solid line) tends to decrease over time. Importantly, in such cases the harm reduction method can lead the individual to not only reduce consumption of the original addictive good but quit it altogether (i.e. $c^{ss} > c_d^{ss} = 0$).

Remark 4 (Quitting the original addictive good) *For an individual previously using the original addictive good, quitting is more likely*

a. *The lower the consumption of the original addictive good c^{ss};*

b. *The lower the addictiveness of the harm reduction method u_{vA};*

c. *The greater the use of the harm reduction method v_d^{ss}.*

Condition (a) states an intuitive condition: it is harder for a heavy smoker than for a light smoker to quit. The logic of condition (b) is that, all else equal, a less addictive harm reduction method promotes quitting because it contributes less to the addictive stock and thus does less (via adjacent complementarity) to increase the marginal utility of the original addictive good. Condition (c) is of particular interest, because there may be concern about the health harms due to high consumption of the harm reduction method, but the trade-off is that greater use of the harm reduction method increases the likelihood that previous users will quit the addictive good.

Another important insight from Figure 2 is that the introduction of a harm reduction method can cause a short-run increase in the consumption of the original addictive good, even if the steady state consumption of the original addictive good will eventually be zero. Importantly, the success of a harm reduction approach depends critically on when one examines outcomes – in Figure 2, if one looked at the periods immediately after the introduction of the harm reduction method, things seem to have gotten worse, as both consumption of the original addictive good and the consumption of the harm reduction method are high. However, after some time periods the steady-state consumption of the original addictive good goes below the level it would have in the scenario where the harm reduction method is not available.

3.2 When do harm reduction policies decrease or increase health harms?

In the previous sections we have examined the conditions under which the introduction of a harm reduction method can increase or decrease consumption of the original addictive good. In this section, we examine the conditions under which it decreases or increases health harms, which may be the most important outcome to policymakers.

For individuals who were abstaining from the addictive good (e.g. non smokers, non opioid users), the answer is intuitive: if they begin using the harm reduction method, health harms worsen. The reason is that, although the harm reduction method is less harmful than the original good, it is still harmful. And for those previously abstaining, the consumption of the harm reduction method is not accompanied by any reduction in the use of the original addictive good, so there are no offsetting reductions in health harms from that source. Moreover, if the harm reduction method is highly addictive, it is possible that it acts as a gateway drug, inducing initiation of the addictive good (see Table 3). This is clearly the worst case scenario, with dual addictive consumption and worse health among persons who were previously abstaining from the original addictive good.

Remark 5 *If a previous abstainer of the original addictive good uses the harm reduction method, health harm increases.*

We next consider people who were previously using the original addictive good. To examine the conditions under which harm reduction can decrease or increase health harms for this group, rewrite equation 13 in terms of long-run change in health after the introduction of a harm reduction method:

$$H_d^{ss} - H^{ss} = \theta_H \left(u_{vA} - r_L \left(u_{vc}\right)\right) v_d^{ss}, \tag{16}$$

Proposition 2 *After the introduction of a harm reduction method, health harms among users of the original addictive good eventually decrease if the harm reduction method is mildly addictive, and increase if the harm reduction method is moderately or highly addictive.*

The key factors determining whether health harms increase or decrease among this group are the addictiveness of the harm reduction method; and the substitutability of the harm reduction method for the original addictive good.

Consider addictiveness. If the harm reduction method is mildly addictive, $u_{vA} < r_L \left(u_{vc}\right)$, then the stock of health harms H falls (see Table 2). Moreover, as shown in the previous Section, for a previous consumer of the original addictive good c, the consumption of c also falls. This represents an unambiguous success of the harm reduction approach – introducing the new addictive option leads to a reduction in consumption of the original addictive good that is large enough to compensate the health harm due to the use of the harm reduction method. Hence health ultimately improves (see Figure 2, left-most panel, for an illustration).

Note that quitting the original addictive good does not guarantee that health improves; it is possible for health to worsen despite quitting the original addictive good, if the harm reduction method is particularly harmful. Even if consumption of the addictive good falls after the introduction of the harm reduction method, the steady state level of health harms can rise if the addictiveness of the harm reduction method is moderate (see eq. 16). The reason is that, even though consumption of the addictive good has declined, the individual is also using the harm reduction method, which itself contributes to both the addictive stock A and the stock of health harms H. In one sense the harm reduction approach has been successful - it has reduced consumption of the original addictive good – but in another sense it has failed because it has worsened the health of those who were previously using the addictive good.

Finally, if the harm reduction is highly addictive, not only does health harm increase, but so does consumption of the original addictive good. In such a case, the harm reduction policy is unambiguously a failure. Thus, if the harm reduction method is highly addictive, then it backfires.

Note that threshold $r_L \left(u_{vc}\right)$, like the threshold $r_H \left(u_{vc}\right)$, is a function of the substitutability of the harm reduction method for the original addictive good. In particular, $\partial r_L / \partial u_{vc} < 0$ and the following Remark holds:

Remark 6 *The greater the substitutability of the harm reduction method for the original addictive good (i.e. the lower u_{vc}), the higher the threshold level $r_L \left(u_{vc}\right)$.*

Hence, the greater the substitutability between the harm reduction method and the original addictive good, the more likely that the introduction of the harm reduction method will lead to a decrease in health harm. The intuition is straightforward: since the harm reduction method is less harmful than the original addictive good, the more the harm reduction method is perceived by the consumer as a close substitute for the original addictive good, the more likely is the consumer to reduce consumption of c, increase consumption of v, resulting in a decline in health harms. In contrast, if the consumer perceives the harm reduction method is a complement with the original addictive good, that would make joint consumption of the two substances more likely, worsening health harms.

Table A1 provides an overview of the results for both abstainers and previous consumers of the addictive good. Note that, when examining the impact on health harms, timing is once again critical. If one examined only early time periods, one might miss later quitting of the original addictive good and perhaps even quitting of the harm reduction method. Immediate evaluation of harm reduction methods may give a misleading impression of steady-state effects.

4 Extension 1: Taxing the harm reduction method

In this Section we show how changes in the price of the harm reduction method due to, e.g. changes in taxation, can influence the demand for addictive consumption and health harm. (We show the effect of changes in the price of the harm reduction method here, and we show the effect of a change in the price of the original addictive good in Appendix A.3.)

