Introduction of a Harm Reduction Method: When Does It Help and When Does it Backfire?*

John Cawley[†] Davide Dragone[‡]

Cornell University Università di Bologna

July 26, 2022

Abstract

Some harm reduction strategies encourage 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 addicts benefit because they can switch to a less harmful substance, but opponents argue that this could increase addiction and even encourage abstainers to begin using the addictive goods. 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; i.e. the conditions under which introducing a harm reduction method can lead to quitting the original addictive good, and the conditions under which it can lead previous abstainers to begin using the harm reduction method and even the original, more harmful, addictive good. Likewise, we demonstrate the conditions under which the introduction does in fact reduce health harms, and when it backfires and results in a worsening of health harms. The three key factors determining these outcomes 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 is helpful for understanding the consequences of harm reduction methods.

*For helpful comments, we thank Jesse Matheson and participants in the following seminars and conferences: the University of Duisburg-Essen seminar, the Essen Health Conference 2022, and the ASHEcon 2022 conference.

[†]Cornell University, Brooks School of Public Policy; 2312 MVR Hall, Cornell University, Ithaca NY 14853, USA. e-mail: jhc38@cornell.edu

[‡]Università di Bologna, Department of Economics, Piazza Scaravilli 2, 40126, Bologna, Italy; e-mail: davide.dragone@unibo.it

Keywords: addiction, harm reduction, initiation, moral hazard, Peltzman effect, risk compensation
JEL codes: I12, I18, D11

1 Introduction

Due to the substantial morbidity and mortality attributable to addictive behaviors such as 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 (Single, 1995; Erickson, 1995; Harm Reduction International, 2022).

Harm reduction methods have proven to be controversial, with advocates touting the potential health benefits to addicts, 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.

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 addicts cease consuming 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 needle exchanges, supervised injection facilities, legalized prostitution, condom distribution, Naloxone access laws, and Good Samaritan Laws (SAMHSA, 2022). For the purposes of this paper, we focus on harm reduction methods that have the following two properties: 1) 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 relievers. In both cases, the harm reduction method binds to similar opioid receptors as the

[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).

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), i.e. derived from the demand for elevated levels of neurotransmitters associated with feelings of pleasure and reward, such as dopamine. Harm reduction methods are substitutes for the original addictive good in the production of the release of those neurotransmitters. 2) The harm reduction method is less harmful to health than the original product. Obviously, if the product was more dangerous, it would not be a harm reduction method. Note that the harm reduction method may still be harmful to health, just not as harmful than the original product.

There are numerous examples of products that satisfy this definition of harm reduction. 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).

When the concern is opioid addiction, harm reduction methods include opioid agonist therapy, which uses methadone as a substitute for heroin or buprenorphine as a substitute for opioid pain relievers. 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 A1 provides examples of harm reduction for cigarettes, heroin, opioid pain relievers, and marijuana, explaining why the harm reduction methods are substitutes for the original addictive substance, and how they may reduce harm.

Harm reduction methods have proven controversial. Advocates see the following advantages: 1) a harm reduction method may increase quitting by addicts, because it allows them to gradually phase out their addiction. One could transition from the original addictive good to the harm reduction method to nothing at all, using the harm reduction method to alleviate the symptoms of withdrawal and gradually reduce one's addiction; 2) even if addicts never fully quit, and instead switch to the harm reduction method, that is still beneficial because the new addiction is less harmful to health than the original one. 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 by addicts. The rationale is: the very dangerousness of the original substance may motivate addicts to quit. However, if there is a substitute product that is less dangerous they may switch to that substitute and stay addicted rather than quit altogether. 2) Harm reduction methods may encourage new people to become addicted. Some abstainers may be abstaining precisely because the original substance is dangerous; introducing a product that is less dangerous 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 dangerous, substance in search of a bigger high. The absolute worst case is that these new users may be youth, who develop a lifelong addiction that could have been avoided.[2]

Both of these arguments 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 A1 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 prescription 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. 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 30 patients in the first year, and 100 patients in the second and subsequent years (Waters, 2019). As a result, 40% of U.S. counties have no waivered physicians who can prescribe buprenorphine to opioid addicts, and another 24% of counties have a patient capacity considered to be low (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).

[2] 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.

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 reduces or increases 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. The conditions for these different outcomes depend on three key drivers: 1) the enjoyableness, and 2) the addictiveness of the harm reduction method, and 3) the substitutability with the original addictive good.

This paper relates to several existing 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. Doleac and Mukherjee, 2022, Frio and França, 2021; Dave et al., 2019; Cotti et al., 2019; Simon et al., 2017; Margolis et al., 2014; Bhattacharya et al., 2011). 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. Pesko and Warman, 2022; Abouk et al., 2021; Allcott and Rafkin, 2020; Pesko et al., 2020; Cotti et al., 2020, Abouk et al., 2019; Marti et al., 2019; Pesko and Currie, 2019; Abouk and Adams, 2017; Friedman, 2015). Third, we contribute to the literature consisting of economic studies of methadone and buprenorphine (e.g. Doleac and Mukherjee, 2022; Allen et al., 2022; Barrette et al., 2021; Maclean et al., 2021; Abouk et al., 2019; Rees et al., 2019; Bishai et al., 2008).

