CESifo
AREA CONFERENCES 2021

Energy and Climate Economics

Munich, 4–5 March 2021

Differential Fiscal Policy Induced Innovations in Consumer Markets

Marcus Hagedorn and Yikai Wang

CESifo

Differential Fiscal Policy Induced Innovations in Consumer Markets

Marcus Hagedorn[1] Yikai Wang[2]

August 8, 2020

Abstract: Many consumer goods - sugar, cigarettes, alcohol, fossil fuels - are considered "sin" goods as they cause externalities like CO_2 emissions or internalities such as addiction. The standard response is then to appeal to the Pigouvian principle and tax these goods to correct these ex- and internalities.

This paper builds on this fundamental Pigouvian insight but argues that the effectiveness of the traditional approach is limited. The main reason is that in many cases, close substitutes are missing, which provide similar benefits to the consumer but are less harmful. As a result behavior does not change and the fiscal intervention is regressive.

We then make two points. First, if those substitutes exist, then a *differential fiscal policy* which directs consumer behavior from the "sin" good to the less harmful substitute complements the pure Pigouvian approach. We survey the nascent empirical literature on this topic and find that this differential approach is promising.

Second, if those substitutes do not exist, then a commitment to implement a differential fiscal policy ("fiscal forward guidance") in the future (not now) will induce innovations today and finally deliver a substitute in the future.

[1]University of Oslo. Email: marcus.hagedorn@econ.uio.no
[2]University of Essex. Email: yikai.wang@essex.ac.uk

Contents

1 Introduction 2

2 Model 4

2.1 Environment 5

3 Differential Fiscal Policy: Theory and Evidence 10

3.1 Theory 10

3.2 Empirical Results 14

3.2.1 Taxation of Sugar-sweetened Beverages 15

3.2.2 Taxation of Novel Tobacco/Nicotine Products 17

3.2.3 Taxation of CO_2 emissions 20

4 Fiscal Forward Guidance 21

5 Conclusion 24

References 27

Bibliography 27

1 Introduction

Many consumer goods - sugar, cigarettes, alcohol, alcohol, fossil fuels - are considered "sin" goods as they cause externalities like CO_2 emissions or internalities such as addiction, for example as a result of self control problems for smokers. The standard response is then to appeal to the Pigouvian principle and tax these goods to correct these ex- and internalities.

This paper builds on this fundamental Pigouvian insight but argues that the effectiveness of the traditional approach is limited. The main reason is that in many cases close substitutes are missing, which provide similar benefits to the consumer but are less harmful. As a result behavior does not change in the desired way and the fiscal intervention is regressive. Either consumption of the sin good is not reduced and if it is lowered then the burden of reduction is on low-income households who cannot afford the higher taxes. Consumption of an alternative less harmful substitute is largely unchanged.

This paper aims at adding two points to this discussion. First, if a close substitute for the traditional sin good exists, then a *differential fiscal policy* which directs consumer behavior from the "sin" good to the less harmful substitute complements the pure Pigouvian approach. A differential fiscal policy makes the traditional sin good less attractive and the new good more attractive. An example would be a tax increase of the traditional good - a car with a high CO_2 emission - and a subsidy or a tax decrease on the new good, for example electric cars. We first derive the theoretical properties of a differential fiscal policy. We show when it works and point out some caveats in the design of such a policy. We then survey the nascent empirical literature on this topic and find that this differential approach is promising and successful in inducing consumers to switch from the traditional to the new good. Specifically we consider three markets, which differ with respect to the availability of substitutes. The market for cigarettes is a market where a close substitute exists as e-cigarettes are close substitutes for traditional cigarettes. The empirical literature finds that an increase in the tax on e-cigarettes leads to a quite strong substitution from e-cigarettes to traditional cigarettes or equivalently that fewer people switch from traditional to e-cigarettes than would have in the absence of the e-cigarette tax increase. The empirical literature on the taxation of sugar-sweetened beverages also confirms our theoretical predictions. Households substitute from taxed to untaxed goods. However, the design of the tax was not ideal as it did not tax sugar independently of the type of consumption good but exempted for example sugary fruit juices from taxation but taxed diet drinks. As a result, while calorie usage was not reduced significantly, the presence of consumer switching suggests that a better designed policy could be successful in reducing calories through inducing substitution. The third market concerns the CO_2 reduction, which turns out to be the most difficult one to induce substitution due to the absence of close substitutes. One stark example is the high taxation of fuel in Germany which renders the usage of combustion engine cars significantly more expensive without having induced significant substitution towards CO_2-emission reduced cars, simply because these cars are not considered close substitutes by consumers. The market for electricity provides some evidence for substitution as a CO_2 tax lead to a coal usage reduction in electricity generation in Great Britain from 40% in 2013 to 3% in 2019.