Due to the absence of income effects, direct price effects are negative: $\partial v_d^{ss}/\partial p_v < 0$. When considering cross-price effects, the addictiveness of the harm reduction method is a key driver. Specifically:

$$\frac{\partial c_d^{ss}}{\partial p_v} > 0 \qquad \Longleftrightarrow \qquad u_{vA} < r_H\left(u_{vc}\right) \tag{17}$$

That is, when the harm reduction method is mildly or moderately addictive, an increase in the price of the harm reduction method induces an increase in the consumption of the original addictive good. Empirically, the evidence shows that when the total price of vaping rises (e.g. due to taxes) then smoking of combustible cigarettes increases (see, e.g. Pesko and Currie, 2019; Pesko et al., 2020; Pesko and Warman, 2022), which suggests that vaping is moderately or mildly addictive for consumers. As shown in the previous Section, these two cases correspond to the scenarios in which introducing the harm reduction method leads to a reduction in the consumption of the original addictive good and, possibly, a reduction in health harms. Under such conditions, taxing harm reduction methods may worsen health harms.

The health consequences of higher taxes on e-cigarettes (vaping or v) can be explicitly assessed considering that

$$\frac{\partial H_d^{ss}}{\partial p_v} > 0 \qquad \Longleftrightarrow \qquad u_{vA} < r_L\left(u_{vc}\right) \tag{18}$$

That is, an increase in the price of the harm reduction method increases health harm if the harm reduction method is mildly addictive. Consistent with the predictions presented in the previous Section, we conclude that taxing the harm reduction method can produce different results, depending on its addictiveness.

Remark 7 *Taxing the harm reduction method:*

- *Increases consumption of the original addictive good if the harm reduction method is either mildly or moderately addictive,* $u_{vA} < r_H\left(u_{vc}\right)$
- *Increases health harm if the harm reduction method is mildly addictive,* $u_{vA} < r_L\left(u_{vc}\right)$
- *Decreases consumption of the original addictive good and decreases health harm if the harm reduction method is highly addictive,* $u_{vA} > r_H\left(u_{vc}\right)$

An implication of the above Remark is that, in the intermediate case in which the harm reduction method is moderately addictive, $u_{vA} \in (r_L, r_H))$, taxation of the harm reduction method will increase consumption of the original addictive good, and yet lead to a health improvement, because the health benefit of reduced vaping outweighs the increased harm from greater smoking. If the harm reduction method is highly addictive then we know from the earlier results summarized in Table 2 that the harm reduction method is backfiring, and causing people to actually consume more of the addictive good. In this case, raising taxes on the harm reduction method has the benefit of reducing consumption of the original addictive good and reducing health harms.

5 Extension 2: The harm reduction method is an antagonist

In the interests of clarity, some simplifying assumptions were made. We now discuss how relaxing those assumptions affects the predictions and implications of the model. For example, we assumed that the harm reduction method acts like an agonist, in that it binds with and activates the same receptors of the original drug ($\varepsilon_A > 0$), it is pleasurable (i.e. gives an euphoric effect, $u_v > 0$), and it is addictive ($u_{vA} > 0$). Essentially, the agonist harm reduction method impersonates the original addictive good, possibly inducing substitution effects that induce lower consumption of c and lower health harm. Our model is flexible enough to study the effects of harm reduction methods that instead act as an antagonist, like the opioid antagonists naloxone and naltrexone, or the alcohol antagonist Antabuse (disulfiram). Similarly to agonist methods, antagonists bind with the same receptors of the original addictive drug. The main difference is that antagonists block these receptors, reducing (or preventing) the pleasure of consumption the original addictive good ($u_{vc} < 0$). Moreover, they are typically not enjoyable ($u_v = 0$) or addictive ($\varepsilon_A = u_{vA} = 0$).

By imposing the above restrictions in our model, it is easy to show that a previous abstainer of the addictive good would not use the antagonist harm reduction method (see eq. 14), as it does not provide positive marginal utility of consumption. The case of a previous user of the original addictive good is more interesting. If we relax the earlier rational choice model to incorporate behavioral economics concepts such as time inconsistent preferences (hyperbolic discounting) and limited self-control, such people may demand an antagonist harm reduction method as a precommitment device (see, e.g., Schelling, 1978; Thaler and Shefrin, 1981; Laibson, 1997). Accordingly, demand for antagonist harm reduction methods by individuals consuming addictive goods can be interpreted as evidence that such individuals are sophisticated about their time-inconsistent preferences and/or self-control failures (Strotz, 1955; Gruber and Köszegi, 2001).

However, this is not the only possible explanation, as the demand for a harm reduction method can be also the result of a time-consistent, rational choice along an optimal consumption trajectory that will ultimately lead to quitting of both the original addictive good and the harm reduction method. As shown in the previous Section, along this optimal trajectory patterns of intermittent consumption are possible. Accordingly, there may be periods in which the harm reduction method is used, even if eventually the individual will not demand them in the long-run.

6 Extension 3: Banning the harm reduction method

So far our analysis has focused on the introduction of a harm reduction method, but the model can also be used to examine the reverse: when a harm reduction method is withdrawn from the market. An example of this is when, in 2022, the U.S. Food & Drug Administration (FDA) banned Juul from marketing its ENDS products in the United States, while cigarettes remain legal to sell (U.S. Food and Drug Administration, 2022).

As shown earlier, when the harm reduction method is not highly addictive (i.e. it is either moderately or mildly addictive), its use is negatively related to the consumption of the addictive good (i.e. they are substitutes). As a result, if the harm reduction method is banned, the consumption of the original addictive good is expected to increase. If the harm reduction method is mildly addictive (as opposed to moderately or highly addictive), the increase in consumption of c is so large that health harm increases. However, if the harm reduction method is highly addictive, then c and v move together as complements. As a result, when the harm reduction method is no longer available, consumption of the addictive good c decreases, and due to the reduced consumption of both goods, health harm decreases.

7 Discussion

Harm reduction methods are controversial. Advocates argue that they can increase the quitting of addictive substances, and, even if not, will still reduce overall health harms. Opponents argue that there is a risk of moral hazard – that introducing a harm reduction method may make quitting less likely and may lead to increased addiction, and that previous abstainers may initiate use of the harm reduction method precisely because it is safer. Opponents also warn that a harm reduction method could be a gateway drug that leads some people who previously abstained to begin using the original addictive good.

