RWJF Scholars in Health Policy Research Program

Some Basic Economics of Tobacco Harm Reduction

Carl V. Phillips

June 2013
Working Paper Series

WP-51

Some basic economics of tobacco harm reduction

Carl V. Phillips, PhD
Scientific Director, Consumer Advocates for Smoke-free Alternatives Association
cphillips@casaa.org

Working Paper version 4 June 13

Comments welcome (about the accuracy of the analysis, the clarity of exposition for the
target audiences, or anything else).

Abstract
Tobacco harm reduction (THR) – the substitution of low-risk tobacco/nicotine products
for smoking and efforts to encourage such substitution – cannot be understood without
understanding the motives and beliefs of consumers and how these affect action. Yet the
discussion of THR has largely taken place without consideration of basic welfare
economics. This has led to uncertainty about points that are actually rather obvious, as
well as conclusions that are easily shown to be wrong. This paper offers and introductory
overview of some of the basic consumer economics that provides insights into THR,
attempting to present these in a form that is accessible to readers unfamiliar with
economics. Among the conclusions from this simple analysis are that THR offers large
improvements in welfare to smokers compared to either smoking or abstinence; that a
recognition that alternative products are low-risk will lead to an increase in total
tobacco/nicotine use, and that the new users will experience an increase in welfare; and
that recent declarations about the optimal relative tax rates for different products are
wrong for any plausible objective.

Introduction
There is remarkably little understanding of what basic welfare economics[1] tells us about
the consumption of cigarettes and other tobacco/nicotine products. To understand the
behavior of populations when making choices, it is not necessary to employ the advanced
math that fills economics journals, but some basic arithmetic and high-school level
economics is needed. The absence of such understanding from discussions tobacco harm
reduction (THR) – the substitution of low-risk tobacco/nicotine products for cigarettes –
is unfortunate because almost all positive or normative analysis requires such an
understanding. The stakes associated with misunderstanding are great: Tobacco harm

---

[1] Welfare economics, which can also be called consumer economics, is the study of costs
and benefits of consumption decisions (and thus, the resulting impact on welfare – or to
use the economics jargon, on "utility") at the individual level, as well as the associated
incentives and decisions.

reduction offers the greatest unrealized potential improvement in public health in developed countries today, and yet the policies discussions surrounding it are generally quite naïve.

The failure to understand the economics is caused not just by a lack of familiarity, but by active efforts by anti-tobacco activists to create a belief that using tobacco/nicotine products[2] is somehow fundamentally different from other consumer choices. Most analyses of tobacco use behavior imply that such consumption involves no consumer volition or welfare, but rather is some kind of involuntary tic or perhaps a possession by evil spirits. This engineered misperception is made worse by a lack of understanding about the flexibility of welfare economic models. While there definitely are some complications that are absent for most other products, few of them cannot be captured in a welfare economics model.

Even a basic understanding of the economics informs several common questions which seem to baffle most commentators. A deeper exploration could also help answer more complicated questions that have been largely ignored because it was not possible to conceptualize them without a useful model of human behavior. The purpose of this paper is to provide the basic outline of the simple consumer economics, primarily to help readers with an interest in THR to gain a general understanding of the concepts and to lay the groundwork future analyses. A few specific questions are answered as a basis for organizing the general exposition. There is nothing novel about the analytics (though there are a few departures from the standard economic models where the standard simplification is clearly not useful); anyone capable of teaching applied microeconomics and familiar with the subject matter could present this, and the basic points can be found in an introductory textbook. But the basic overview is missing from the literature on tobacco harm reduction.

As part of the goal of introducing the economics to those who might have limited familiarity with many of the concepts, footnotes are included to explain some of the economic terms and concepts and to expand upon a few complicated points that might distract from the basic presentation.

Common questions that can be addressed using the framework presented here include:

1. What can we say about the net benefit to smokers who adopt tobacco harm reduction, and how does it compare to the welfare from either smoking or abstinence?

---

[2] The construction "tobacco/nicotine products" is used to collect consumer products that contain tobacco leaf (cigarettes, smokeless tobacco) along with products that put the nicotine extracted from tobacco in another substrate (e-cigarettes, pharmaceutical nicotine products (NRT)). Because there is debate about whether the latter category should be included in the simpler phrase, "tobacco products", the more awkward terminology has developed.

2. What will happen to total consumption prevalence when low-risk alternatives to cigarettes become established, and will people who never smoked be inclined to initiate use of these products?
3. What is the optimal excise tax ("sin tax") structure on the various categories of products?
4. Can we predict the social dynamics of the market shifting from cigarettes to low-risk alternatives?
5. How can we assess the incentives for black or grey market demand and its effects on the legal market?

The first of these is illustrated by the model, though it is not a product of the theory alone (in particular, it requires introducing what we know about the low risk of alternative products). The second is a simple result of the model, and the mere act of illustrating the concepts answers the question. The third is used as the organizing principle for the latter part of this exposition. The fourth is addressed in other ongoing social dynamics research (e.g., http://ep-ology.blogspot.com/2012/10/agent-based-model-of-thr-adoption-and.html). The fifth is widely discussed, with vague allusions to the welfare economics, but the analyses generally appear to be faulty due to a failure to incorporate consumer incentives.

For practical purposes, the market for tobacco/nicotine products consists of two types of products: (1) Cigarettes, the currently dominant product, which substantially hasten the deaths of about a third of lifelong users, and also create substantial disability and non-fatal illnesses; (2) Low-risk smoke-free alternatives, including Western-style smokeless tobacco products, inhaled nicotine products (electronic cigarettes (e-cigarettes) and a few emerging variations), and pharmaceutical-style nicotine products (nicotine gum, patches, lozenges). There is a lot of noise about comparisons of risk among the different low-risk product categories, but for practical purposes they are equal, with a health risk that is estimated to be about 99% less lower than cigarettes (Phillips et al. 2006). Because of this rough equivalence, the present analysis just simplifies this to a collective category of "low-risk products".

While there are products that are expected to fall somewhere in between low-risk products and cigarettes, none are in widespread use, so there are basically just two points, not a distribution across the range (convex combinations[3] of the two options are beyond the scope of the present analysis). Variations in risks among different varieties of cigarettes are undoubtedly substantially greater than the differences among low risk products and between those risks and zero. But since it is not entirely clear which

---

[3] Convex combination is the economics jargon for a total that is divided between two choices, so in this case it means that some portion, x, of the total consumption is cigarettes while the remaining 1-x is the low-risk product. Anti-tobacco activists have dubbed this "dual use", with the suggestion that it is somehow substantially worse than simply the reduced level of smoking that it represents, though there is no reason to believe that this is the case.

varieties are lower risk, and the risks are sufficiently similar for present purposes, this is ignored.[4]

These product categories are substitutes, in both the common-language and economic senses of that word. Increases in the popularity of low-risk products come mostly at the expense of cigarettes though, as discussed below, not entirely. The more attractive one product is – which for the present analysis is simplified to two features, purchase price and health risk – the more substitution from would-be users of the other product that will occur. Almost all users of one of these products are more inclined than nonusers to use another, though some would prefer abstinence to switching.[5]

All the products deliver nicotine which is generally (though not universally) identified as the most important source of benefits from use. They are not perfect substitutes (though for simplicity, this analysis starts by assuming they are). Due to a combination of real physical advantages and history, cigarettes are typically considered the optimal product (ignoring the health costs), and other products imperfectly mimic the delivery of psychoactive chemicals, the speed of delivery, or the physical and social ritual. However, the low-risk products also have advantages beyond lowering the risk for serious disease, including a reduction in acute health effects, reduced impact on others, and a wider variety of flavors and other features.