Our second point addresses markets, where close substitutes do not exist. For such markets we show that a commitment to implement a differential fiscal policy ("fiscal forward guidance") in the future (not now) will induce innovations today and finally deliver a substitute in the future. We develop an innovation model building on the seminal contributions of Grossman and Helpman (1991) and Aghion and Howitt (1992) and in particular of Acemoglu and Zilibotti (1997). In this model firms decide whether to invest in developing new less harmful goods. A key variable determining their efforts is the potential profits that they expect in the future. These profits in turn are proportional to the market size of the new good in the future. This is where our two points are connected. A commitment to implement differential regulation in the future will enlarge the demand and thus the market size of the new good in the future. Firms today anticipate this and as a response start innovating. Differential fiscal policy and innovation are complements which reinforce each other. The larger is the future market for less harmful goods, the stronger are firms innovation incentives today. Policy makers should therefore consider to foster a large market for new goods even though the market size might seem too large ex-post. For example, a commitment to a large market for electric cars will foster innovation in electric cars but such a policy might increase the total number of future cars. However, although the number of future cars could increase, the number of cars with combustion engines has decreased, but by less than the increase in electric cars.

The paper is organized as follows. Section 2 presents the model. It features two stages. In the second stage both the sin good and a substitute are available. The first stage is an innovation stage where firms invest into $R\&D$ taking into account consumer behavior and policy actions in stage 2. Section 3 shows theoretically that in the second stage where both the sin good and a substitute are available that a differential regulation will induce consumer switching. We do not stop at this simple theoretical insight but also address the essential question whether this theory work in practice. We therefore discuss several empirical papers in this nascent area which in our view apply the highest methodological standards. We find that differential fiscal policy, i.e. mainly differential taxation, works and that the empirical and theoretical results are in line. Section 4 then develops the fiscal policy forward guidance principle for consumer markets. We show that a commitment to differential regulation in Stage 2 will induce innovations in Stage 1 which eventually deliver a substitute which is close enough to make consumers switch and at the same time is harmless or less harmful than the sin good. Section 5 concludes.

2 Model

In this Section we lay out a standard innovation model. There are two types of goods. A traditional good - for example a high CO_2 emission good - without any room for further improvement and thus without any innovative activity. And a second new good - for example a low or zero CO_2 emission good - with room for improvement and thus innovative activity.

2.1 Environment

Time is discrete and for simplicity there are only two periods. The first period is the innovation period and households consume in the second period. The environment for the variety of goods is similar to but more general than the basic models in Grossman and Helpman (1991) and Aghion and Howitt (1992). For analytical tractability, there is a continuum of goods $i \in [0, 1]$. The traditional goods are indexed with $i \in [0, \alpha)$ and their quality is fixed and thus innovation is ruled out. We make this stark and somewhat unrealistic assumption to focus on the large difference in innovative possibilities between traditional and new goods. Allowing for a small amount of innovation in traditional goods would not affect our conclusions. In contrast, the quality of new goods $i \in [\alpha, 1]$ can be improved through innovation. The modeling and the cost of innovation follow the groundbreaking work of Acemoglu and Zilibotti (1997). In addition to consuming the traditional and the new good - for example cars with combustion engines or with emission free engines - there is also a general consumption good c, including for example food and housing.

Household Problem. The representative household has an (experienced) utility function and maximizes:[3]

$$\begin{aligned} &\max \frac{1}{1-\sigma}\left(\int_0^\alpha q_i x_i^{\frac{\epsilon-1}{\epsilon}} di + \int_\alpha^1 q_i y_i^{\frac{\epsilon-1}{\epsilon}} di\right)^{\frac{\epsilon}{\epsilon-1}(1-\sigma)} + c \\ &\text{s.t.} \quad \int_0^\alpha (1+\tau_x) p_i x_i di + \int_\alpha^1 (1+\tau_y) p_i y_i di + c = m. \end{aligned}$$

where x_i are the traditional goods and y_i are the (potential) new goods. We assume (later) $\epsilon > 1$, that is traditional and new goods are substitutes to capture that both goods provide

[3]This is the utility perceived by the household, that is s/he maximizes this function. The true utility function could be different (Farhi and Gabaix, 2020). Our positive and not normative focus does not allow us to take a stand on whether the true and the experienced utility functions are different or not.

similar services to consumers while one has negative externalities and the other does not. The price of the traditional good x_i is p_i with quality q_i. The price of the new good y_i is p_i with quality q_i. Traditional goods are taxed at rate τ_x while new goods are taxed at rate τ_y. The price of the numeraire good c is normalized to one and m are the available household resources. The assumption of a representative household is standard in the innovation literature and allowing for a (realistic amount of) heterogeneity would complicate our analysis but not affect our conclusions.