This paper outlines the conditions under which each of these predictions is correct. We provide a model of harm reduction, an implication of which is that the introduction of a novel harm reduction technique is neither always good nor always bad. Depending on the characteristics of the harm reduction method, it may not be consumed at all, may be consumed by those previously taking the original addictive good, and/or may be consumed even by those who previously abstained from the original addictive good. Also, depending on the characteristics of the harm reduction method, it can lead current users of the addictive good to quit, it can lead current users to increase their consumption of the original addictive good, or it can lead past abstainers to initiate the original addictive good.

There are three critical characteristics of the harm reduction method that determine which of these outcomes will occur. The first is its enjoyableness – do the benefits of the harm reduction method in terms of marginal utility of consumption exceed the costs in terms of monetary price and future health harms? This will determine whether people consume the harm reduction method. For those who do not consume it, nothing changes. They continue to have the same steady state consumption of the original addictive good as before.

For those who do consume the harm reduction method, whether or not it leads previous users of the original addictive good to quit or not, and whether it leads previous abstainers to begin using the original addictive good, is determined by the second and third critical factors. The second factor is the addictiveness of the harm reduction method. This is critical because the more the harm reduction method contributes to the addictive stock, the more it increases the marginal utility not just of the harm reduction method but also the original addictive good.

A harm reduction method that is highly addictive will be more likely to lead previous users of the addictive good to increase their consumption, and will be more likely to induce previous abstainers to initiate use of the addictive good. The third critical factor is the extent to which the new harm reduction method is a substitute for (as opposed to a complement with) the original addictive good. The greater the extent to which it is a substitute, the less likely it leads previous abstainers to initiate and the less likely it worsens health harms.

An important insight from the model is that the effect of the new harm reduction method depends critically on which time period is examined. Depending on the time period examined, one might see use of original addictive good increasing or decreasing. In our simulations, the situation sometimes seems worse in the early periods, with use of both the harm reduction and original addictive good rising initially. Under some conditions, however, the outlook improves with time, as people decrease their consumption over time as health consequences mount. Thus, one should be careful that evaluations conducted immediately after the introduction of the harm reduction may be misleading, and it may take time to determine how the harm reduction method has affected steady-state consumption of the original addictive good.

The model also indicates that there are trade-offs to reducing access to harm reduction methods. On the one hand, restricting access to the harm reduction method can reduce the health harms that arise specifically from the harm reduction method, but on the other hand restricting access to the harm reduction method makes it harder for consumers to switch away from the original addictive good, potentially leaving them in worse health and more heavily addicted.

The model of harm reduction used in this paper applies to a variety of cases, including ENDS and NRT for combustible cigarettes, methadone and buprenorphine for heroin and other opioids, and edible THC products for combustible marijuana.

The model implies a variety of policy levers that the government can use to affect the likelihood that the introduction of the harm reduction method succeeds in reducing health harms and consumption of the original addictive good:

1) Whether the government chooses to allow the harm reduction method on the market at all. For example, a government may decide whether to give regulatory approval for a new prescription drug, such as buprenorphine, or a new over-the-counter product such as Electronic Nicotine Delivery Systems (ENDS) or e-cigarettes. In recent examples from the U.S., the FDA authorized the marketing of ENDS devices in 2021 (U.S. Food and Drug Administration, 2021) but the same agency a year later issued marketing denial orders to Juul, prohibiting them from selling their ENDS products in the U.S. (U.S. Food and Drug Administration, 2022).

2) The government can selectively regulate access to the harm reduction method. Govern-

ments may in particular want to restrict access by youth (Abouk and Adams, 2017; DeSimone et al., 2022). In 2019 the U.S. raised the minimum age to purchase e-cigarettes from 18 to 21 (U.S. Food and Drug Administration, 2020b). It may also impose limits on a doctor's ability to prescribe prescription harm reduction products; for example, the U.S. limits the number of patients to whom a physician may prescribe buprenorphine (SAMHSA, 2021). The government may require that these prescription methods be administered under certain conditions; for example, methadone is often provided only in a clinic; it is rarely given to patients for home consumption. In addition, harm reduction methods can be made available only with a prescription, rather than over-the-counter. This may help ensure that the quantity of harm reduction method consumed will not be so great as to worsen health harms.

3) The government may regulate the addictiveness of the harm reduction method (u_{vA}). This factor turns out to be critical in determining what happens to consumption of the original addictive good. Such regulation could, for example, limit the potency of buprenorphine doses, the amount of THC in edible marijuana products, and the amount of nicotine that can be delivered in an increment of time by an e-cigarette.

4) Governments may seek to reduce the health harms of the harm reduction method (ε_H in our model). For example, they might set high safety requirements for vaping.

5) Governments may seek to decrease the marginal utility u_v of the harm reduction method. For example, in 2020, the FDA banned flavored ENDS that might appeal to youth, including fruit and mint flavors (U.S. Food and Drug Administration, 2020a). Regulation of advertising is another potential way of reducing the marginal utility of consumption.

6) Governments may tax either the original good or the harm reduction methods in order to raise their monetary price and decrease demand for them (Pesko et al., 2020). There is substantial variation in the rate at which U.S. states tax e-cigarettes; 21 states do not tax them at all, and among states that do tax them the rates vary from 8% to 95% (IGEN, 2022). Clean indoor air laws that ban vaping increase the total cost of vaping by raising its time cost.

A limitation of the paper is that, while the model does yield equations for the steady-state consumption of the harm reduction method and the original addictive good, as well as the consumption paths leading to the steady state, there are difficulties in empirically estimating them because one cannot easily measure or observe key parameters such as the addictiveness of the two substances, the health harms of the two substances, and the marginal utility of the two substances. At a minimum, however, the model offers a way of identifying the factors critical to determining whether the introduction of a harm reduction method succeeds in reducing health harms and consumption of the original addictive good.

Despite its limitations, this paper contributes to the literature by proposing a model of harm reduction, the implications of which indicate that neither advocates nor opponents are always correct. The introduction of a harm reduction method can facilitate quitting and reduce health harms, as advocates claim, or can backfire and lead to not just to increased use of the addictive good and worsening health harms but the initiation of the addictive good by previous abstainers, as opponents fear. The model also indicates the key factors that determine which of these outcomes occur; these are the enjoyableness of the harm reduction method, the addictiveness of the harm reduction method, and the extent to which the harm reduction method is a substitute for (as opposed to a complement to) the original addictive good.