Unlike many areas of harm reduction, the risk from the low-risk alternatives is so low that substitution is practically the same as abstinence. The low risk means that a lifetime of use of a substitute product poses a similar risk for the average smoker as continuing to

---

[4] Aside: Very little is known about the comparative risks of different cigarettes because the party line from anti-tobacco activists is that all cigarettes are equally harmful, and it would be political suicide for anyone in the "public health" establishment to work on that topic. Ironically, in spite of this, many anti-tobacco activists have embraced policy recommendations that would mandate changes in cigarette chemistry, despite the lack of knowledge about what could make a cigarette lower risk.
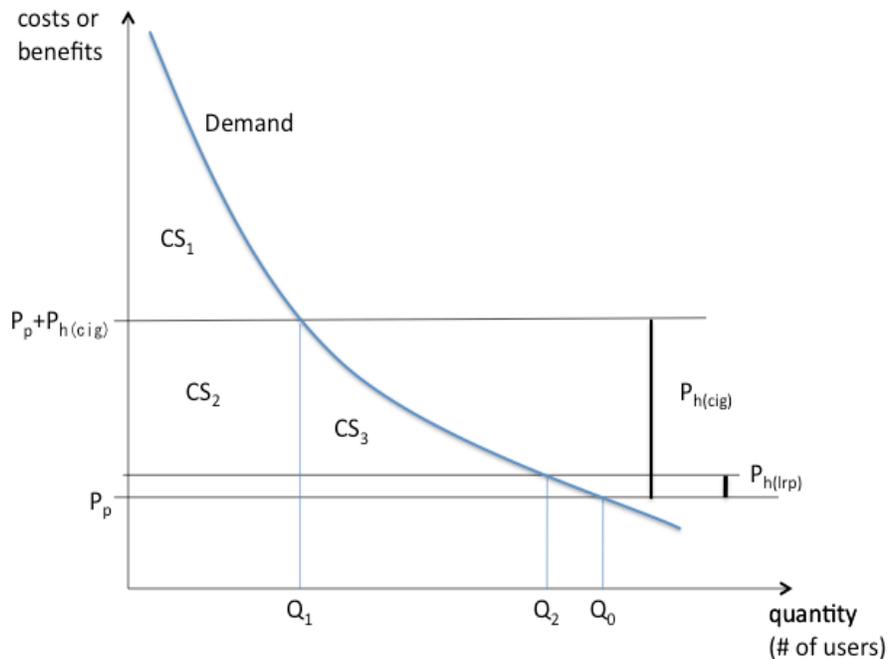
[5] Aside: It is not captured in the present analysis, it is a simple extension to demonstrate some implications of this. Since users of any tobacco/nicotine product have demonstrated a preference for use over abstinence, they are more likely to also prefer another tobacco/nicotine product to abstinence. A popular claim in anti-harm-reduction activism is that low-risk tobacco/nicotine products are a "gateway" to smoking. To the extent that any evidence is cited in support of this claim, it is simply the fact that American smokeless tobacco users were historically somewhat more likely to take up smoking later in life than were lifelong abstainers (in the studied populations, there was little awareness that smokeless tobacco was low-risk, so those who switched probably did not realize they were making a very unhealthy switch). But by that logic, every brand of cigarettes is a "gateway" to other brands, and so the existence of every single brand increases smoking. Obviously the observed data is the inevitable result of some people liking tobacco and others not liking it, and in no way suggests that the use of the smokeless product caused smoking that would not have otherwise occurred.

smoke for about two more months, even if complete abstinence were achieved after that (Phillips, 2009). Still, there is resistance and controversy about tobacco harm reduction, and since there is no consequential uncertainty about the health science, most of the genuine confusion surrounds questions of consumer behavior. Some relatively simple economics resolves some of the common controversies.

Preference, purchase price, and health: some basic implications
Some who are not familiar with basic economic models may believe that the mere presence of health costs renders consumer economics inapplicable, but this is not true. Even the simplest model allows costs to consumers to exceed the purchase price. Figure 1 is a simple supply curve model of consumption of a single tobacco product, cigarettes.

Figure 1

costs or benefits

Demand

CS_1

$P_p + P_{h(cig)}$

CS_2

$P_{h(cig)}$

CS_3

$P_{h(lrp)}$

P_p

Q_1 Q_2 Q_0 quantity (# of users)

It is easiest to think of quantity (q) in this analysis as the number of people in the population who use the product. Typically the q-axis is conceptualized as total quantity consumed, ignoring whether this is a few people consuming a lot or everyone consuming a bit. But in this case it is more useful to abstract out variations in individual consumptions and just count the number of users.

The downward sloping curve, D, is the demand curve that relates the total cost to the number of people choosing to consume.[6] The p-axis measures the total gross benefits that each consumer gets from consuming the product, which is equivalent to the maximum cost each would be willing to pay (if the cost were any higher than that, then the net benefits would be negative and the consumer would choose to forgo the product). Thus, any point on the demand curve can be interpreted as either the gross benefit to one particular consumer from consuming the product (with the net benefit then being calculated by subtracting the cost) or the cost level at which the consumer would choose to forgo the product. The monotonic downward slope exists by construction – that is, it is created by ordering the consumers in descending willingness to pay. The shape illustrated for the demand curve is stylized, with the actual shape determined by real-world preferences.[7]

The horizontal lines represent the costs of consumption faced by each consumer. The quantity where the top horizontal line (the total cost) crosses the demand curve divides the population into those on the left-hand side who get net benefit and choose to consume and those on the right-hand side for whom cost exceeds benefit, who choose to not consume.

All consumers face the same purchase price, no matter what they are willing to pay.[8] Adding in health costs in this simplified model requires assuming that all consumers face the same health risks, as well as the more significant assumptions that all consumers perceive the risks to be the same and that their tradeoff between money and health costs

---

[6] Some readers will be familiar with versions of the figure that include a supply curve (basically the cost of production) along with the demand curve. This is used, for example, to illustrate the market equilibrium where they cross, where the "invisible hand" results in the right amount of production and consumption to maximize net welfare (assuming there are no costs not reflected in those curves). Because of such applications of the graph, the vertical axis is generally labeled p, for price, though it also measures all costs and benefits, which is its role in the present analysis.

[7] The curve that is used here is probably more realistic than the straight line that is often drawn. Indeed, the curve drawn may be too close to a straight line. It illustrates that there are a few people to the left who get enormously more benefits from the product than others do, and that there is a long tail of people who get a little benefit, but not much. As will become apparent in the analysis, the former are unlikely to choose abstinence, even if the cost is driven very high, while the latter are likely to have already quit or never to have started using the product.

[8] Suppliers would generally like to be able to "price discriminate", to charge consumers with high willingness to pay a lot while keeping prices low for others to avoid losing the sale, and thereby capture more profit at the expense of consumers. Service providers (e.g., airlines, telecom) have various ways to move toward this goal, but sellers of retail commodities that anyone can buy at a store for the same price lack this options and are largely forced to try to persuade those who are willing to pay more will buy higher-profit premium varieties.

is the same. [9] The assumption that the perception of risk is accurate is relaxed below, but departing from homogeneity across the population requires individual-level models rather than population average models. Costs and benefits resulting from personal interaction and social acceptability also cannot be usefully modeled at the population average level, they are omitted from this analysis. [10]

The failures to mention concepts like "addiction", which some readers might already be objecting to, is not a simplifying assumption. Rather, as is addressed below, the concept actually does not call for any substantial departure from standard economic analysis.

The details of the supply side are not addressed in this analysis. Supply is treated as a competitive homogeneous product market with constant production costs that yield a constant wholesale purchase price to which taxes are then added. [11] The purchase price, P_p, is represented by the horizontal line in the graph. Since we are focused entirely on consumers, there is no reason to separate out the free-market price (the revenue collected

---

[9] Consumers treat different costs as equivalent when they are in the same currency (e.g., almost no one cares how the total purchase price is divided between the seller's and the government's revenue), but not everyone has the same tradeoff between different categories of costs. One person's response to a particular health risk might equivalent to their response to a $1/day increase in price, while another might give it the same weight as a $5/day price increase. Thus treating the two costs as if they can just be added vertically, at the same rate for everyone, is an oversimplification. However, relaxing this assumption requires moving to a model where individual consumers are modeled, since the variations in that tradeoff will not match the ordering of the consumers that generates the demand curve, and so there is no way to just replace the P_h horizontal line with some other shaped curve.

[10] A reader unfamiliar with economic analysis might be struck by the frequent statements about assumptions. Economics is no more assumption-laden than the fields of inquiry that might be more familiar. Rather, there is an ethic of clearly stating the assumptions/simplifications to facilitate a discussion about which of them might be inaccurate and affect the conclusions of the analysis and which can be improved upon. This is serves as a method of reducing complexity to manageable levels, rather than trying to use the complexity to hide assumptions or use it as an excuse for not doing any analysis. Other fields of inquiry simply tend to try to try to hide their critical assumptions by never making such statements "assume that this very artificial situation represents real-world human behavior and health" (in clinical trials or toxicology) or "assume that the model of the relationship among the variables that I forced onto the data is correct" (in epidemiology), and actively avoid any scrutiny of which assumptions might be wrong and might affect the conclusions.