2 A Model of Harm Reduction and Addictive Consumption

We begin with a model in which no harm reduction method is available and then introduce a harm reduction method. In both cases, we build on the two-stock model of rational addiction developed by Becker and Murphy (1988) and extended by Dockner and Feichtinger (1993). For clarity, we use the example of the original addictive good of combustible cigarettes (c) and the harm reduction method of e-cigarettes or vaping (v).

2.1 Benchmark: Addictive Consumption in the Absence of a Harm Reduction Method

We define c as the consumption of the addictive good (e.g. combustible cigarettes). Consumption of this addictive 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 the following equation of motion:

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

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. The equation of motion of H is:

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

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

As is common in the rational addiction literature, we assume that the instantaneous utility function is strictly concave, linear in the composite good q (to rule out income effects) and linear-quadratic. Specifically, the utility function is:

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

where u_A, u_H and $u_{cc}, u_{vv}, u_{AA,} u_{HH}$ are negative. Variable q represents a numeraire composite good. 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 c one has consumed in the recent past, the more enjoyment one gets from consuming c (or, put another way, the greater the withdrawal one experiences from not consuming, i.e. the greater the foregone marginal utility of c).[3]

As in the model of rational addiction, the consumer is assumed to be forward-looking, i.e. at least partially aware of the future consequences of consuming c for their addictive stock A and health harms H. Given a discount rate ρ, the intertemporal rational addiction problem of the agent is

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

subject to 1, 2, and the budget constraint

$$M\left(t\right) = q\left(t\right) + p_c c\left(t\right) \tag{5}$$

where M is income and p_c describes the monetary price of the addictive good, and it may include taxes on the addictive good.[4]

To determine the optimal quantity of addictive consumption, we apply the Pontryagin's maximum principle. 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 to obtain the optimal trajectory of state and costate variables leading to the steady state (see Appendix A.1 for details). Replacing back into the first order conditions yields the steady state level of consumption of the addictive good for this benchmark case:

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

[3]In the rational addiction literature, this key feature has also been linked to adjacent complementarity, reinforcement, and the serial correlation between past and present addictive consumption (see, e.g., Becker and Murphy, 1988; Becker et al., 1991). To avoid confusion, we refer to $u_{cA} > 0$ as the degree of addictiveness of the good.

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

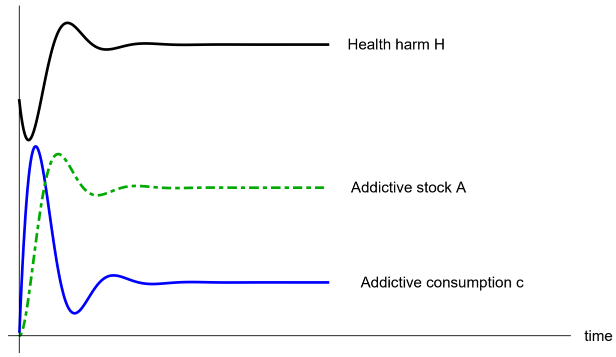

Figure 1: Possible trajectories of c, A and H for a consumer of the addictive good

where $\pi_c > 0$ describes the decrease in future utility resulting from consumption raising addiction and health harms (see eq. 29). Together, the monetary price p_c and the future consequences π_c of consumption represent the full (shadow) price of the addictive good c.

Since $\alpha > 0$ (see eq. 28), steady-state consumption of the addictive good will be positive only if $u_c > p_c + \pi_c$. That is, a person would only consume the addictive good 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 trajectory approaching the steady state, consumption at each point in time depends on current addiction and health harm according to

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

Since $a_c > 0$, for a given stock of addiction A and health harm H, $\hat{c}$ will be higher, the higher the steady-state consumption (which depends on its enjoyableness when the addictive stock is zero, i.e. u_c). 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). This implies that, even if steady state consumption is zero, it does not preclude the possibility that the person had earlier experimented with the addictive good.

2.2 A Model of 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 is cigarettes then v is vaping of e-cigarettes. The harm reduction method v adds to the addictive stock A and the stock of health harms H :

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

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

Given that the contribution of the original addictive good c to the stocks A and H is normalized to one, the harm reduction method v is assumed to contribute less to the addictive stock than the original good, $\varepsilon_A \in (0,1)$. 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, $\varepsilon_H \in (0,1)$ 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.