References

Abouk, R. and S. Adams (2017). Bans on electronic cigarette sales to minors and smoking among high school students. *Journal of Health Economics 54*, 17–24. 1, 7

Abouk, R., C. J. Courtemanche, D. M. Dave, B. Feng, A. S. Friedman, J. C. Maclean, M. F. Pesko, J. J. Sabia, and S. Safford (2021). Intended and unintended effects of e-cigarette taxes on youth tobacco use. NBER working paper 29216, National Bureau of Economic Research. 1

Abouk, R., R. L. Pacula, and D. Powell (2019). Association between state laws facilitating pharmacy distribution of naloxone and risk of fatal overdose. *JAMA Internal Medicine 179*(6), 805–811. 1

Agrawal, V., N. Sood, and C. M. Whaley (2022, October). The ex-ante moral hazard effects of covid-19 vaccines. Working Paper 30602, National Bureau of Economic Research. 1

Allcott, H. and C. Rafkin (2022). Optimal regulation of e-cigarettes: Theory and evidence. *American Economic Journal: Economic Policy 14(4)*, 1–50. 1, 2.1

Allen, L., M. Burns, and B. Saloner (2022). The consequences of removing prior authorization for buprenorphine in Medicaid — Building an evidence base. *JAMA Health Forum 3(6)*, e220189–e220189. 1

Alpert, A., W. N. Evans, E. M. Lieber, and D. Powell (2022). Origins of the opioid crisis and its enduring impacts. *The Quarterly Journal of Economics 137*(2), 1139–1179. 1

Barrette, E., L. Dafny, and K. Shen (2021). Do policies to increase access to treatment for opioid use disorder work? NBER working paper 29001, National Bureau of Economic Research. 1

Becker, G. S., M. Grossman, and K. M. Murphy (1991). Rational addiction and the effect of price on consumption. *American Economic Review 81*(2), 237–241. 2.1

Becker, G. S. and K. M. Murphy (1988). A theory of rational addiction. *The Journal of Political Economy 96*(4), 675–700. 2, 2.1, 4

Bhattacharya, J., M. K. Bundorf, N. Pace, and N. Sood (2011). Does health insurance make you fat? In *Economic Aspects of Obesity*, pp. 35–64. University of Chicago Press. 1

Bishai, D., J. Sindelar, E. Ricketts, S. Huettner, L. Cornelius, J. Lloyd, J. Havens, C. Latkin, and S. Strathdee (2008). Willingness to pay for drug rehabilitation: Implications for cost recovery. *Journal of Health Economics 27*(4), 959–972. 1

Campbell, D. (2009). Vatican's stance on UN drugs policy "risks lives". `https://www.theguardian.com/world/2009/feb/27/vatican-un-drugs-policy`, last accessed on 2022-07-07. 3

Chaloupka, F. J. (1991). Rational addictive behavior and cigarette smoking. *Journal of Political Economy 99*(4), 722–742. 4

Clarke, E., K. Thompson, S. Weaver, J. Thompson, and G. O'Connell (2019). Snus: A compelling harm reduction alternative to cigarettes. *Harm Reduction Journal 16*, 62. 1

Cotti, C., E. Nesson, and N. Tefft (2019). Impacts of the ACA medicaid expansion on health behaviors: evidence from household panel data. *Health Economics 28*(2), 219–244. 1

Cotti, C. D., C. J. Courtemanche, J. C. Maclean, E. T. Nesson, M. F. Pesko, and N. Tefft (2020). The effects of e-cigarette taxes on e-cigarette prices and tobacco product sales: Evidence from retail panel data. NBER working paper 26724, National Bureau of Economic Research. 1

Dave, D. M., R. Kaestner, and G. L. Wehby (2019). Does public insurance coverage for pregnant women affect prenatal health behaviors? *Journal of Population Economics 32*(2), 419–453. 1

DeSimone, J. S., D. S. Grossman, and N. R. Ziebarth (2022). Regression discontinuity evidence on the effectiveness of the minimum legal e-cigarette purchasing age. Working Paper 30614, National Bureau of Economic Research. 7

Dockner, E. J. and G. Feichtinger (1993). Cyclical consumption patterns and rational addiction. *American Economic Review 83*(1), 256–263. 2

Doleac, J. L. and A. Mukherjee (2022). The effects of naloxone access laws on opioid abuse, mortality, and crime. *The Journal of Law and Economics 65*(2), 211–238. 1, 1

Dragone, D. and D. Raggi (2018). Testing rational addiction: When lifetime is uncertain, one lag is enough. Working paper 1119, Department of Economics, Università di Bologna. 4

Dragone, D. and D. Raggi (2021). Resolving the milk addiction paradox. *Journal of Health Economics 77*, 102452. 4

Dragone, D. and P. Vanin (2022). Substitution effects in intertemporal problems. *American Economic Journal: Microeconomics 14*(3), 791–809. 4

Erickson, P. G. (1995). Harm reduction: What it is and is not. *Drug and Alcohol Review 14*(3), 283–285. 1

Friedman, A. S. (2015). How does electronic cigarette access affect adolescent smoking? *Journal of Health Economics 44*, 300–308. 1

Frio, G. S. and M. T. A. França (2021). Human papillomavirus vaccine and risky sexual behavior: Regression discontinuity design evidence from Brazil. *Economics & Human Biology 40*, 100946. 1

Grimm, C. A. (2020). *Geographic disparities affect access to buprenorphine services for opioid use disorder*. US Department of Health and Human Services - Office of the Inspector General, OEI-12-17-00240. 1

Gruber, J. and B. Köszegi (2001). Is addiction "rational"? theory and evidence. *The Quarterly Journal of Economics 116*(4), 1261–1303. 4, 5

Harm Reduction International (2022). What is harm reduction? `https://www.hri.global/what-is-harm-reduction`, last accessed on 2022-07-07. 1

IGEN (2022). Vape & e-cig taxes by State. https://igentax.com/vape-tax-state, last accessed on 2022-07-07. 1, 7