[11] Higher and lower quality products within the category, and their different prices, are not considered. Production costs are considered constant so the typical upward-sloping supply curve that might appear in such a graph, which assumes production gets more expensive as quantity increases (and is actually seldom realistic), is replaced with the horizontal line for price.

by the sellers) from the taxes (which are a large majority of the purchase price in many jurisdictions), and so only the sum of the two costs is considered.[12]

Quantity Q_0 is the number of people who would smoke if there were no health risk and no other costs other than the purchase price. For concreteness, if the purchase price were the free-market (no substantial "sin tax") price, Q_0 would probably be 50% of the total population or a bit more, since this is the rate of smoking in populations where it is socially acceptable, consumers have enough wealth to afford it, products are not aggressively taxed, and there is not widespread concern about the health risks (e.g., men in Western countries c.1960). When $P_{h(cig)}$, the health costs from cigarettes, are added to the total costs, the portion choosing to consume drops to Q_1. More precisely, when the risk becomes *known*, and thus is incorporated into consumers' assessment of the costs, this shift occurs; if the marginal[13] consumers' perception of the health costs is inaccurate, the result will deviate from Q_1, as discussed below. Again, for concreteness (though nothing in the present analysis is anchored to these empirical observations), this can be thought of as about 25-30% of the population, a typical smoking prevalence in populations where there is common knowledge of the health effects, but taxes and social opposition are modest and smoke-free alternatives are not popular.

The huge welfare benefits from low-risk alternative products
Now consider a low-risk product as an alternative to cigarettes. Simplify the model by letting the purchase price be the same and assuming that every consumer considers the good to be perfect substitutes, differing only in their health effects. These assumptions result in the same demand curve for the low-risk product as the original one for cigarettes.[14] The health cost of these products is $P_{h(lrp)}$ (which is illustrated with a vertical distance that is far more than 1% the height of $P_{h(cig)}$ so that it shows up on the graph; it is not actually possible to illustrate the relative health risk of the low-risk products on the same linear scale as cigarettes on a reasonably sized page and still be able to distinguish them from zero).

_____

[12] An analysis examining both supply and demand side factors would typically include the supply curve and then add taxes as a second line on top of that.

[13] In economics, "margin" or "marginal" refers to the actor or action that is right on the borderline of changing. So in the case of using a tax to optimize behavior, we focus on the consumers who are just at the border of ceasing (or starting) consumption, which means that their net consumer surplus from their current choice is just barely higher than it would be under the alternative choice. Those who are completely dedicated to their choice (those at a point on the demand curve that is much higher or lower than the cost) will not be motivated to change by a small variation in the perceived cost.

[14] Readers thinking one step ahead will realize that the demand curve for cigarettes will not be the same in a world where low-risk alternatives are available, as discussed immediately after this note. Thus the word "original". The game-changing interaction of these two products makes the simplest economic analysis, in which everything is currently close to equilibrium and the shocks are relatively minor, inadequate. This comes into play in the discussion of price elasticity of demand at the end, and requires a departure from what can easily be illustrated in these graphs.

The consumer surplus gain from the new product is quite large.[15] This shift is clearly welfare increasing. The total consumer surplus in the cigarette-only world was CS_1 in Figure 1, the area above the total cost curve and below the demand curve for those who chose to consume. The vertical distance between these is the consumer surplus for one individual, and the area results from adding each of these together from left to right. The consumer surplus gain from all smokers switching to the (hypothetically perfect) substitute is CS_2, which is $P_{h(cig)}$-$P_{h(lrp)}$ multiplied by the number of users.

Put simply, this is the obvious point that health costs are one contribution to overall welfare. Thus, when health costs are almost eliminated, there is a huge increase in welfare. (Note that relative sizes of CS_1 and CS_2 does not represent a theoretical or empirical claim. It is quite possible that the true shape of the demand curve is such that the existing consumer surplus from the benefits of smoking is either greater or less than the health costs (imagine the slope and shape of the demand curve toward the left of the graph changing without changing the places where the lines cross.)

It is worth immediately reiterating the fundamental implication of this. Under the simplifying assumptions, every smoker is better off using the low-risk product than smoking, by a difference of $P_{h(cig)}$-$P_{h(lrp)}$ each, or a total of CS_2. But they are even *more* better off using the low-risk product as compared to being abstinent. They preferred smoking to abstinence, generating a surplus of CS_1 so if they prefer the low-risk product to smoking, it is clearly much better than abstinence (total surplus of CS_1+CS_2). It is easy to argue that someone might still prefer smoking despite this simplified conclusions – that merely requires recognizing that the low-risk products are inferior substitutes for many consumers. But it is difficult to find any defense for trying to push smokers to be abstinent when low-risk products are so clearly a higher-welfare option for them, other than by abandoning the pretense that anti-smoking policies are intended to make smokers better off.

Another conclusion that is immediately obvious from even this simple analysis is that there will be upward shift in the equilibrium value of Q if cigarettes are replaced with low-risk alternatives. With P_h reduced, the new demand is higher. Because $P_{h(lrp)}$ is so low, the new equilibrium, Q_2, where the demand curve crossed the total cost line, will be just short of Q_0. Q_2 might be much closer to Q_0 than is illustrated in Figure 1 (recall that the height of $P_{h(lrp)}$ is exaggerated), or it might be that the simplifying assumptions are wrong and it falls well short. Predicting the magnitude of the increase in use is much more difficult than predicting the direction, of course, but there is no doubt that the Q will increase if the health costs are reduced. Given the peak popularity of cigarettes and the popularity of coffee – the product that is most similar to smoke-free tobacco/nicotine – a reasonable guess is that the new equilibrium will be in the range of almost half the population using these products.

---

[15] Consumer surplus is the measure of the net gain to a consumer of consuming a good and paying its costs. In the graph, it is the vertical distance between the demand curve (which represents each individual's gross benefits) and the costs.

It is probably difficult for many readers to imagine that the direction of the effect was ever in doubt. After all, the question "will more people use a tobacco/nicotine product when it is low risk?" is equivalent to "do some people choose not to smoke because it is high risk?", and no one doubts that the answer to the latter is "yes".[16] Yet many regulatory discussions about tobacco harm reduction include, as if it were in question, calls to determine whether the introduction of low-risk products might increase total usage prevalence. (An alternative interpretation to genuine ignorance is that those setting the agenda for such discussions – a role generally captured by anti-tobacco extremists who oppose tobacco harm reduction for a variety of reasons[17] – do recognize the obvious truth. But they are trying to trick others into accepting this question as legitimate, so that when the answer is "discovered", they will have rhetoric in place that the "discovery" should be interpreted as a condemning harm reduction.)

When the increase in quantity occurs, some consumers who adopt the low-risk products will be those who would not be smokers in a world where the only tobacco/nicotine option were smoking. Some of the additional users will be in new cohorts coming of age who adopt low-risk tobacco/nicotine rather than ever smoking, as has been observed in the one population (Sweden) where tobacco harm reduction has been most effective. (The question of whether factors like rebellious nature of smoking in a society that prefers low-risk alternatives might still attract youth is beyond the present model.) Some adopters of low-risk products will be members of older cohorts who never smoked or quit smoking because their net consumer surplus from smoking was negative but becomes positive once P_h is reduced. Again, the portrayal of this as if it were not easily predictable seems to be based entirely on a combination of failing to understand the basic economics and political trickery.

---

[16] A previous observation of increased consumption resulting from reduced health risk is the increase in risky sexual behavior that followed the emergence of technologies that changed HIV/AIDS from a rapid death sentence to a somewhat manageable disease (thus dramatically lowering the cost of the behavior). Reducing the harm from consumption of a particular behavior makes those who are doing it better off, but also lowers the cost so that others are willing to jump into the "market". Despite the apparent genuine surprise about this, any economist would have predicted that reducing the cost of HIV would increase the number of HIV cases. Similarly, safer cars have led people to drive in ways that they prefer but that they would have avoided before (e.g., going faster on wet roads). Every improvement in medical technology or safety equipment makes people more likely to take risks. Few would rock climb or fly on airplanes if a broken leg were still a likely death sentence as it was for most of human history or planes crashed as often as they did in 1930. For the HIV or auto safety examples, it is possible that there might be a net increase in total health costs (depending on the shape of the demand curves), though for the case of low-risk tobacco/nicotine products, the reduction in risk is so close to elimination of risk that this is not possible, as discussed below.