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, $u_v > 0$ describes the enjoyableness of the harm reduction method, absent previous and current consumption, and $u_{vA} > 0$ 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.[5]

Overall, the harm reduction method v affects individual consumption choices through two main channels. The first one concerns individual preferences, in that v is enjoyable $u_v > 0$,

[5]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). Our model is flexible and allows the harm reduction method to be an antagonist, as shown in Section 4. Here, for clarity and focus we assume it is a full or partial agonist rather than 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 or substitutes that is used to describe how the demand for one good responds to changes in the price of another good.

addictive ($u_{vA} > 0$) and affects the marginal utility of the original addictive good (u_{vc}). The second one is 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. These channels imply that use of the harm reduction method can affect consumption of the original addictive good c directly, through preferences, and indirectly through the accumulation of addiction experiences and health harms.

The steady state consumption of the harm reduction method is:[6]

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

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).

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 u_v of the harm reduction method, and the costs include the monetary costs of purchasing the harm reduction method and the future harmful consequences of consuming it, $p_v+\pi_v$. Due to the interdependence between the two addictive goods, in the long run the 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 (possibly zero) level of consumption of the original addictive good, is higher the greater the degree of addictiveness u_{vA} and u_{cA}. 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. 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.

The new steady-state level consumption of the addictive good and health harm is[7]

$$\begin{aligned} c_d^{ss} &= c^{ss} + \theta_c \left(u_{vA} - r_H \left(u_{vc}\right)\right) v_d^{ss} \quad (12) \\ H_d^{ss} &= H^{ss} + \theta_H \left(u_{vA} - r_L \left(u_{vc}\right)\right) v_d^{ss} \quad (13) \end{aligned}$$

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.

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

[7]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.

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 the original addictive good and the magnitude of health harms.

3 Consequences of the introduction of a harm reduction method

In this section we determine the consequences of the introduction of a harm reduction method v on initiation of the harm reduction method, which was previously unavailable, consumption of the original addictive good, and health harms.

Our results hinge on three key factors: 1) the enjoyableness of the harm reduction method (u_v); 2) the addictiveness of the harm reduction method (u_{vA}); and 3) the substitutability of the harm reduction method for the original addictive good (u_{vc}). For later reference, consider the following:

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* $u_{vA} > r_H(u_{vc})$.

We will organize the subsequent analysis by considering two categories of individuals. First, we consider consumers of the original addictive good and how they respond to the introduction of the harm reduction method. This category of consumers (e.g. adult smokers or opioid consumers) is a main target for policies that aim at reducing health harms by inducing substitution with the original addictive good until one eventually quits. However, it is also the category of consumers for which concerns about dual addictive consumption (e.g. smoking and vaping, heroin and methadone use) are raised by the opponents of harm reduction policies.

Then we consider individuals who were abstaining from consuming the addictive good before the harm reduction method is introduced. This category includes, e.g., non-smoking teenagers, which is a group of particular interest because opponents of harm reduction may fear that introduction of e-cigarettes may lead non-smoking teenagers to begin vaping and eventually to become smokers.[8]

[8] Interestingly, concerns that methadone might be a gateway for heroin use are not common. As shown later, according to our model this would be due to methadone being mildly addictive.

3.1 Do consumers use harm reduction methods?

Obviously, one critical factor for harm reduction policies to be relevant in the long run is whether individuals end-up consuming the harm reduction method. Based on 11, the following holds

Proposition 1 *An individual eventually uses the harm reduction method if*

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

If the individual is not using the harm reduction method because it is not sufficiently enjoyable (i.e. u_v is too low), 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 because consumption, addiction and health remain unaffected.

Whether enjoyableness u_v is "large enough", depends on the full price of the harm reduction method – which depends on the monetary price p_v and on the awareness of the future consequences of addictive harmful consumption π_v – and on the previous consumption (if any) of the original addictive good. Accordingly, using the harm reduction method is less likely the greater the full price of the harm reduction method, everything else equal.

If the sign of γ is positive, a previous experience of addictive consumption makes it more likely that expression 14 holds and the individual uses the harm reduction method. The reason lies in the fact that the difference between being a previous consumer or abstainer depends on the enjoyableness u_c of the original addictive good relative to its full price $p_c + \pi_c$. Previous consumers of the original addictive good have accumulated addictive stock A, which increases the desirability of the harm reduction method through u_{vA}. For this mechanism to occur, however, γ must be positive. As shown in equation 45 in the Appendix, this occurs when c and v are addictive enough, and/or sufficiently substitutes.

Remark 1 *Previous consumers of the original addictive good are more likely to eventually use the harm reduction method then previous abstainers if c and v are sufficiently addictive and substitutes*

If instead the original addictive good and harm reduction method are not sufficiently substitutes or addictive, the sign of γ is negative, which implies that, for a previous consumer of the addictive good, it is less likely the individuals uses the harm reduction method in the long run.

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

Table 1: 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 ($u_{vA}>r_H$).

3.2 Effects on consumption of the original addictive good: Gateway effect, substitution effects, and quitting

We now focus on scenarios in which the individual eventually decides to use the harm reduction method because condition 14 is satisfied.

We first consider how the harm reduction method affects consumption of the original addictive good c. Two factors are key: the degree of addictiveness u_{vA} of the harm reduction method, and the degree of substitutability with the original addictive good, as measured by u_{vc}.