Krawczyk, N., S. T. Allen, K. E. Schneider, K. Solomon, H. Shah, M. Morris, S. J. Harris, S. G. Sherman, and B. Saloner (2022). Intersecting substance use treatment and harm reduction services: Exploring the characteristics and service needs of a community-based sample of people who use drugs. *Harm Reduction Journal 19(95)*, 1–10. 2

Laibson, D. (1997). Golden eggs and hyperbolic discounting. *The Quarterly Journal of Economics 112*(2), 443–478. 5

Lakdawalla, D., N. Sood, and D. Goldman (2006). HIV breakthroughs and risky sexual behavior. *The Quarterly Journal of Economics 121*(3), 1063–1102. 1

Lillard, D. R. (2020). *The Economics of Nicotine Consumption, in: K. F. Zimmermann (ed.), Handbook of Labor, Human Resources and Population Economics*. Springer. 1

Maclean, J. C., J. Mallatt, C. J. Ruhm, and K. Simon (2021). Economic studies on the opioid crisis: A review. NBER working paper 28067, National Bureau of Economic Research. 1

Margolis, J., J. Hockenberry, M. Grossman, and S.-Y. Chou (2014). Moral hazard and less invasive medical treatment for coronary artery disease: The case of cigarette smoking. NBER working paper 28067, National Bureau of Economic Research. 1

Marshall, A. (1890). *Principles of Economics*. Macmillan and Co. Ltd.: London. 1

Marti, J., J. Buckell, J. C. Maclean, and J. Sindelar (2019). To "vape" or smoke? Experimental evidence on adult smokers. *Economic Inquiry 57*(1), 705–725. 1

National Harm Reduction Coalition (2023). Harm reduction issues: Medication for opioid use disorder. https://harmreduction.org/issues/medication-for-opioid-use-disorder/, last accessed on 2023-01-16. 2

Packham, A. (2022). Syringe exchange programs and harm reduction: New evidence in the wake of the opioid epidemic. *Journal of Public Economics 215*(1), 104733. 1

Peltzman, S. (1975). The effects of automobile safety regulation. *Journal of Political Economy 83*(4), 677–725. 1

Pesko, M. F., C. J. Courtemanche, and J. C. Maclean (2020). The effects of traditional cigarette and e-cigarette tax rates on adult tobacco product use. *Journal of Risk and Uncertainty 60*(3), 229–258. 1, 4, 7

Pesko, M. F. and J. M. Currie (2019). E-cigarette minimum legal sale age laws and traditional cigarette use among rural pregnant teenagers. *Journal of Health Economics 66*, 71–90. 1, 4

Pesko, M. F. and C. Warman (2022). Re-exploring the early relationship between teenage cigarette and e-cigarette use using price and tax changes. *Health Economics 31*(1), 137–153. 1, 4

Piccoli, L. and S. Tiezzi (2021). Rational addiction and time-consistency: An empirical test. *Journal of Health Economics 80*, 102546. 4

Powell, D., R. L. Pacula, and E. Taylor (2020). How increasing medical access to opioids contributes to the opioid epidemic: Evidence from medicare part D. *Journal of Health Economics 71*, 102286. 1

Rees, D. I., J. J. Sabia, L. M. Argys, D. Dave, and J. Latshaw (2019). With a little help from my friends: The effects of Good Samaritan and naloxone access laws on opioid-related deaths. *The Journal of Law and Economics 62*(1), 1–27. 1, 1

SAMHSA (2021). FAQs about the new buprenorphine practice guidelines. https://www.samhsa.gov/medication-assisted-treatment/become-buprenorphine-waivered-practitioner/new-practice-guidelines-faqs, last accessed on 2022-07-07. 7

SAMHSA (2023). Harm reduction. https://www.samhsa.gov/find-help/harm-reduction, last accessed on 2023-01-18. 1

Satel, S. (2019). The vaping overreaction. The Atlantic, https://www.theatlantic.com/ideas/archive/2019/10/danger-vaping-bans/600451, last accessed on 2022-07-07. 3

Schelling, T. C. (1978). Egonomics, or the art of self-management. *The American Economic Review 68*(2), 290–294. 5

Simon, K., A. Soni, and J. Cawley (2017). The impact of health insurance on preventive care and health behaviors: evidence from the first two years of the ACA medicaid expansions. *Journal of Policy Analysis and Management 36*(2), 390–417. 1

Singer, J. A. (2018). Harm reduction: Shifting from a war on drugs to a war on drug-related deaths. *Cato Policy Analysis 858*, 1–19. 2

Single, E. (1995). Defining harm reduction. *Drug and Alcohol Review 14*(3), 287–290. 1

Stancliff, S., H. Joseph, C. Fong, T. Furst, S. D. Comer, and P. Roux (2012). Opioid maintenance treatment as a harm reduction tool for opioid-dependent individuals in NYC: the need to expand access to buprenorphine in marginalized populations. *Journal of Addictive Diseases 31*(3), 278–287. 1, 2

Strotz, R. H. (1955). Myopia and inconsistency in dynamic utility maximization. *The Review of Economic Studies 23*(3), 165–180. 5

Thaler, R. H. and H. M. Shefrin (1981). An economic theory of self-control. *Journal of Political Economy 89*(2), 392–406. 5

U.S. Department of Health and Human Services (2023). Harm reduction. https://www.hhs.gov/overdose-prevention/harm-reduction,last accessed 2023-18-01. 1

U.S. Food and Drug Administration (2020a). FDA finalizes enforcement policy on unauthorized flavored cartridge-based e-cigarettes that appeal to children, including fruit and mint. FDA News Release, https://www.fda.gov/news-events/press-announcements/fda-finalizes-enforcement-policy-unauthorized-flavored-cartridge-based-e-cigarettes-appeal- last accessed on 2022-07-07. 7

U.S. Food and Drug Administration (2020b). Newly signed legislation raises federal minimum age of sale of tobacco products to 21. https://www.fda.gov/tobacco-products/ctp-newsroom/newly-signed-legislation-raises-federal-minimum-age-sale-tobacco-products-21, last accessed on 2022-07-07. 7

U.S. Food and Drug Administration (2021). FDA permits marketing of e-cigarette products, marking first authorization of its kind by the agency. FDA Press release. https://www.fda.gov/news-events/press-announcements/fda-permits-marketing-e-cigarette-products-marking-first-authorization-its-kind-agency, last accessed on 2022-07-07. 7