[17] Some of the motives for opposing tobacco harm reduction are identified at http://antithrlies.com/about/

To see that it is rational for some consumers who would choose abstinence over smoking, return to the consumer surplus analysis. CS_3 is the addition consumer surplus that accrues to those who rationally chose to not smoke, because the total costs exceeded the benefits, but would rationally choose to use low-risk products. As an aside for those focused only on the health costs and who do not care about overall consumer welfare, the size of the new health cost rectangle determined by multiplying P_h by Q – wider by perhaps as much as double, but with such a small height that it really cannot even be drawn on the graph – is much smaller in spite of the increase in consumption.

It is probably obvious that this equilibrium-based analysis – which is based on full knowledge and frictionless decision making, as well as the assumption that the low-risk products are a perfect substitute for cigarettes and differed only in the lower risk – implies everyone would have already switched, since the low-risk products exist and have been known to be low-risk for many years. Everyone's consumer surplus is much higher if they choose the low-risk alternative. A dynamic model with learning and other sources of friction, rather than this equilibrium-based analysis, along with preference for one product over another, is needed to show why this switching is not immediate or complete.

A rational choice?
Following this exposition of how a simple consumer choice model illustrates and informs, it is useful to circle back to a question that is undoubtedly vexing some readers: Is it legitimate to model tobacco/nicotine use as rational choice? The typical discourse does not even acknowledge that there are benefits from smoking or alternative product use, let alone that they outweigh the costs. But claiming that hundreds of millions of people are making a choice that does not provide net benefits, is a rather extraordinary claim that defies our most basic knowledge of consumer choices. The extraordinary level of evidence and analysis needed to support such a claim does not exist, and proponents of that position seems to just be counting on no one ever mentioning neither economics nor the explicit testimony of consumers that they are deriving benefit.

Tobacco/nicotine use behavior, like all human behavior to some extent, involves important departures from the simple rational economic model. People do not have perfect information, and they sometimes react irrationally to what they know. The model can be improved to account for this possibility, as is done to some extent below. People use rules of thumb rather than optimizing every decision. But it should be remembered that the choice to smoke, in a world where the health risks are well-known and aggressively communicated, is one of the weightiest decisions in someone's life, not a causal choice.

There are various claims by those who seek to deny that this is a benefit-motivated rational choice. Some are so outlandish that they are not even worth addressing – e.g., that people are making a consumption choice that have enormous impacts on their psyches, lifestyles, disposable income, and health because they are permanently mesmerized by advertising or pretty packages. Those who seek a legitimate argument, rather than just counting on people to believe anything they claim, do not appeal to such

absurdities or to the standard modest departures from the rational, presumably realizing that some more dramatic departure is needed to avoid acknowledging rationality and the basic accuracy of the standard economic model. But these major departures are either just content-free hand waving or can be expressed in terms of welfare economics.

One substantive claim is that most smokers really want to quit (which in the present context translates into "would have higher welfare if they quit"), a claim that is supported by surveys in which respondents report as much. But economics favors revealed preference over stated preference, and with good reason: talk is cheap. People make many claims about their preferences that in practice, when they are asked to put their money where their mouth is (act according to their costs and benefits), are not really what they do. Because of this, stated preferences are generally considered unreliable in economic analysis when revealed preference is available, and are highly controversial in cases where revealed preference is unavailable.[18] Since there is an observed market in this case, claiming that hundreds of millions of people really do not want to make the choice they are making, and that they really prefer an alternative that is available and that they are aware of, requires rather more evidence than the cheap talk from some surveys.

This is especially true since there are obvious explanations for the survey answers that explain why they deviate from revealed preference. Respondents know that "I want to quit" is the socially acceptable answer to give (just as many dutifully declare that they are trying to quit in conversation). Moreover, to the extent that people are sincere in their declaration, it is apparent that they are stating a second-order preference, a preference to be someone who does not prefer to smoke, or a preference to have the benefits of smoking without smoking, not an actual preference to not smoke.[19]

The more substantive approach is to attempt to justify the extraordinary claim using the concept of "addiction". The scare quotes refer to the fact that, when not further clarified, the word does not represent a well-defined claim. Since it lacks an accepted scientific or even medical/psychological definition, and is merely an "I know it when I see it" concept, it is not terribly useful for informing analysis. If pushed to identify what behaviors constitute addiction, most people who use the word will describe something that is massively disruptive to short-term functioning or health, and thus does not characterize tobacco/nicotine use behavior. Indeed, when the word is used in the context of tobacco, it seems to mean nothing more than "they persist in a behavior even though it

---

[18] For some goods, there is no real market that would allow revelation of preferences. There are extensive literatures on such topics as quality-of-life and the value of environmental preservation that attempt to put numbers to these values so they can be used in cost-benefit type analyses. It is a worthwhile goal, and some version of those estimated values it is needed to make rational policy. Nevertheless, it is easy to observe that the estimates that derive from the research are often out of touch with reality and even internally inconsistent.

[19] For more on this concept, see the series at http://antithrlies.com/tag/second-order-preferences/

has downsides" (which describes *every* consumption choice) or merely "I personally disapprove of that behavior".

We can try to make a better case for the claim than most users of the word do. There seem to be two substantive categories of candidate semi-definitions. The first starts with the notion that someone's day-to-day benefits are higher because he has a history of smoking. This in itself is neither harmful nor unusual; benefits of a particular good often increase as someone moves up a learning curve and comes to appreciate it more. What makes this concept of addiction different and harmful is that the history of consumption lowers someone's baseline welfare, so that some of the benefits of each day's consumption are needed to merely bring the consumer back to the level of welfare he would have had he never used the products. Making a choice that leads down such a path can still be welfare improving, considering both the positive and negative effects (in Becker and Murphy's (1988) presentation of this economic conceptualization of addiction, their thesis is that such a choice can be rational and welfare maximizing). It is clearly inaccurate to claim that people want to smoke *only* because of the effects of past consumption (as some of the unsupported rhetoric in the field implies), but it is still probably the case that some of the day-to-day benefits are used up getting back to the abstinence baseline.

But if this is addiction, it turns out to be moot for purposes of analyzing harm reduction. The preferences of current smokers are what they are, regardless of how they got that way. The benefits of current consumption are still positive, even for the most extreme versions of this claim. This phenomenon might mean that the demand curve drops sharply for the nonsmokers who lack this learned appreciation for tobacco/nicotine (which would move Q_2 toward Q_1), but this does not change any of the analysis presented here.[20]

A second concept of addiction, sometimes described in terms of the slightly better defined concept of *dependence*, could plausibly exist at a level that that has practical implications. The claim is that individuals have multiple consumption equilibria and their consumer surplus is affected by their recent past consumption, and they may be stuck at local maximum that provides lower welfare than some global maximum.[21] That

---

[20] The extreme version is that *all* the benefits of smoking come from past use, and there are no positive effects other than digging out of the hole that previous use created, and moreover that every smoker would be happier as a nonsmoker had he never smoked. While some people seem to actually believe this, it is implausible and inconsistent with the evidence. However, for those readers who want to come back to this thought after finishing the paper, you will notice that though even this story does not affect most of the analysis, it would mean that $E_{abs,lrp}=0$, since abstainers have absolutely no desire to start using a low-risk product at any price. This would mean that this claim that "people only smoke because they are addicted" by itself is sufficient to argue that there should be no excise tax on low-risk products, and also that low-risk products will not attract new users.
[21] "Multiple equilibria" describes a situation where there is more than one consumption choice that is locally best – that is, it is better than any incremental change away from it.

is, there is a *short-term* change in tastes from *recent* past consumption, which might mean that while a smoker's net benefits from smoking are currently positive, so she wants to smoke today, if she were to avoid smoking for a while this would no longer be the case, so she would then not want to smoke that day. In this story, each of these states is an equilibrium, since she wants to keep doing what she is doing, smoking or not smoking, but they have different net welfare. It is possible that the nonsmoking state has higher net welfare for a current smoker, but she is not changing behavior either because she does not realize that she would be happier at the other equilibrium or the short-term costs to make the transition are too daunting.