The role of these factors is clear in expression 12, which can be conveniently rewritten to show the difference in steady-state consumption of the original addictive good c after the harm reduction method v becomes available:

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

Since θ_c is a positive parameter and $v_d^{ss}>0$, simple inspection of expression 15 allows to observe the following:

Proposition 2 *In the long run, consumption of the original addictive good c increases if the harm reduction method v is highly addictive, and it decreases otherwise*

Table 1 helps organizing the results for an individual already consuming the original addictive good prior to the introduction of a harm reduction method. Row 1 concerns the case where the individual uses the harm reduction method, and row 2 concerns the case where the individual does not use the harm reduction method, after it has been introduced. Columns 1, 2, and 3 concern the cases where the addictiveness of the harm reduction method is mild, moderate, and high (as defined in Definition 1 in Section 3).

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 of consumption 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 backfires: using v induces increased consumption of the original addictive good so that dual consumption and higher overall addiction result.

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

The bottom panel of Figure 2 shows a case in which consumption increases due to the harm reduction method being highly addictive. A particularly undesirable outcome results when the introduction of the harm reduction method leads abstainers to initiate consumption of the original addictive good. In this case the harm reduction method has been 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.

As shown in equation 52, the threshold r_H depends on the substitutability in preferences between the two addictive goods. In particular, $\partial r_H / \partial u_{vc} < 0$, hence the following remark holds:

Remark 3 *The greater the substitutability between the original addictive good and the harm reduction method (i.e. the lower u_{vc}), the higher the threshold level $r_H(u_{vc})$ for the harm reduction method to be classified as highly addictive.*

An implication of Remark 3 is that long-run consumption of the original addictive good is more likely to decrease when it is a substitute (in preferences) with the harm reduction method. In such a case, the harm reduction method is more likely to replace the original addictive good and dual consumption is less likely:

Remark 4 *Conditional on using a harm reduction method, in the long run the consumption of the original addictive good is more likely to decrease:*