U.S. Food and Drug Administration (2022). FDA denies authorization to market JUUL products. FDA Press release. https://www.fda.gov/news-events/press-announcements/fda-denies-authorization-market-juul-products, last accessed on 2022-07-07. 1, 6, 7

Vestal, C. (2016). In opioid epidemic, prejudice persists against methadone. Pew Charitable Trusts, https://www.theatlantic.com/ideas/archive/2019/10/danger-vaping-bans/600451, last accessed on 2022-07-07. 3

Washington State Health Care Authority (2023). Harm reduction strategies. https://www.hca.wa.gov/assets/free-or-low-cost/82-0376-harm-reduction-strategies-trifold.pdf, last accessed on 2023-01-16. 2

Waters, R. C. (2019). In support of buprenorphine deregulation. *Family Practice Management 26*(4), 6–8. 1

Weber, E. and A. Gupta (2019). *State Medicaid programs should follow the 'Medicare Model': Remove prior authorization requirements for buprenorphine and other medications to treat opioid use disorders*. Legal Action Center, New York, NY. 1

WHO (2021). *WHO Report on the Global Tobacco Epidemic, 2021: Addressing new and emerging products*. World Health Organization. 1

WHO (2022a). Alcohol. https://www.who.int/news-room/fact-sheets/detail/alcohol, last accessed on 2022-07-07. 1

WHO (2022b). Opioid overdose. https://www.who.int/news-room/fact-sheets/detail/opioid-overdose, last accessed on 2022-07-07. 1

WHO (2022c). Tobacco. https://www.who.int/health-topics/tobacco, last accessed on 2022-07-07. 1

WHO (2022d). Tobacco: E-cigarettes. https://www.who.int/news-room/questions-and-answers/item/tobacco-e-cigarettes, last accessed on 2022-07-07. 1

A Appendix

A.1 Solving the benchmark model with one addictive good

After constructing the Hamiltonian function associated to the objective function 3 and the laws of motion 1 and 2, we compute the first order conditions with respect to c, and the law of motions of the costate variables associated to the addictive stock A and health harms H. By replacing the optimal consumption of the addictive good in the laws of motion of states an costates, we obtain a system of four differential equations that can be solved in closed-form. Specifically, replacing q from the budget constraint, the Hamiltonian function associated to the consumer's problem is

$$\begin{aligned}\mathcal{H}(c, v, q; A, H) = & \left(u_c + \frac{u_{cc}}{2}c + u_{cA}A\right)c + M - p_c c \\ & + \left(u_A + \frac{u_{AA}}{2}A\right)A + \left(u_H + \frac{u_{HH}}{2}H\right)H + u_{AH}AH \\ & + \lambda(c - \delta_A A) + \mu(c + \omega A - \delta_H H) \end{aligned} \tag{19}$$

where λ and μ are the shadow prices of A and H, respectively.

The necessary conditions for an internal solution are

$$\mathcal{H}_c = 0 \quad \Leftrightarrow \quad \underbrace{u_c + u_{cc}c + u_{cA}A}_{\mathcal{U}_c} = p_c - \lambda - \mu \tag{20}$$

$$\dot{\lambda} = (\rho + \delta_A)\lambda - \omega\mu - \underbrace{(u_A + u_{AA}A + u_{AH}H + u_{cA}c)}_{\mathcal{U}_A} \tag{21}$$

$$\dot{\mu} = (\rho + \delta_H)\mu - \underbrace{(u_H + u_{HH}H + u_{AH}A)}_{\mathcal{U}_H} \tag{22}$$

$$\dot{A} = c - \delta_A A \tag{23}$$

$$\dot{H} = c + \omega A - \delta_H H \tag{24}$$

together with the appropriate initial and transversality conditions. The first order condition 20 implies that the marginal benefit of consuming c must be equal to the marginal cost of consuming, which depends on the market price as well as on the shadow price of A and H. Note that the addiction stock affects consumption of the original good c directly (through $u_{cA}A$) and indirectly (through its shadow value λ), while health harms plays only an indirect role through μ.

The equation of motion of the shadow value of addiction (eq. 21) depends also on shadow value μ of health. Moreover, the marginal utility of addiction $\mathcal{U}_A$ directly depends on the addictiveness of c . In particular, u_{cA} reduces the shadow price of building up addiction because it increases the marginal utility of consuming the addictive goods. The law of motion of the shadow value of health harms (eq. 22), instead, does not depend on addiction nor on c.

Solving the foc for c yields the optimal consumption of the addictive good:

$$c^{*}=\frac{u_{c}-p_{c}+u_{cA}A+\lambda+\mu}{-u_{cc}} \tag{25}$$

Replacing in 20 to 24 and imposing $\dot{\lambda}=\dot{\mu}=\dot{A}=\dot{H}=0$ yields the steady state values λ^{ss}, μ^{ss}, A^{ss} and H^{ss}

Replacing them in 25 gives the steady state consumption of c:

$$c^{ss}=\alpha\left(u_{c}-p_{c}-\pi_{c}\right) \tag{26}$$

where

$$\alpha=\frac{\delta_{A}\delta_{H}\left(\delta_{A}+\rho\right)\left(\delta_{H}+\rho\right)}{-u_{cc}|J|}>0. \tag{27}$$

$|J|$ is the determinant of the Jacobian matrix (not shown) computed at the steady state,[11] and

$$\pi_{c}=-\frac{1}{\delta_{A}+\rho}u_{A}-\frac{\delta_{A}+\omega_{A}+\rho}{\left(\delta_{A}+\rho\right)\left(\delta_{H}+\rho\right)}u_{H}>0 \tag{28}$$

describes the non-monetary cost associated to the consumption of the addictive good. Moreover

$$H^{ss}=\frac{\delta_{A}+\omega_{A}}{\delta_{A}\delta_{H}}c^{ss},\qquad A^{ss}=\frac{1}{\delta_{A}}c^{ss} \tag{29}$$