This has face validity and is not just an appeal to some untestable and unsupportable claim about what his preferences would have been in a different world. Ignorance of potential changes in tastes or transition costs preventing people from improving their life is not an unusual economic story. Under this story, someone's welfare could be improved by breaking him out of his current behavior.

However, for many smokers, this story is contradicted by the revealed preference evidence. A large portion of smokers go through fairly long periods of abstinence, trying to quit, only to make the choice to start again. After overcoming the short-term effects of dependence and settling into an abstinence, many discover that it is not an equilibrium – they still feel like their net welfare would be improved by smoking, and it is a constant fight to not do so. Others find themselves in a new equilibrium but then realize they preferred the smoking equilibrium and switch back. Failure to recognize that abstinence, even if it is a locally stable equilibrium, is not the global optimum for many smokers seems to be the main reason why most smoking cessation methods are almost useless. They are based on the premise (and heavily marketed promise) that if someone can be transitioned from smoking to abstinence that he will be happy with the new state, but that turns out to seldom be true for those smokers who are not inclined to quit by just quitting.

Though it clearly does not apply to many smokers, if this story applies to some smokers, it *could* challenge the observation that pushing smokers to be abstinent rather than to switch to a low-risk product lowers their welfare. However, this would require a very extreme version of the story: Someone would have to be so much better off in the abstinence equilibrium that not only is his welfare higher than when he was smoking, but it is higher by more than $P_{h(cig)}$-$P_{h(lrp)}$, an enormous margin that seems to defy the "did not realize" or "not worth the short term cost" explanations for being stuck in the wrong equilibrium. Thus, even for smokers for whom the claim is accurate, it is still consistent

---

One of these is a global maximum (better than any other possible choice), but it is possible for someone to be "stuck" at one of the others and not want to move away from their current local maximum. Think of a ball rolling on hilly ground; it will stop at a point where it is immediately uphill every direction, but this might not be lowest point in the area. This is further complicated when the possible better equilibrium is not best today or tomorrow because it takes a while to come to like it better. Thus, someone might be happier in six months if change their current consumption, but they might be less happy than they are now for a while after making the change.

with tobacco harm reduction being welfare maximizing compared to either abstinence or smoking.

In sum, once the claims about the necessary huge departures from rational choice are expressed in concrete terms, it is easy to see where they are moot for present purposes, or how extreme the claims need to be to matter. The rational actor model is imperfect, and the present simplified version of it even more so, but there does not seem to any basis for denying its basic implications.

Optimal rates of taxation

An analysis of tobacco/nicotine tax policy can help to further illustrate the core points about the economics, as well as answering some specific questions about how to optimize taxes to achieve some particular goal. It turns out that for most goals the optimal tax rates to affect demand are quite simple to determine, though in many cases they are so extreme that they must be balanced against political realities and supply-side issues.[22]

One particular proposed simple formula turns out to be clearly wrong. There is an oft-repeated notion that the optimal excise tax on low-risk alternatives to cigarettes should be proportional to the comparative risk, with the ratio of the taxes matching the ratio of the risk. As shown below, this claim is wrong no matter what the goal of the taxation is. In fairness, some making the claim probably realize it is baseless, and are simply looking for a sciency-sounding claim that keeps the taxes on low-risk products at approximately zero, which this would. It turns out, however, that there is usually a much stronger case for setting the taxes to exactly zero (assuming that a negative tax – a subsidy – is not an option).

---

[22] Aside (included because it responds to some recently published errors, but readers interested in a more simplified version can just skip it): The cigarette market, and to a lesser extent the low-risk product market, is not perfectly competitive, but rather is an oligopoly with brand loyalty and other stickiness that limits price competition. An oligopoly – where there is more than one supplier (not a monopoly) but still not many – creates a situation where suppliers can increase their profits by raising prices a bit, trading off a small loss in quantity for a larger gain in per-unit income. This contrasts with a perfectly competitive market where any such move by one seller will cost them nearly all their sales because others will keep their prices as low as possible and take all their customers. In this situation, producers will further raise their wholesale prices if cigarette taxes increase, increasing their per-unit profit, because the profit-maximizing price for an oligopolist increases when their share of the total purchase price decreases (any given increase will reduce sales by less because it is a smaller fraction of the purchase price, but will still provide as much additional per-unit profit). This means that any tax will raise prices by somewhat more than the tax rate, and thus any stated level of taxation actually understates the net impact, with some extra "tax" collected by the supplier. One implication of this is that the higher the taxes, the higher the per-unit profit of the suppliers. Another is that if someone is interested in precisely quantifying optimal tax rates, this additional increase in price that would occur needs to be accounted for.

The question of optimal taxation depends on the goal, which could be any of:
1. aligning consumer preferences with their real costs to correct perception errors (i.e., maximizing consumers' welfare),
2. maximizing net social welfare,
3. minimizing health costs,
4. minimizing consumption,
5. maximizing revenue,
6. some objective function that balances two or more of these.

Goal 3 or 4 seems to be what many commentators have in mind when discussing the goal of taxation, though they seldom make this clear. In many jurisdictions, 5 is the real goal; this is not addressed here, since in addition to seldom being explicitly defended as an ethical goal, the analysis would hinge an assessment of consumers' preferences regarding to black market (unregulated, untaxed) or grey market (casual smuggling from lower-tax jurisdictions) competitive products. Category 6 includes infinite possibilities and absent anyone clearly stating a particular more complicated objective function it is difficult to know which to analyze.

Maximizing consumer welfare
Since normal analyses of using taxes as incentives focus on maximizing welfare, consider this first. If costs comprise purchase price and health costs, consumers fail to maximize their own welfare if and only if there is a difference between perceived and actual health effects. If consumers at the margin underestimate the magnitude of P_h, either due to factual error or the natural tendency to avoid thinking about risks one is accepting, then at the margin there will be too much consumption. Some consumers who perceive their net benefits to be positive will be mistaken and will be making the wrong choice. In that case, a tax equal to the difference between the perceived health cost (P_h^* in Figure 2) and the real health cost will maximize their welfare by dissuading them from making the wrong choice. The tax will also be borne by all inframarginal consumers, and so will decrease their welfare, which will widely perceived to be a cost, though the net social effects a transfer, not a real resource cost.[23] Of course, if consumers accurately estimate

---

[23] Inframarginal refers to anyone who is part of the market but not at the margin – i.e., far enough to the left of the point where the benefit and cost lines cross. Those who are far enough to the left will not change their behavior due to higher taxes but will still have to pay it the increased purchase price. Real resource costs are inputs that are consumed (e.g., the goods and labor that go into manufacturing, or the health costs), and thus are true costs to society. Transfers are losses to one party that are fully captured by another such as taxes (the consumer's loss is the government's gain). Transfers create the same incentives for those paying them that any other cost would, but from a social perspective they do not create a cost because no resources are consumed. However, there are often arguments about whether it is fair to burden a particular group of consumers with more than their share of taxes. In this particular case, though, there are many anti-tobacco activists who see the added burden as a benefit because they believe that smokers deserve to be punished for their behavior.

the health costs or even overestimate them (which seems to be common in populations where there is aggressive anti-smoking messaging) then any positive tax rate moves consumers *further* from their optimal choice.