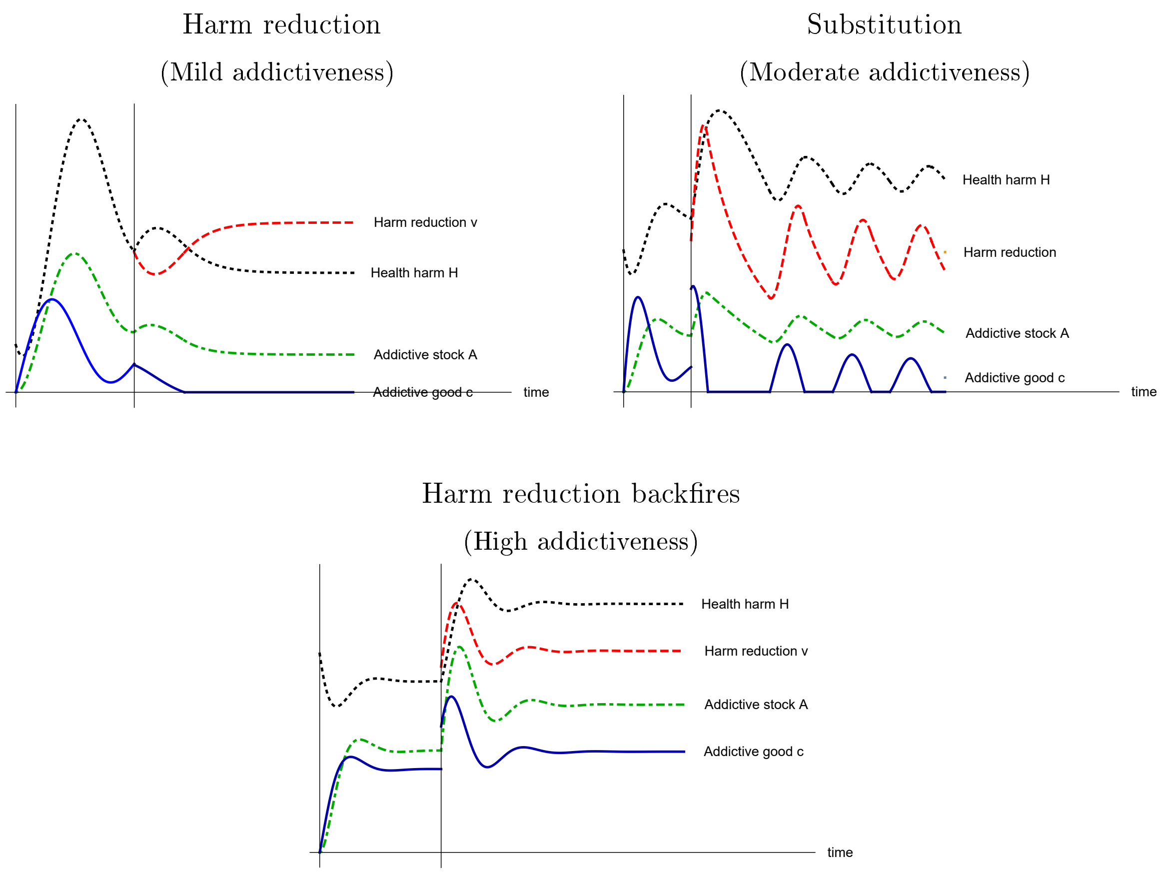

Figure 2: Possible trajectories of consumption, addiction and health harm for an individual already consuming the addictive good c before the harm reduction method is introduced. The vertical line denotes when the method becomes available. In all panels, the individual ends up using the harm reduction method because condition 14 is satisfied. Solid line: consumption of the original addictive good c; Dashed line: use of the harm reduction method v; Dotted line: health harm H; Dot-dashed green line: addictive stock A.
Top-left panel: the consumer quits the original addictive good and health harm is reduced. Top-right panel: the harm reduction method substitutes the original addictive good, health harm increases. Bottom panel: addictive consumption and health harm increase.

- *The smaller the addictiveness of the harm reduction method (i.e. u_{vA} small);*

- *The greater the substitutability between the two addictive goods (i.e. u_{vc} small)*

On the contrary, if the harm reduction method is not highly addictive, or it is a complement to the original addictive good (so that consuming one good enhances the marginal utility of the other good, i.e. $u_{vc} > 0$, which is not the case we are considering in this section), joint (dual) consumption is more likely.

If the harm reduction method is less than highly addictive, then consumption of the original addictive good will fall (i.e. $c_d^{ss} < c^{ss}$). The top panels of Figure 2 illustrate this case when v

is mildly or moderately addictive. In both graphs, consumption of the original addictive good (solid line) tends to decrease over time, eventually with quits and relapses (top-right panel). Over time, the harm reduction method has replaced consumption of the original addictive good, possibly with a reduction in the addiction stock.

Remark 5 (Substitution effects) *If the harm reduction method is not highly addictive, i.e. it is mildly or moderately addictive, its use reduces consumption of the original addictive good*

Importantly, the harm reduction method, if not highly addictive, can lead the individual to not only reduce consumption of the original addictive good but quit it altogether ($c^{ss} > c_d^{ss} = 0$). Based on equation 15 we can state:

Remark 6 (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 to quit for a heavy smoker than for a light smoker. The reason is that the level of consumption of the original addictive good (e.g. smoking) depends, absent the harm reduction method, on how much it is enjoyable, i.e. on the level of u_c. Heavy smokers find smoking particularly enjoyable, and thus quitting is particularly difficult. Condition (b) shows that, if the policy goal is a reduction in the consumption of the original addictive good (which includes increasing quits of the original addictive good), then a less addictive harm reduction method is more effective than a more addictive one. 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. In our example, the higher the quantity of vaping, the higher the substitution effect with smoking and the associated likelihood of quitting.

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 eventually goes to zero. Importantly, the success of a harm reduction approach depends critically on when one examines outcomes – in

Harm reduction method used?	Addictiveness of harm reduction method Mild/Moderate	 High
Yes	**Initiation with the harm reduction method** $c_d^{ss} = 0, H \uparrow$	**Gateway effect: Use of addictive good and harm reduction method** $c \uparrow, H \uparrow$
No	**Harm reduction method is irrelevant** $v_d^{ss} = c_d^{ss} = H_d^{ss} = 0$	

Table 2: 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.

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 to zero.

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

In the previous sections we have determined the conditions under which the harm reduction method is used. When this is the case, consumption of the original addictive good will increase if the harm reduction method is highly addictive, a worst-case scenario in which the individual becomes a dual consumer of addictive good. We have also shown that, if the harm reduction method is mildly or moderately addictive, then the harm reduction method can substitute for the original addictive good, eventually leading to quit. Although these results are relevant for understanding the drivers of addictive consumption in presence of a harm reduction method, they do not address a related concern, namely: when do harm reduction policies decrease or increase health harms?

For individuals that were abstaining from the addictive good (e.g. non smokers, non opioid users), the answer is intuitive: if they began using the harm reduction method, health harms worsen. The reason is that the harm reduction method is assumed to be harmful (although less than the original good), which clearly does no good to an individual previously abstaining from addictive consumption. Moreover, if the harm reduction method is highly addictive it is possible that use of the harm reduction method acts as a gateway drug inducing initiation with the addictive good (see Table 2). This is clearly the worst case scenario, with dual addictive consumption and worse health.

Remark 7 *If a previous abstainer of the original addictive good uses the harm reduction method, health harm increases*

It is worth emphasizing that harm reduction policies are typically not meant for abstainers, but for individuals who are already consuming an addictive good. To address under what conditions harm reduction policies can decrease or increase health harms for this category of people, rewrite equation 13 in terms of long-run change in health as 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 3 *After the introduction of a harm reduction method, health harm eventually decreases if the harm reduction method is mildly addictive, and it increases if the harm reduction method is moderately or highly addictive*

The main implication of Proposition 3 is that harm reduction policies have both the potential to decrease or increase health harms. The key factors are: 1) the addictiveness of the harm reduction method; and 2) 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 1). Moreover, as shown in the previous section, for a previous consumer of the original addictive good c, also consumption of c 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, top-left 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 (while still being less harmful than the original addictive good). 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 on the other hand it has failed in that it has worsened the health of those who were previously using the addictive good.