Household demand Before formulating the firm problem, we first have to solve the maximization problem to derive households demand functions. Denote therefore aggregated consumption of the $x-y$ goods as

$$Z=\left(\int_0^{\alpha} q_i x_i^{\frac{\epsilon-1}{\epsilon}}\,di+\int_{\alpha}^{1} q_i y_i^{\frac{\epsilon-1}{\epsilon}}\,di\right)^{\frac{\epsilon}{\epsilon-1}}$$

so that the first-order conditions (FOC) are:

$$\begin{aligned}/\partial x_i :&Z^{-\sigma}Z^{\frac{1}{\epsilon}}q_i x_i^{-\frac{1}{\epsilon}}=\left(1+\tau_x\right)p_i,\\ /\partial y_i :&Z^{-\sigma}Z^{\frac{1}{\epsilon}}q_i y_i^{-\frac{1}{\epsilon}}=\left(1+\tau_y\right)p_i.\end{aligned}$$

This implies that

$$y_i^{-\frac{1}{\epsilon}}=\frac{\left(1+\tau_i\right)p_i}{q_i}Z^{-\frac{1}{\epsilon}}Z^{\sigma}$$

such that the the demand function for y_i is

$$y_i=\left(\frac{\left(1+\tau_y\right)p_i}{q_i}\right)^{-\epsilon}Z^{1-\sigma\epsilon}\doteq\overline{p}_i^{-\epsilon}Z^{1-\sigma\epsilon},$$

where aggregate quality adjusted demand Z is taken as given by firm i and we define the adjusted price

$$\bar{p}_i = \frac{(1+\tau_y)\, p_i}{q_i}.$$

Since $\epsilon > 0$, the demand for good y_i is decreasing in the adjusted price $\bar{p}_i$, implying that it is decreasing in the tax τ_y and the price p_i and increasing in the quality q_i.

Firm Problem First, given the quality q_i, firm i solves the following maximization problem choosing optimal price (and therefore demand), given the demand function, τ_i and Z:

$$\max_{p_i} p_i y_i (p_i, q_i) - \psi y_i (p_i, q_i).$$

Second, firm i decides whether to innovate or not. Following Acemoglu and Zilibotti (1997), we assume that the cost of innovation for product $i \in [\alpha, 1]$ is an increasing function of i:

$$f_i = d_f (i - \alpha),$$

and the outcome of innovation is that the quality increases from q to λq, where λ is the step of quality improvement. This cost function captures the fact that some products are more expensive to innovate, and the aggregate innovation continuously increases if the profits of innovative firms increase due to favorable changes in innovation subsidy, tax policies or demand.

We first solve the firm's optimal choice of the price p_i and then derive the innovation decision through comparing the costs and benefits of innovation. First, the optimal pricing problem of firm i with quality q_i,

$$\pi(q_i) = \max_{p_i} p_i \left(\frac{(1+\tau_i)\, p_i}{q_i}\right)^{-\epsilon} Z^{1-\sigma\epsilon} - \psi \left(\frac{(1+\tau_i)\, p_i}{q_i}\right)^{-\epsilon} Z^{1-\sigma\epsilon},$$

gives the FOC

$$y_i + (p_i - \psi)(-\epsilon)\frac{y_i}{p_i} = 0$$

which is equivalent to the standard markup rule

$$p_i = \frac{\epsilon}{\epsilon - 1}\psi,$$

where $\varepsilon = -\frac{dy_i/y_i}{dp_i/p_i}$ is the price elasticity of demand for y_i, capturing the percentage demand change of y_i in response to a price change p_i. Then we can calculate the profit given the optimal price:

$$\begin{aligned}\pi(q_i) &= (p_i - \psi)\, y_i \\ &= \frac{1}{\epsilon - 1}\psi\left(\frac{(1+\tau_i)\frac{\epsilon}{\epsilon-1}\psi}{q_i}\right)^{-\epsilon} Z^{1-\sigma\epsilon} \\ &= d_\pi (1+\tau_i)^{-\epsilon} q_i^{\epsilon} Z^{1-\sigma\epsilon},\end{aligned}$$

where $d_\pi = \epsilon^{-\epsilon}(\epsilon - 1)^{\epsilon-1}\psi^{1-\epsilon}$ is a constant. The profit evaluated at the optimal price is positively related to the quality q_i, negatively related to the tax τ_i, and ambiguous in the aggregate quality adjusted demand Z (positive if σ is close to zero and utility function is almost linear in Z, as in the classic innovation literature, and negative if σ is large).

Second, the firm decides whether to innovate by comparing the cost and benefit:

$$\begin{aligned}\pi(\lambda q) - d_f(i - \alpha) \geq \pi(q) \Rightarrow \\ d_\pi (1+\tau_i)^{-\epsilon} Z^{1-\sigma\epsilon}(\lambda - 1)^{\epsilon} q^{\epsilon} \geq d_f(i - \alpha) \Rightarrow \\ i \leq \alpha + d_I (1+\tau_i)^{-\epsilon} Z^{1-\sigma\epsilon}\end{aligned}$$

where $d_I = d_\pi (\lambda - 1)^\epsilon q^\epsilon / d_f$ is a constant. We find that the innovation incentives are negatively related to the tax τ_i and ambiguous in the aggregate quality adjusted demand Z.

We can summarize the solution for firm i producing y_i conditional on Z. It charges a price

$$p_i = \frac{\epsilon}{\epsilon - 1}\psi,$$

the innovation choice leads to qualities

$$q_i = \begin{cases} \lambda q & \text{if } i \le \alpha + d_I (1+\tau_i)^{-\epsilon} Z^{1-\sigma\epsilon} \\ q & \text{if } i > \alpha + d_I (1+\tau_i)^{-\epsilon} Z^{1-\sigma\epsilon} \end{cases} \tag{1}$$

and the demand for good y_i is

$$\begin{aligned} y_i &= \left(\frac{(1+\tau_i)\, p_i}{q_i}\right)^{-\epsilon} Z^{1-\sigma\epsilon} \\ &= \begin{cases} \left(\frac{(1+\tau_i)\frac{\epsilon}{\epsilon-1}\psi}{\lambda q}\right)^{-\epsilon} Z^{1-\sigma\epsilon} & \text{if } i \ge \alpha + d_I (1+\tau_i)^{-\epsilon} Z^{1-\sigma\epsilon} \\ \left(\frac{(1+\tau_i)\frac{\epsilon}{\epsilon-1}\psi}{q}\right)^{-\epsilon} Z^{1-\sigma\epsilon} & \text{if } i < \alpha + d_I (1+\tau_i)^{-\epsilon} Z^{1-\sigma\epsilon}, \end{cases} \end{aligned} \tag{2}$$

A firm which produces good x_i also charges a price

$$p_i = \frac{\epsilon}{\epsilon - 1}\psi,$$

and produces

$$\left(\frac{(1+\tau_x)\frac{\epsilon}{\epsilon-1}\psi}{q_x}\right)^{-\epsilon} Z^{1-\sigma\epsilon}$$

with quality $q_i = q_x$.

Equilibrium Assume that the tax rate τ_y is the same for all y goods, and τ_x is the same for all x goods, then the aggregate innovation can be summarized by the innovation cutoff i_I:

$$q_i = \begin{cases} \lambda q & \text{if } i < i_I \\ q & \text{if } i \geq i_I. \end{cases}$$

where $i_I = \alpha + d_I \left(1 + \tau_y\right)^{-\epsilon} Z^{1-\sigma\epsilon}$.

3 Differential Fiscal Policy: Theory and Evidence

In this Section we first explore the theoretical implications of a differential tax and other regulatory policies on the consumption choices of households in stage 2. Note that our analysis is positive and not normative since we made no assumption on how agents' perceived utility is related to a social welfare function. In the second step we provide an overview of empirical work corroborating our theoretical findings.