Figure 2
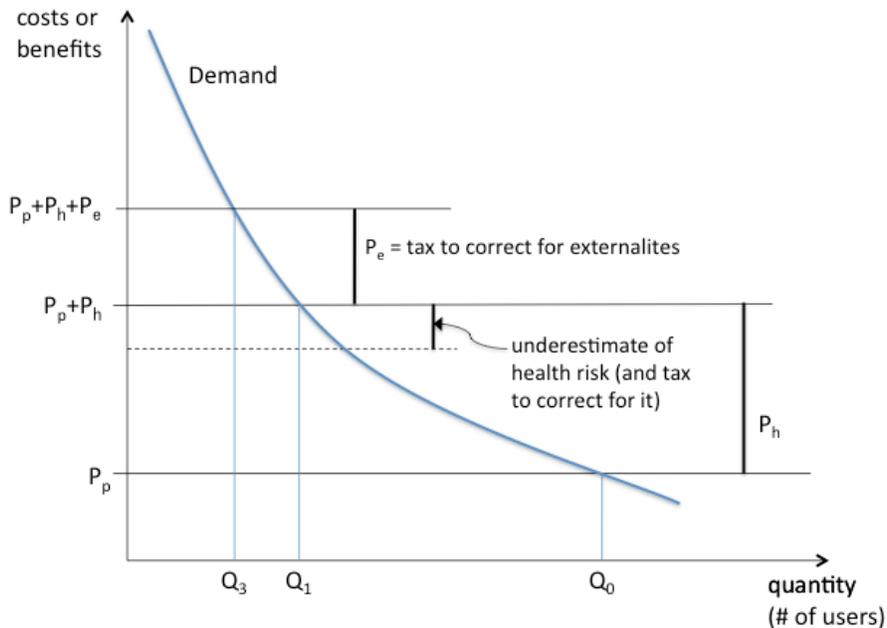
Note: This illustrates the optimal tax rates if they are truly positive. As noted in the text, this is often not the case.

Conceptualizing such a tax based entirely on consumers' technical knowledge of the risks has its flaws. We know that people sometimes respond to risk in ways that do not optimize their lifetime welfare. But trying to deal with this is fraught with enormous difficulty and controversy. Demonstrating that someone who understands the risks is still not optimizing their behavior (i.e., maximizing their welfare) based on this knowledge, let along quantifying the failure, is extremely difficult. Almost all commentators who claim that people are acting irrationally in the face of risk are actually just claiming, "I, personally, would react to that risk differently, so they must be acting irrationally." But the accepted ethics of free markets as part of a free society that are embedded in economics see this claim as equivalent to saying "I, personally, dislike the taste of beetroot, and so anyone who eats it must be acting irrationally." Further analysis of the point is beyond the present scope, but suffice to say that anyone who claims that someone else's informed preferences are not merely different, but are fundamentally irrational at a level that calls for correcting them as if they were incorrect knowledge, is making an extraordinary claim that requires extraordinary evidence, not just political rhetoric.

Thus, if someone's goal in setting the tax level were maximizing consumers' own welfare (unlikely to be the case, but a good starting point), it seems likely that the optimal level of

tax on cigarettes would be about zero. It might even be negative due to the overestimated health effects, though obviously this is not a possibility.

The optimal tax on low-risk products, based on this goal, should also compensate for any misperception of the health cost. Since the risk from those products is almost always grossly overestimated, a large negative tax (subsidy) would be in order. While this is unlikely to happen in general, it is not an entirely ridiculous notion: UK regulators have moved to approving low-risk alternatives to cigarettes as a "medicine" for tobacco harm reduction, which may result in them being subsidized by the medical system. The pharmaceutical industry's entry in this category are already subsidized in many jurisdictions (albeit only for a few months, not as a long-term harm reduction solution).

When considering the claim that the tax rates should be proportional to risk, it is clear that it cannot be based on this goal. If taxes on both products are stuck at a zero lower bound (i.e., there is no possibility of subsidies, and thus the lowest possible tax rate is zero), then the would turn out to be proportional (since $1/100 \cdot 0 = 0$), but this is presumably not what the commentators have in mind. If the taxes were actually optimized to correct for consumer misperception, there is no reason to expect they would be proportional to risk. However, the faulty intuition that the tax rates should be proportional might trace back to this goal. If a commentator has the mistaken belief that consumers completely ignore health costs when making consumption decisions, then the perception-correcting taxes would indeed be proportional to the risk. But since that premise is clearly wrong, so is the conclusion.

Maximizing net social welfare
To expand to the full social cost-benefit analysis, we need to consider any net externalities.[24] It is generally believed that smoking creates net negative externalities, primarily because of non-consumer-paid medical costs for treatment of resulting disease. In reality, it is not entirely clear that such costs are not offset by medical costs savings due to early mortality, and it is quite clear that any net cost is exceeded by total foregone consumption due to early mortality (most of the years of potential life lost due to smoking occur after retirement, so there is little foregone production). However, there are also aesthetic externalities (smell of smoke, litter) and some health costs imposed on others, and some argue that the lost "productivity" of companionship due to early mortality after retirement should be counted.

Setting aside the substantial uncertainty about whether it is true, assume the net externalities are negative. In that case, the level of consumption is greater than the optimum and a simple Pigouvian tax should be imposed to make the price paid by the

---

[24] Externalities are costs and benefits that someone's consumption choice imposes on others who are not party to the choice. Most of the time it refers specifically to *negative* externalities, net costs imposed on the world that are not reflected in the price on which the consumer makes his decision.

consumer reflect the total social cost and not just his own personal costs.[25] This is illustrated in Figure 2 by the line at $P_p+P_h+P_e$ where the last term reflects the total net cost of the externalities and thus the tax.[26] Figure 2 also illustrates that correcting for both misperception of health costs and externalities can be accomplished by adding the two taxes, or subtracting one from the other if the needed adjustment is actually a price reduction. Consumption with both of those corrections included would drop to Q_3 (assuming that the net externalities are actually negative and thus a positive tax of P_e should be added).

If we apply the same logic to the low-risk alternatives, there seems to be little possibility the net externalities are negative, and thus no justification for a Pigouvian tax. Unlike the expensive lingering diseases caused by smoking (lung diseases, cancer), to the extent that there is any health risk from smoke-free alternatives, it seems to be dominated by cardiovascular stress which, if it has any effect at all, tends to cause a rapid death with minimal medical costs. If there is such a risk, it almost certainly causes a net savings in medical costs alone, to say nothing of the reduction in other consumption. The products create no aesthetic or health effects from environmental smoke, and littering is substantially reduced.

Returning to the question of relative tax rates, there is clearly no basis for suggesting that the externalities, and thus the optimal tax rate to maximize social welfare, are proportional to the risk. There is no reason to believe there are net negative externalities from low-risk product use, and if there are, they are down in the range of typical consumer consumption choices where there is no serious consideration of imposing a Pigouvian tax. Moreover, even if the net externalities from low-risk product use were negative, there would be no reason to predict they are proportional to risk. Once again, the optimal taxes are only proportional if they are both exactly zero.

Minimizing consumption
If the goal is to minimize consumption of all tobacco/nicotine products, the tax rate should simply be infinity (i.e., a ban on legal sales at any price) on all products. If the goal is to minimize *smoking* and the infinite tax on cigarettes successfully eliminated use,

---

[25] A Pigouvian tax is just what is implied by this sentence – a tax specifically designed to make sure the consumer is feeling the full social costs he is imposing, and thus is not inefficiently overly-inclined to make the consumption choice.

[26] An additional complication is that the above observations about oligopoly mean that there is some *producer surplus* that increases net social benefits. Because the producers are able to charge a price higher than what is needed to just cover their costs, some of the purchase price is a pure transfer from consumers to the capitalists. That is, some of what the consumers pay is not consumed in producing their product, but is pure profit to the seller, meaning it is not a net cost to society, and so net social effects of consumption are somewhat more positive than depicted. However, there are arguments that this should not be counted as a wash (ranging from the fact that it is a transfer from relatively poor people to the rich, to profits accruing to evil tobacco companies should not be counted as a social benefit) so to keep this simpler and ignore those fights, this surplus is ignored.

then the tax on low-risk products is moot, and so could be anything. But if we recognize that there would be a black market that kept cigarette consumption nonzero in spite of the ban, then the infinite tax on cigarettes should be accompanied by zero tax (or a subsidy if possible) for the substitute products. While these are not very interesting scenarios to analyze, they are the apparent goal of many activists – to use taxes and any other policy to move as close to a ban as possible.

The more realistic version is to realize that cigarette taxes remain finite due to black or grey market competitive pressure and political constraints, even if their champions really would prefer a successful ban. In that case, the optimal level of taxes on low-risk alternatives remains zero (or as negative as possible) for the goal of minimizing just smoking. If the goal is to minimize total product use, however, the optimal tax level for the low-risk products is *higher* than that for cigarettes. If the potential for smuggling is about the same, then the higher tax needed to discourage the use of products with lower health costs will mean that the optimal tax rate (trading off the effect on discouraging consumption and the effect on encouraging smuggling) will be higher.