Finally, if the harm reduction is highly addictive, not only health harm increases, but also consumption of the original addictive good. In such a case, the harm reduction policy is unambiguously a failure. We conclude that, if the harm reduction method is highly addictive, the introduction of a harm reduction method backfires, in the sense that not only might health harms worsen, not only might previous consumers of the addictive good decline to quit, they may actually increase their consumption of the addictive good.

Note that threshold $r_L(u_{vc})$ responds to the substitutability in preferences between the two addictive goods in a similar way to the threshold $r_H(u_{vc})$ that defines the change in consumption of the addictive good. In particular, $\frac{\partial r_L}{\partial u_{vc}} < 0$ and the following remark holds:

Remark 8 *The greater the substitutability in preferences between the original addictive good and the harm reduction method (i.e. the lower u_{vc}), the higher the threshold level $r_L(u_{vc})$.*

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, factors (such as the substitutability in preferences) that incentivize using the former instead of the latter, will likely produce a reduction in health harm. On the contrary, factors that reinforce the joint consumption of the two addictive goods (e.g. u_{vc} being positive) will likely produce more health harm.

Once again, it is critical when one observes the onset of health harms. If one examined early time periods, it may be impossible to distinguish early experimentation with the original addictive good followed by quitting, from a steady-state positive consumption of the addictive good. Accordingly, a short-run evaluation of the harm reduction policy may give different insights with respect to an evaluation carried over longer time-periods.

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.

Direct price effects are negative: $\partial v_d^{ss}/\partial p_v < 0$. When considering cross-price effects, we obtain

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

That is, when the harm reduction method is mildly or moderately addictive, i.e. $u_{vA} < r_H(u_{vc})$, 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 this is indeed the case for vaping and smoking (see, e.g. Pesko and Warman, 2022; Pesko et al., 2020; Pesko and Currie, 2019), 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.

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

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

That is, an increase in the price of the harm reduction method increases health harm if the harm reduction method is mildly addictive, i.e. $u_{vA} < r_L(u_{vc})$. 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 9 *Taxing the harm reduction method:*

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

An implication of the above Remark is that, in the intermediate case in which the harm reduction method is moderately addictive, $u_{vA} \in (r_L(u_{vc}), r_H(u_{vc}))$, 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 compensates for the increased harm of greater smoking. If the harm reduction method is highly addictive then we know from the earlier results summarized in Table 1 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 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 addictive ($\varepsilon_A = u_{vA} = 0$), nor enjoyable ($u_v = 0$).

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 would deliver no benefits. The case of a previous user is more interesting, because antagonist methods imply that term γ in equation 14 is strictly negative. Since, for any level of consumption of the original addictive good,

$$0 < p_v + \pi_v - \frac{\gamma}{\theta_v} c^{ss}, \tag{19}$$

a previous consumer of the original addictive good (who rationally consumes the good by balancing its benefits and costs) would not rationally use the antagonist method in the long run (even if it were for free, i.e with $p_v = 0$). The intuition is simple: since the individual enjoys consuming the addictive good (despite being harmful and addictive), it is not rational to use another good whose only effect is that of reducing the benefits of addictive consumption.

This result is at odds with the evidence showing a demand for antagonists as a method to reduce consumption of the addictive good. One possible explanation relies on the use of the antagonist method as a precommitment device to be used before the temptation of consuming the addictive good arises. This explanation relies on the notion of addictive consumption as a result of self-control failure, and of the use of precommitment methods that "bind the hands" of an individual that wants to avoid falling into the temptation of addictive consumption (see, e.g., Laibson, 1997; Thaler and Shefrin, 1981; Schelling, 1978). Accordingly, evidence on the demand for antagonists by individuals consuming addictive goods can be interpreted as evidence of sophisticated consumers being prone to time-inconsistent preferences and self-control failures (Gruber and Köszegi, 2001; Strotz, 1955).

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 no use of 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 Discussion

Harm reduction methods are controversial. Advocates argue that they can help addicts quit and, even if not, will still reduce health harms. Opponents argue that there is a risk of moral hazard – that introducing a harm reduction method may make addicts less likely to quit and may lead them to become more addicted, 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 see people quitting the original addictive product. 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 improved with time, as people decreased their consumption over time as the health consequences mounted. Thus, it may take time to determine whether the harm reduction method has led to short-term experimentation that fizzles out or a steady-state increase in 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 users 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. One can even apply it to less addictive substances, such as diet colas as a substitute for caloric colas (which are somewhat addictive due to caffeine).