3.1 Theory

Our theory shows that the aggregate consumption of the traditional goods equals

$$X := \alpha \left(\frac{(1 + \tau_x) \frac{\epsilon}{\epsilon - 1} \psi}{q_x} \right)^{-\epsilon} Z^{1-\sigma\epsilon}$$

and consumption of new goods equals

$$Y := (I^* - \alpha) \left(\frac{(1 + \tau_y) \frac{\epsilon}{\epsilon - 1} \psi}{\lambda q_y} \right)^{-\epsilon} Z^{1-\sigma\epsilon} + (1 - I^*) \left(\frac{(1 + \tau_y) \frac{\epsilon}{\epsilon - 1} \psi}{q_y} \right)^{-\epsilon} Z^{1-\sigma\epsilon},$$

where I^* is the innovation decision in the first state which is taken as given here. The quality is equal to q_x for all traditional goods and equal to q_y for all new goods. The relative consumption of new to traditional goods is thus equal to:

$$\frac{Y}{X} = \frac{(I^* - \alpha)\left(\frac{(1+\tau_y)\frac{\epsilon}{\epsilon-1}\psi}{\lambda q_y}\right)^{-\epsilon} + (1 - I^*)\left(\frac{(1+\tau_y)\frac{\epsilon}{\epsilon-1}\psi}{q_y}\right)^{-\epsilon}}{\alpha\left(\frac{(1+\tau_x)\frac{\epsilon}{\epsilon-1}\psi}{q_x}\right)^{-\epsilon}} \tag{3}$$

$$= \left(\frac{(I^* - \alpha)}{\alpha}\lambda^{\epsilon} + \frac{1 - I^*}{\alpha}\right)\left(\frac{1+\tau_y}{1+\tau_x}\frac{q_x}{q_y}\right)^{-\epsilon} \tag{4}$$

The substitution effects between X and Y consumption can be summarized through the wedge

$$\omega = \frac{1+\tau_y}{1+\tau_x}\frac{q_x}{q_y}.$$

An increase in the wedge leads to substitution from Y to X, and a decrease of the wedge induces the desired substitution from X to Y. The consequences for tax policy, which aims at decreasing the wedge, are clear. Both an increase in the tax on the traditional good and a decrease in the tax on the new good accomplish this. This is intuitive. If the traditional good gets more expensive relative to the new good then households substitute away. Similarly, an increase in the quality of the new good q_y relative to the traditional good quality q_x decreases the wedge and induces substitution. Example of relative quality increases are preferential parking for electric cars but not for combustion engines cars and allowing flavors, such as menthol, in e-cigarettes but not in traditional cigarettes.

Result 1. [Substitution through Regulation/policies: Substitution]

- *The substitution effects from regulation can be summarized through the wedge* $\omega = \left(\frac{1+\tau_y}{1+\tau_x}\frac{q_x}{q_y}\right)$

- *A decrease in the tax ratio* $\frac{1+\tau_y}{1+\tau_x}$ *induces a substitution from traditional to new goods.*

- *A decrease in the quality ratio* $\frac{q_x}{q_y}$ *induces a substitution from traditional to new goods.*

The interpretation of the results should be clear by now. All that matters is the difference induced through regulation and taxes between the two goods. From a theoretical perspective it is inessential in this simple framework what exactly causes this difference.

While the previous result is in line with standard intuition, we want to point out a potential caveat for implementing such a policy right away. While the substitution channel is straightforward, the total effect of a change in taxes or regulation is more involved since one has to consider the effect on the aggregate quality adjusted demand Z. Consider for example an increase in the tax τ_x on the traditional good. It is possible that the consumption of the new good shrinks although the tax increase leads to a substitution towards the new good, simply because the market size shrinks. Then consumption of both the traditional and new good fall, where the first falls more than the latter because of the substitution effect. A naive policy might therefore have counterproductive effects. In a different scenario, however, the increase in τ_x might lead to an increase in Y. The increase in Y might even be so strong that the total demand $X + Y$ increases, that is the increase in Y is larger than the fall in X. Such a scenario would arise if the incentives to substitute between X and Y goods are strong - $\sigma\epsilon$ is large. Households might then substitute one X item for more than one Y item. For example, households might switch from medium-sized combustion engine cars to larger electric cars. Others might buy an electric car who did not own a car before.

A policy that is successful in increasing consumption of the new good Y and reducing consumption of the traditional good X might thus lead to an increase in total consumption $X+Y$. In the case of the tobacco market this would mean more alternative nicotine products are consumed, fewer traditional cigarettes are smoked but total nicotine consumption goes up. In the automobile market, we would observe more electric cars, fewer combustion engines cars but the total number of cars increasing. Theoretically, an increase in τ_x could lead to a decrease in X and at the same time to an increase in Y which outweighs the decrease in x

such that $X+Y$ increases.