Tax structures and other regulations (e.g., the European Union ban on popular smokeless tobacco products) that make cigarettes cheaper compared low-risk alternatives are widely interpreted by harm reduction supporters as attempts to benefit the tobacco control industry, whose business model depends on tobacco/nicotine use staying highly risky. However the economic perspective shows that these regulations can also be interpreted as attempts to minimize total use – at the expense of both public health and consumer welfare – in the context of cigarettes having natural disincentives due to the health effects but low-risk products not having such disincentives.

Once again, it is clear is that the taxes to support any of these goals are not proportional to risk unless they are both set to infinity.

Minimizing health risks
Minimizing health risks seems to be the implicit goal most commentators have in mind when they make claims about optimal tax policy. Some readers may protest that, "of course that is the goal, why even question it or bother with the others?" To respond to such protests, it is worth noting that just like the goal of minimizing use, this is a niche political view that lacks justification in economics or any other ethical systems or empirical observations. The "public health" notion of elevating health to trump all other consumer costs and benefits – which is not limited to discussions of tobacco/nicotine – does not fit any defensible ethical model, and is clearly contrary to actual human decision making. The political position that "it is better, if possible, to cause someone to be abstinent rather than use a low-risk tobacco/nicotine product, no matter how enormous their consumer surplus from such use" runs contrary to not just normal policy ethics but to the most fundamental tenet of health ethics, that people should be free to make informed health-affecting choices to maximize their welfare.

If there were no black/grey market competition, the goal of minimizing health risks would be identical to the "minimize use" goal: The tax on cigarettes should be infinite,

since there is no possible advantage from letting anyone choose the most hazardous option. Moreover, since the assumption is that this ban is effective, the tax on the low-risk products should also be infinite because, given that there was no one to rescue from smoking, there is no health gain to be had by allowing the use of the alternative.

The more useful scenario is to not consider an unrealistic ban, but rather a high but finite tax on cigarettes that is determined by politics rather than any optimization. If we take cigarette taxes being as being fixed at some finite level, it is possible to calculate the optimal tax rate for low-risk alternatives by introducing two additions to the model.

The first is that the products cannot be treated as perfect substitutes apart from their P_h and P_p. Otherwise, as noted above, given that the respective values of P_p are not terribly different, the difference in the values of P_h would mean that everyone would switch to the lower risk products. Since this is empirically not the case, it must be that the non-health net benefits of cigarettes are much greater for many tobacco/nicotine users. (This seems quite likely to change over a relatively short period, as technology and social factors, but is true now.) Given this situation, there is some price advantage that would cause a particular current smoker to switch products, thereby dramatically improving their contribution to the objective of minimizing health costs.

But there is a tension that a higher price for the low-risk alternatives will discourage nonusers from adopting them, a choice which worsens their contribution to this objective, albeit by only a small amount. To quantify such a tradeoff, it is necessary to introduce the concept of *cross-price elasticity of demand*. The concept of *own-price elasticity* of demand,[27] the change in consumption that is caused by the change in the price of a good, has been at the core of the analysis so far (it is represented by the slope of the demand curve), though the term has not been used. The cross-price elasticity is the change in consumption of a good when the price of another good changes. This is usually described in terms of the percentage change in consumption resulting from a percentage change in price, so a cross-price elasticity of 1.0 would mean that if the price of the low-risk product decreases by 1% then consumption of cigarettes decreases by 1%.

To complete the equation, we also need a cross-price elasticity for the "good" abstinence, as a function of the price of low-risk alternatives. Consumption prevalence of low-risk products is determined by how many people switch from either of the other consumption choices, cigarettes and abstinence. Sandwiched in between these choices and not yet very popular in most populations, low-risk product use is dominated by the attraction of some portion of the much larger groups. This is a departure from the usual conceptualization, which considers own-price elasticity, but it is the better model in this case (more details in the footnote[28]). Cross-price elasticity for abstinence works the same

---

[27] Which is usually just called "price elasticity" when there is no ambiguity.

[28] The standard own-price elasticity concept, in which price-affected changes in consumption of a good are represented by a change in the current consumption for that product, does not work in this case. The change in the quantity of low-risk product use is minimally affected by the current quantity of users, since it is dominated by the effects of

way, with a particular change in the price for low-risk products affecting the portion of the population choosing to switch between abstinence and low-risk product use. (It does not matter that abstinence does not have a purchase price; it has whatever costs and benefits that it has.)

With this we can calculate the health (or any other) impact of a change in taxes for the low-risk products taking everything else as fixed. The increase in total health impacts for a p% change in the price of the low-risk product equals,

$$p \; E_{cig,lrp} \; Q_{cig} \; (H_{cig} - R \; H_{cig}) - \; p \; E_{abs,lrp} \; Q_{lrp} \; (R \; H_{cig})$$

where $E_{cig,lrp}$ is the cross-price elasticity for cigarette demand as a function of the price for low-risk products, $E_{abs,lrp}$ is the cross-price elasticity for abstinence demand, H_{cig} is health risk from cigarettes, Q_{cig} is the current number of smokers, Q_{lrp} is the current number of low-risk alternative users, and R the comparative risk that converts H_{cig} into the risk for the low-risk product.

This is much simpler than it might look. All its says is that the number of smokers is reduced by a factor of p times $E_{cig,lrp}$, and each smoker who switches removes the health risk from smoking but adds the health risk from the low-risk, but also the number of low-risk product users increases by p times $E_{abs,lrp}$ times the number who currently abstain, and each one of these adds the risk of low-risk product use. When p is positive (a tax is imposed) then the first term will be a positive number, an increase in the total health costs, while the second will be negative, a decrease.

To make this more concrete, imagine a population of 1 million people, with 20% smokers and 75% abstinent. Let $E_{cig,lrp}=0.05$ and $E_{abs,lrp}=0.05$. These are small numbers for price elasticities of goods that are good substitutes in the common-language sense of the word (consumers of, say, butter vs. margarine have cross-price elasticity almost 1), but we would expect that people's long-run decisions about choices this important are not hugely

---

price on the two competing consumption choices (cigarettes and abstinence). Attracting 5% of all smokers to switch would double the low-risk product users in some populations, but increase it by 10-fold in others. Trying to model this with own-price elasticity is hopeless. The current baseline level of low-risk product use simply does not matter much (though it does affect the social dynamics that would attract smokers, but that is a different analysis). Strictly speaking, in the usual cross-price elasticity conceptualization, it is not assumed that the effect is purely substitution of one product for another when the price of one good affects the consumption of another. In this case, though, we can focus just on substitution – attracting or discouraging someone's crossing over from another choice to low-risk product use – so any loss of consumption in one of the other categories is an increase in low-risk product use. All of the available modeling choices are simplifications, of course, but the simplification that is used here is much better than the usual simplification (which will someday be better, when we are closer to an equilibrium level of low-risk product users, rather than the current rapid and uneven uptake).

affected by price. If smokers are not already switching to the low-risk products, guessing that a 20% reduction in price would only attract 1% of them to switch seems to be in the right range. To be conservative in the illustration, the same number is chosen for non-users though this seems far higher than is realistic (very few non-users of tobacco/nicotine are motivated to remain abstinent because of the price of low-risk products). Then, assuming that R is about 1/100th, the net public health impact of a 1% *increase* in the purchase price of alternative products (a 1% ad valorem excise tax) is (with the H_{cig} term is just left as the unit of analysis rather than trying to convert it to some other units):

$$1\% \cdot .05 \cdot 200,000 \cdot (.99\ H_{cig}) - 1\% \cdot .05 \cdot 750,000 \cdot (.01\ H_{cig}) = 95\ H_{cig}$$

This means that the 1% increase in the price of low-risk products would increase total health costs in this population by about one hundred smokers worth. Breaking it down, this would include 375 more people who are abstinent rather than adopting low-risk products, but their health benefits are dwarfed by the 100 additional smokers. Even giving the tax too much benefit of the doubt the result is clear: any positive tax is harmful. Indeed, even if we made the scenario more extreme and dropped the cross-price elasticity of smokers to $E_{cig,lrp}=.005$, there would still be a net improvement in public health from any positive tax on low-risk products.

To provide the obvious sensitivity analysis, if R were increased to .05 (an implausibly high number given what we know about the products), the result would drop only to 76, still representing a clear net health cost from taxing the low-risk products.