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 alternatively more selectively regulate access to the harm reduction method. For example, 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 u_{vA} of the harm reduction method. This factor turns out to be critical in determining what happens to consumption of the original addictive good in our model. 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 of course tax either the original good or the harm reduction methods in order to raise their monetary price and decrease demand for them. As noted in the Introduction, 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).

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. These factors determine whether the strongly divergent predictions of advocates or opponents turn out to be correct.

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.

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.

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.

Allcott, H. and C. Rafkin (2020). Optimal regulation of e-cigarettes: Theory and evidence. NBER working paper 27000, National Bureau of Economic Research.

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

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.

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.

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

Bhattacharya, J., M. K. Bundorf, N. Pace, and N. Sood (2011). Does health insurance make you fat? pp. 35. University of Chicago Press.

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.

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 2022-07-07.

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.

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.

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.

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

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

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

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

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.

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*.

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

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

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

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

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.

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.

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

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.

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

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.

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.

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.

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.

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 2022-07-07.

SAMHSA (2022). Harm reduction. https://www.samhsa.gov/find-help/harm-reduction, last accessed 2022-07-07.

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

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

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.

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

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

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

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-c last accessed 2022-07-07.

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 2022-07-07.

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 2022-07-07.

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 2022-07-07.

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 2022-07-07.

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

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.

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

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

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

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

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

A Appendix

A.1 One good

Replacing q from the budget constraint, the Hamiltonian function associated to the consumer's problem is

$$\begin{aligned}\mathcal{H}(c, v, q; A, H) &= \left(u_c + \frac{u_{cc}}{2}c + u_{cA}A\right)c + M - p_c c \\ &+ \left(u_A + \frac{u_{AA}}{2}A\right)A + \left(u_H + \frac{u_{HH}}{2}H\right)H \\ &+ \lambda\left(c - \delta_A A\right) + \mu\left(c + \omega A - \delta_H H\right)\end{aligned} \tag{20}$$

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

The necessary conditions for an internal solution are

$$\mathcal{H}_c = 0 \quad \Leftrightarrow \quad \underbrace{u_c + u_{cc}c + u_{cA}A}_{\mathcal{U}_c} = p_c - \lambda - \mu \tag{21}$$

$$\dot{\lambda} = (\rho + \delta_A)\lambda - \omega\mu - \underbrace{(u_A + u_{AA}A + u_{cA}c)}_{\mathcal{U}_A} \tag{22}$$

$$\dot{\mu} = (\rho + \delta_H)\mu - \underbrace{(u_H + u_{HH}H)}_{\mathcal{U}_H} \tag{23}$$

$$\dot{A} = c - \delta_A A \tag{24}$$

$$\dot{H} = c + \omega A - \delta_H H \tag{25}$$

together with the appropriate initial and transversality conditions. The first order condition 21 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. 22) 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. 23), instead, does not depend on addiction nor on c.

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

$$c^* = \frac{u_c - p_c + u_{cA}A + \lambda + \mu}{-u_{cc}} \tag{26}$$

Replacing in 21 to 25 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 26 gives the steady state consumption of c:

$$c^{ss} = \alpha \left(u_c - p_c - \pi_c \right) \tag{27}$$

where

$$\alpha = \frac{\delta_A \delta_H \left(\delta_A + \rho \right) \left(\delta_H + \rho \right)}{-u_{cc} |J|} > 0 \tag{28}$$

$|J|$ is the determinant of the Jacobian matrix (not shown) computed at the steady state,[9] and

$$\pi_c = -\frac{1}{\delta_A + \rho} u_A - \frac{\delta_A + \omega_A + \rho}{\left(\delta_A + \rho \right) \left(\delta_H + \rho \right)} u_H > 0 \tag{29}$$

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

$$H^{ss} = \frac{\delta_A + \omega_A}{\delta_A \delta_H} c^{ss}, \qquad A^{ss} = \frac{1}{\delta_A} c^{ss} \tag{30}$$

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 21 into the system of differential equations 22–25, 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

$$\hat{c}\left(A, H \right) = a_c c^{ss} + a_A A + a_H H \tag{31}$$

where

$$\alpha_c = \frac{e_1 e_2}{\delta_A \delta_H} > 0 \tag{32}$$

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

$$\alpha_H = -\frac{\left(\delta_H + e_1 \right) \left(\delta_H + e_2 \right)}{\delta_A - \delta_H + \omega} \tag{34}$$

[9] 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.

and $e_1, e_2 < 0$ are the eigenvalues with negative real parts associated to the Jacobian matrix of 22 to 25. 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

$$\begin{aligned}\mathcal{H}(c,v,q;A,H) = &\left(u_c + \frac{u_{cc}}{2}c + u_{cA}A\right)c + \left(u_v + \frac{u_{vv}}{2}v + u_{vc}c + u_{vA}A\right)v \\ &+ \left(u_A + \frac{u_{AA}}{2}A\right)A + \left(u_H + \frac{u_{HH}}{2}H\right)H + M - p_c c - p_v v \\ &+ \lambda\left(c + \varepsilon_A v - \delta_A A\right) + \mu\left(c + \varepsilon_H v + \omega A - \delta_H H\right)\end{aligned} \tag{35}$$