It is possible to determine the level of consumption of the original addictive good along the optimal trajectory directed to the steady state, as a function of the states A and H. One possible way to obtain it is to replace 20 into the system of differential equations 21– 24, and then solve the system for given boundary conditions A_0, H_0, A^{ss} and H^{ss}. The solution is going to be a function of time, the initial conditions and a set of four eigenvalues and eigenvectors. Out of the four eigenvalues, two have always positive real parts. Imposing asymptotic stability and replacing the two expressions that depend on time, it is possible to obtain the following expression

$$\hat{c}\left(A,H\right)=a_{c}c^{ss}+a_{A}A+a_{H}H \tag{30}$$

where $\hat{c}$ denotes the "policy function", i.e. the optimal level of consumption as a function of

[11] We assume that the trajectories to the steady state are asymptotically stable, which implies that we focus on case in which two eigenvalues of the 4x4 Jacobian matrix associated to the dynamic system have non-positive real parts. When this is the case, $|J|$ is positive.

the state variables, and

$$\alpha_c = \frac{e_1 e_2}{\delta_A \delta_H} > 0 \tag{31}$$

$$\alpha_A = \frac{(\delta_A + e_1)(\delta_A + e_2) + (\delta_A + \delta_H + e_1 + e_2)\,\omega}{\delta_A - \delta_H + \omega} \tag{32}$$

$$\alpha_H = -\frac{(\delta_H + e_1)(\delta_H + e_2)}{\delta_A - \delta_H + \omega} \tag{33}$$

and $e_1, e_2 < 0$ are the eigenvalues with negative real parts associated to the Jacobian matrix of 21 to 24. If the eigenvalues are complex numbers, the policy function features oscillations.

A.2 Two goods

The solution follows the same procedure used in the previous Section. The Hamiltonian function associated to the consumer's problem is

$$\begin{aligned}\mathcal{H}(c, v, q; A, H) = {}& \left(u_c + \frac{u_{cc}}{2}c + u_{cA}A\right)c + \left(u_v + \frac{u_{vv}}{2}v + u_{vc}c + u_{vA}A\right)v \\ & + \left(u_A + \frac{u_{AA}}{2}A\right)A + \left(u_H + \frac{u_{HH}}{2}H\right)H + u_{AH}AH + M - p_c c - p_v v \\ & + \lambda\left(c + \varepsilon_A v - \delta_A A\right) + \mu\left(c + \varepsilon_H v + \omega A - \delta_H H\right)\end{aligned} \tag{34}$$

where λ and μ are the shadow prices of A and H, respectively.

The necessary conditions for an internal solution are

$$\mathcal{H}_c = 0 \quad \Leftrightarrow \quad \underbrace{u_c + u_{cc}c + u_{cA}A + u_{vc}v}_{\mathcal{V}_c} = p_c - \lambda - \mu \tag{35}$$

$$\mathcal{H}_v = 0 \quad \Leftrightarrow \quad \underbrace{u_v + u_{vv}v + u_{vA}A + u_{vc}c}_{\mathcal{V}_v} = p_v - \lambda\varepsilon_A - \mu\varepsilon_H \tag{36}$$

$$\dot{\lambda} = (\rho + \delta_A)\lambda - \omega\mu - \underbrace{(u_A + u_{AA}A + u_{cA}c + u_{vA}v + u_{AH}H)}_{\mathcal{V}_A} \tag{37}$$

$$\dot{\mu} = (\rho + \delta_H)\mu - \underbrace{(u_H + u_{HH}H + u_{AH}A)}_{\mathcal{V}_H} \tag{38}$$

$$\dot{A} = c + \varepsilon_A v - \delta_A A \tag{39}$$

$$\dot{H} = c + \varepsilon_H v + \omega A - \delta_H H \tag{40}$$

together with the appropriate transversality conditions. The left hand sides of the first order conditions 35 and 36 describe the instantaneous marginal utility of consuming c and v, respectively. The right hand sides describe the marginal cost of consuming, which depends on the market price as well as on the shadow prices of A and H.

The equation of motion of the shadow value of addiction (eq. 37) depends also on shadow value μ of health. Moreover, the marginal utility of addiction $\mathcal{V}_A$ directly depends on the addictiveness of c and v. In particular, u_{cA} and u_{vA} reduce shadow price of building up addiction, due to the fact that addictiveness increases the marginal utility of consuming the addictive goods. The law of motion of the shadow value of health harms (eq. 38), instead, does not depend on addiction nor on c or v.

Solving the focs for c and v yields the optimal consumption of the addictive good and of the harm reduction method

$$c^* = a_1 u_{vc} - a_2 u_{vv}; \qquad v^* = a_2 u_{vc} - a_1 u_{cc} \tag{41}$$

where

$$a_1 = \frac{u_v - p_v + u_{vA} A + \varepsilon_A \lambda + \varepsilon_H \mu}{u_{cc} u_{vv} - u_{vc}} \tag{42}$$

$$a_2 = \frac{u_c - p_c + u_{cA} A + \lambda + \mu}{u_{cc} u_{vv} - u_{vc}} \tag{43}$$

In the special case in which the harm reduction method does not affect the marginal utility of the addictive good, $u_{vc} = 0$, c^* does not depend on the price or the marginal utility of the harm reduction method and, conversely, v^* does not depend on p_c, nor on the marginal utility of c.

Replacing c^* and v^* in 37 to 40 allows derive the steady state values of λ, μ, A and H. Replacing in the expressions for c^* and v^*. and rearranging allows to write the steady state consumption of the harm reduction method as

$$v_d^{ss} = \gamma c^{ss} + \theta_v \left(u_v - p_v - \pi_v \right)$$

where

$$\gamma = \theta_v \left\{ \underbrace{\frac{\varepsilon_A}{\delta_A+\rho} u_{cA} + \frac{1}{\delta_A} u_{vA}}_{>0} + u_{vc} + \underbrace{\frac{\varepsilon_A}{\delta_A(\delta_A+\rho)} u_{AA} + \frac{(\delta_A+\omega)\left[\varepsilon_A\omega + \varepsilon_H(\delta_A+\rho)\right]}{\delta_A\delta_H(\delta_A+\rho)(\delta_H+\rho)} u_{HH}}_{<0} \right. \quad (44)$$

$$\left. + \underbrace{\left[\frac{\varepsilon_A}{\delta_H(\delta_A+\rho)} + \frac{\varepsilon_H}{\delta_A(\delta_H+\rho)} + \frac{\varepsilon_A(2\delta_H+\rho)\omega}{\delta_A\delta_H(\delta_A+\rho)(\delta_H+\rho)}\right] u_{AH}}_{<0} \right\} \quad (45)$$

$$\theta_v = \frac{-u_{cc}}{u_{cc}u_{vv} - u_{vc}} \frac{|J|}{|J_d|} > 0 \quad (46)$$

$$\pi_v = -\frac{\varepsilon_A}{\delta_A+\rho} u_A - \frac{\varepsilon_A\omega + \varepsilon_H(\delta_A+\rho)}{(\delta_A+\rho)(\delta_H+\rho)} u_H > 0 \quad (47)$$

and $|J_d|$ is the determinant of the Jacobian matrix at the steady state (not shown).