As has been pointed out for decades, given how low the risk of alternative products is, it is not actually possible for their existence to increase *total* health costs, no matter how popular they are with those who would otherwise be abstinent. With only a few percent of would-be smokers using the low-risk alternatives instead, the net effect would be a reduction in risk even if every non-user adopted the products. But that observation was never sufficient to conclude that for prices *at the margin* the incentives for switching from each other group might not justify a tax to discourage use. However, the present analysis clearly demonstrates that this too is the case.

Until we get to a time where the number of smokers is very low, no remotely realistic parameters for this equation suggest that the tax should be higher than zero. Otherwise, to get to a case where $E_{abs,lrp}$ is enough higher than $E_{cig,lrp}$ so that a positive tax would reduce net health impacts, we need some very extreme scenarios. For example, if the low-risk products were almost free, anyone still paying a lot to smoke would be unlikely to respond to a further price drop, whereas many non-users would probably adopt them if they were given away free.

The latter observation means that to calculate the true optimal *subsidy* (since a positive tax moves us the wrong way regarding this particular goal, a negative tax must move the right way), we would need to account for how the elasticities changed for very large changes away from the status quo. That is, as the net purchase price of low-risk products

moved close to zero, we would reach a point where enough more non-users were being attracted to them, as compared to more smokers, that we should stop increasing the subsidy. Calculating that optimum is well beyond the present scope, but in any case is moot in a world where only price increases, not subsidies, are being considered.

It is clear for the goal of minimizing health risks, the claim that taxes should be proportional to risk is once again wrong. Whatever the tax on cigarettes, the current tax on low-risk products should be zero or, if possible, negative.

Discussion
This relatively simple application of basic economic intuition and modeling provides definitive answers to a number of questions that are not intuited by everyone who is interested in understanding tobacco harm reduction. There is no doubt that substitution of low-risk products, assuming their gross benefits for consumers is anywhere close to that from smoking, is welfare-enhancing for most smokers. Similarly, there is no doubt that their use is welfare-enhancing for some people who prefer abstinence to smoking. The latter tells us that if there is availability of and honest information of low-risk products, there will be more total tobacco/nicotine use as compared to a world where smoking is the only option.

Claims that there is no evidence that low-risk products will attract some would-be abstainers, even as they attract many would-be smokers, are simply not defensible. Proponents of tobacco harm reduction who make such a claim are not only ignoring the obvious, but are setting themselves up to be shown to be wrong. This is not a terrible personal risk in the realm of tobacco analysis, where there seems to be no penalty for emphatically making claims that prove to be clearly wrong (a large portion of those working in the field owe their continuing careers to this lack of accountability). But beyond the personal level, tying the defense of a harm reduction strategy to a claim that will inevitably eventually be recognized as false is not wise.

The better argument is that everyone who deviates from another would-be consumption option does so because they gain an increase in welfare, as is the case with any new option for consumers that some prefer over all current choices. No one is made worse off if we consider a defensible measure of well-being (i.e., we do not pretend people only about health). It also happens to be the case that though some consumers would be increasing their risk, there is no conceivable scenario that could lead to a net increase in total social risk. This is not an inevitable of consumer choice, like the utility benefit is, but is a result of just how low the risks of the low-risk products are.

A bit more analysis shows that for almost any of the obvious goals, the optimal tax rate on low-risk products is zero (or negative if a subsidy is an option). The only exception is the goal of minimizing consumption of all tobacco/nicotine products, in which case the tax should be as high as possible, and if constrained to be finite, the tax should be higher for low-risk products than for cigarettes. The only goals that result in the optimal taxes on the two product categories being proportional to the risk are those that are optimized

when they are both zero or both infinity, and so that ratio never holds if the cigarette taxes are assumed to be positive and finite.

It is certainly possible to argue that the basic welfare economics theory, no matter how clearly accurate and useful it has proven for understanding consumer behavior, misses some important consideration in this case. However, questioning the predictions of well-established theory requires more than waving one's hands and saying "it is not that simple." Denying the established model is an extraordinary claim that calls for extraordinary evidence and arguments. No alternative consumer modeling seems to have ever been attempted by those who claim that hundreds of millions of people are being fantastically irrational in their choices, day after day. The economic model is the established and proven basis for analyzing most policy decision, whereas the competing view seems to be little more than a modern-day version of "they are all possessed by evil spirits and just need an exorcism."

All social science theories are imperfect simplifications of reality, and their value depends on whether they are useful despite the simplifications. Any claim that a model is improved by adding complexity needs to be defended, and a claim that the only available solid model is uninformative even more so.

It is always the case that empirical observation trumps theory; if experience shows that the theoretical conclusions are wrong, then the theory needs to be fixed. But this observation refers to solid, replicable, carefully examined empirical results. When a one-off result with an *ad hoc* analysis contradicts a well-established theory, the better conclusion is that there is probably something wrong with the data or analysis. Moreover, it is important to understand that every interpretation of empirical observations is theory laden, and if that theory is wrong, then the conclusions are likely to be wrong. Health science has a long history of blindly relying on empiricism without much thought about fixing or testing the theory, which is a recipe for self-perpetuating superstition. If every patient who recovers after bloodletting is considered evidence that the theory is correct, then applying leeches is "supported by the evidence". Without John Snow's statistical analysis, the reality of infectious agents in drinking water goes unnoticed because everyone who gets sick was exposed to miasma, the untested accepted explanatory theory.

In many cases, defensible theory alone provides far better information than any available empirical evidence. Because we had an established germ theory of disease, it was quite clear that AIDS was infectious before HIV was identified; empiricism that was "informed" by earlier medical views would have blamed it on witchcraft or the gods hating gay men and black people. Similarly, it was long clear that oral cancer was infectious, even as it was widely blamed on low-risk tobacco products. Predictions based on historical data about the effects of tobacco product price and availability on the use of various products are so thoroughly confounded by other trends and shocks, as well as the complications from disequilibrium and imperfect information, that they are practically useless except in the context of a defensible theory. When the data is interpreted based

on some theory that assumes tobacco use is not caused by rational choice, and that thus there is little interest in low-risk alternatives to smoking, they are particularly suspect.

An economic model of tobacco/nicotine consumption is not merely theoretically defensible and empirically supported by analogy, but it passes the test that is often taught as the *sine qua non* for distinguishing science from religion: making empirically refutable predictions, that if refuted would lead to the conclusion that it was wrong. The alternative view that consumption is caused by latter-day witchcraft (the magical thrall of marketing and some vague concept of addiction) does not past the test; it makes few claims are testable and those that are testable have repeatedly proved to be false. Ironically, it is also possible to make testable predictions about those who refuse to acknowledge that tobacco/nicotine use has benefits: When the population shifts toward the new equilibrium of higher usage prevalence of low-risk products, the advocates of the witchcraft theory will ignore the actual evidence that people like the products and attribute the increase in prevalence to industry marketing efforts.

People who are not empowered by a bit of scientific analysis can easily be tricked into believing clearly inaccurate claims that just sound good ("marketing causes all product use", "taxes should be proportional to risk") or are supposedly backed by data ("the limited adoption of smokeless tobacco in most places means that smokers are not interested in reducing their risks"). Scientific theory, even when it does not offer and answer to a question straightaway, points out how to ask the right questions and where to express doubt. Sadly, the study of tobacco harm reduction is marked by doctrine and a lot of absurd claims are supposedly based on evidence. The vast majority of these come from those who are opposed to all use of tobacco/nicotine as a matter of personal pique, but some come from other quarters (e.g., "there is no evidence that low-risk products will lead to an increase in total consumption"). It is hoped that this presentation and those that will follow can help immunize the readers against some of the naïve conclusions that are produced by this pseudo-science.

References

Becker, G. and K. Murphy (1988) "A theory of rational addiction". Journal of Political Economy, 96, 675-700.

Phillips CV, Rabiu D, Rodu B (2006). Calculating the comparative mortality risk from smokeless tobacco versus smoking. American Journal of Epidemiology, 163 (11):S189.

Phillips CV (2009). Debunking the claim that abstinence is usually healthier for smokers than switching to a low-risk alternative, and other observations about anti-tobacco-harm-reduction arguments. Harm Reduct Journal, 6:29.