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

The necessary conditions for an internal solution are

$$\mathcal{H}_c = 0 \quad \Leftrightarrow \quad \underbrace{u_c + u_{cc}c + u_{cA}A + u_{vc}v}_{\mathcal{V}_c} = p_c - \lambda - \mu \tag{36}$$

$$\mathcal{H}_v = 0 \quad \Leftrightarrow \quad \underbrace{u_v + u_{vv}v + u_{vA}A + u_{vc}c}_{\mathcal{V}_v} = p_v - \lambda\varepsilon_A - \mu\varepsilon_H \tag{37}$$

$$\dot{\lambda} = (\rho + \delta_A)\lambda - \omega\mu - \underbrace{(u_A + u_{AA}A + u_{cA}c + u_{vA}v)}_{\mathcal{V}_A} \tag{38}$$

$$\dot{\mu} = (\rho + \delta_H)\mu - \underbrace{(u_H + u_{HH}H)}_{\mathcal{V}_H} \tag{39}$$

$$\dot{A} = c + \varepsilon_A v - \delta_A A \tag{40}$$

$$\dot{H} = c + \varepsilon_H v + \omega A - \delta_H H \tag{41}$$

together with the appropriate transversality conditions. The left hand sides of the first order conditions 36 and 37 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. 38) 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, do 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. 39), 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

$$c^* = a_1 u_{vc} - a_2 u_{vv}; \qquad v^* = a_2 u_{vc} - a_1 u_{cc} \tag{42}$$

where

$$a_1 = \frac{u_v - p_v + u_{vA} A + \varepsilon_A \lambda + \varepsilon_H \mu}{u_{cc} u_{vv} - u_{vc}} \tag{43}$$

$$a_2 = \frac{u_c - p_c + u_{cA} A + \lambda + \mu}{u_{cc} u_{vv} - u_{vc}} \tag{44}$$

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 38 to 41 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

$$\gamma = \theta_v \left\{ \underbrace{\frac{\varepsilon_A}{\delta_A + \rho} u_{cA} + \frac{1}{\delta_A} u_{vA}}_{>0} + u_{vc} + \underbrace{\frac{\varepsilon_A}{\delta_A \left(\delta_A + \rho\right)} u_{AA} + \frac{\left(\delta_A + \omega\right)\left[\varepsilon_A \omega + \varepsilon_H \left(\delta_A + \rho\right)\right]}{\delta_A \delta_H \left(\delta_A + \rho\right)\left(\delta_H + \rho\right)} u_{HH}}_{<0} \right\} \tag{45}$$

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

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

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

$$c_2^{ss} = c^{ss} + \theta_c \left(u_{vA} - r_H \right) v_2^{ss} \tag{48}$$

$$H_2^{ss} = H^{ss} + \theta_H \left(u_{vA} - r_L \right) v_2^{ss} \tag{49}$$

where

$$\theta_c = \frac{\alpha}{\delta_A + \rho} > 0 \tag{50}$$

$$\theta_H = \frac{\delta_A + \omega}{\delta_A \delta_H} > 0 \tag{51}$$

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

$$r_L = r_H - \frac{\omega \varepsilon_A + \delta_A \varepsilon_H}{\theta_c \left(\delta_A + \omega\right)} < r_H \tag{53}$$

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

$$\frac{\partial v_d^{ss}}{\partial p_v} = \frac{u_{cc}}{\left(u_{cc} u_{vv} - u_{vc}^2\right)} \frac{|J|}{|J_d|} < 0 \tag{54}$$

$$\frac{\partial c_d^{ss}}{\partial p_c} = \frac{u_{vv}}{\left(u_{cc} u_{vv} - u_{vc}^2\right)} \frac{|J_v|}{|J_d|} < 0 \tag{55}$$

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

$$\frac{\partial c_d^{ss}}{\partial p_v} = \frac{\delta_A \delta_H \left(\delta_H + \rho\right)}{\left(u_{cc} u_{vv} - u_{vc}^2\right) |J_d|} \left(r_H \left(u_{vc}\right) - u_{vA}\right) \tag{56}$$

$$\frac{\partial v_d^{ss}}{\partial p_c} = -\frac{\delta_A \delta_H \left(\delta_H + \rho\right)\left(\delta_A + \rho\right)}{\left(u_{cc} u_{vv} - u_{vc}^2\right) |J_d|} \gamma \tag{57}$$

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

$$\frac{\partial H_d^{ss}}{\partial p_v} = \frac{\left(\delta_A + \omega\right)\left(\delta_H + \rho\right)}{\left(u_{cc} u_{vv} - u_{vc}^2\right) |J_d|} \left(r_L \left(u_{vc}\right) - u_{vA}\right) \tag{58}$$

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

Table A1: Examples of Harm Reduction Methods

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 A2: Possible changes in steady-state outcomes after the introduction of a harm reduction method.