Conditional on $v_2^{ss} > 0$, steady state consumption of the original addictive good and health harm are

$$c_d^{ss} = c^{ss} + \theta_c (u_{vA} - r_H) v_d^{ss} \quad (48)$$

$$H_d^{ss} = H^{ss} + \theta_H (u_{vA} - r_L) v_d^{ss} \quad (49)$$

where

$$\theta_c = \frac{\alpha}{\delta_A+\rho} > 0 \quad (50)$$

$$\theta_H = \frac{\delta_A+\omega}{\delta_A\delta_H} > 0 \quad (51)$$

$$r_H = -(\delta_A+\rho) u_{vc} - \frac{(\delta_A+\rho)\varepsilon_A}{\delta_A} u_{cA} - \frac{\varepsilon_A}{\delta_A} u_{AA} - \frac{(\delta_A+\rho+\omega)(\omega\varepsilon_A + \delta_A\varepsilon_H)}{\delta_A\delta_H(\delta_H+\rho)} u_{HH} \quad (52)$$

$$+ \left[\frac{\varepsilon_A+\varepsilon_H}{\delta_H+\rho} + \frac{(\rho+2\omega)\varepsilon_A}{\delta_A(\delta_H+\rho)} + \frac{\rho(\omega\varepsilon_A + \delta_A\varepsilon_H)}{\delta_A\delta_H(\delta_H+\rho)}\right] u_{AH} \quad (53)$$

$$r_L = r_H - \frac{\omega\varepsilon_A + \delta_A\varepsilon_H}{\theta_c(\delta_A+\omega)} < r_H \quad (54)$$

Note that, conditional on the harm reduction method being used, the threshold r_H decreases if u_{vc} increases. In other words, the more the harm reduction method increases the marginal utility of the original addictive good, the more likely that c and H increase and the harm reduction policy fails to reach its objectives.

A.3 Taxation

The effect on steady state consumption of a change in own price (direct price effect) is

$$\frac{\partial v_d^{ss}}{\partial p_v} = \frac{u_{cc}}{(u_{cc}u_{vv} - u_{vc}^2)} \frac{|J|}{|J_d|} < 0 \tag{55}$$

$$\frac{\partial c_d^{ss}}{\partial p_c} = \frac{u_{vv}}{(u_{cc}u_{vv} - u_{vc}^2)} \frac{|J_v|}{|J_d|} < 0 \tag{56}$$

where $|J_v|$ is the determinant of the Jacobian matrix when the original addictive good is not available and the harm reduction method is instead available. Under asymptotic stability of steady state use of the harm reduction method, $|J_v| > 0$.

When considering cross-price effects, we obtain

$$\frac{\partial c_d^{ss}}{\partial p_v} = \frac{\delta_A \delta_H (\delta_H + \rho)}{(u_{cc}u_{vv} - u_{vc}^2)|J_d|} (r_H (u_{vc}) - u_{vA}) \tag{57}$$

$$\frac{\partial v_d^{ss}}{\partial p_c} = -\frac{\delta_A \delta_H (\delta_H + \rho)(\delta_A + \rho)}{(u_{cc}u_{vv} - u_{vc}^2)|J_d|} \gamma \tag{58}$$

Finally, the effect of a price change of the harm reduction method on health harm in the case of dual consumption is

$$\frac{\partial H_d^{ss}}{\partial p_v} = \frac{(\delta_A + \omega)(\delta_H + \rho)}{(u_{cc}u_{vv} - u_{vc}^2)|J_d|} (r_L (u_{vc}) - u_{vA}) \tag{59}$$

Outcome	Conditions	Intuition	Steady-state
Currently Consuming the Original Addictive Good ($c^{ss} > 0$)			
Irrelevant harm reduction method: consumption of the original addictive good does not change	$u_v \leq p_v + \pi_v - \frac{\gamma}{\theta_v} c^{ss}$	The harm reduction method is not appealing enough	$v_d^{ss} = 0$ c, H constant
Substitution: the harm reduction method replaces the original addictive good: Quitting the original addiction may result; Health harm may decrease	$u_v > p_v + \pi_v - \frac{\gamma}{\theta_v} c^{ss}$ (i) $u_{vA} < r_L$ (ii) $u_{vA} \in (r_L, r_H)$	The harm reduction method is appealing. If it is (i) mildly addictive, health harm decreases (ii) moderately addictive, health harm increases	$v_d^{ss} > 0$ $c \downarrow$ (i) $H \downarrow$ (ii) $H \uparrow$
Worst case scenario: Consumption of both addictive goods increases. Health harm increases	$u_v > p_v + \pi_v - \frac{\gamma}{\theta_v} c^{ss}$ $u_{vA} > r_H$	The harm reduction method is appealing and highly addictive	$v^{ss} > 0$ $c \uparrow$ $H \uparrow$
Currently Abstaining from Original Addictive Good ($c^{ss} = 0$)			
Irrelevant: No initiation, no gateway effect	$u_v \leq p_v + \pi_v$;	The harm reduction method is not appealing enough	$v_d^{ss} = 0$ $c_d^{ss} = H_d^{ss} = 0$
Initiation with the harm reduction method only; no initiation with the original addictive good	$u_v > p_v + \pi_v$ $u_{vA} < r_H$	The harm reduction method is appealing but not highly addictive	$v_d^{ss} > 0$ $c_d^{ss} = 0$ $H \uparrow$
Gateway effect: Initiation with both addictive goods	$u_v > p_v + \pi_v$ $u_{vA} > r_H$	The harm reduction method is appealing and highly addictive. It induces initiation also with the original addictive good	$v_d^{ss} > 0$ $c, H \uparrow$

Table A1: Possible changes in steady-state outcomes after the introduction of a harm reduction method.