Example *It is easy to find examples which feature this behavior. Fix parameters:* $\epsilon = 2$, $\alpha = 0.6$, $q = 1$, $\lambda = 2$, $\sigma = 3$ *and* $\psi = 1$ *and consider fixing* $i_I = I^* = 2/3$ *such that* $(i_I - \alpha)\lambda^\epsilon + 1 - i_I = \alpha$, *i.e.,* X *and* Y *sectors are of the same size. This can be achieved by adjusting innovation cost* d_f. *Setting* $\tau_x = 60\%$ *and* $\tau_y = 10\%$ *then* $Z = 0.7721$ *and* $X+Y = 5.3237$. *If* τ_x *is increased to* 65%, *while keeping* i_I *fixed, then* $Z = 0.7689$ *and* $X+Y = 5.3309$. *This means that aggregate quality adjusted demand* Z *falls while the total number of items sold,* $X+Y$, *increases.*

We summarize the caveats in a result.

Result 2. [Caveat: Total Effect]

- *An increase in* τ_x *or* τ_y *decreases* X *and* Y *consumption if* σ *is smaller than* $1/\epsilon$ *but bounded away from* 0.

- *An increase in* τ_x *decreases* X *and increases* Y *consumption if* σ *is large. And might even increase* $X+Y$.

- *An increase in* τ_y *decreases* Y *and increases* X *consumption if* σ *is large. And might even increase* $X+Y$.

There are two possible responses to this caveat. If the naive policy leads to a reduction in both X and Y, households utility falls. One policy option is to ensure that households' perceived utility is unchanged, that is Z is unaffected. Such a policy requires for example increasing τ_x and decreasing τ_y or increasing q_y at the same time. As a result consumption of X falls and Y consumption increases.

Result 3. [Effects of Regulation/policies: Total Effect - I]

If the change in taxes/regulation keeps the aggregate quality adjusted demand Z unchanged then

- *A decrease in the tax ratio $\frac{1+\tau_y}{1+\tau_x}$ induces a reduction in traditional and an increase in new goods.*

- *A decrease in the quality ratio $\frac{q_x}{q_y}$ induces a reduction in traditional and an increase in new goods.*

An alternative strategy would be to design the policy to keep the number of items sold, $X + Y$, unchanged. This means that the policy change ensures that the number of items sold does not change and each additional Y good sold means one less X good sold.

Result 4. [Effects of Regulation/policies: Total Effect - II]

A combination of changes in τ_x and τ_y can be designed which keeps $X + Y$ constant, decreases X and increases Y.

3.2 Empirical Results

This Section provides empirical evidence using three different products to substantiate our main theoretical result that differential taxation induces substitution from the traditional good to the new good. The three different product categories are characteristic of the effects of sin taxes. The product categories differ in the availability of a close substitute, where we emphasize again that being a close substitute is viewed from the consumer's perspective.

Tobacco taxation is a category with a close substitute since alternative nicotine delivery systems that are likely to be significantly less harmful exist

The second category has good substitutes, but the implemented policy design is suboptimal. Highly sweetened sodas have good substitutes, but the regulation was imperfect and had unintended substitutes, that is, unintended from a policy maker's perspective whose intention was calorie intake. But unsurprising given our theoretical framework which predicts that consumers substitute towards close, untaxed substitutes.

The third category is one without substitutes or only imperfect ones. For cars, CO_2 emission has no proper substitutes as alternative engines are not quite ready to be substitutes for combustion engines. For electricity, the substitutes for coal are imperfect or not carbon-free.

3.2.1 Taxation of Sugar-sweetened Beverages

We first consider sugar-sweetened beverages and their taxation. This is a type of sin taxation that leads to substitution effects as described by our theory. However, at the same time, the tax is not ideally designed, and the substitution effects are therefore not as desired.

Taxes on sugar-sweetened beverages have been introduced in several countries including, the USA, Denmark and France. From a methodological perspective the US is the most attractive country. Taxes in the US are introduced at a local level, for example at the city level in Philadelphia or Berkeley, such that neighboring areas without a tax can serve as a control group. Comparing the consumption response in Philadelphia and its neighboring areas will then identify the causal effect of introducing a tax on soda. Obviously, this is not possible if the tax had been introduced nationwide since there would be no control group.

Seiler et al. (2019) use a panel dataset covering beverage sales in several hundred stores in Philadelphia, including small convenience and wholesale stores. The methodology used is a difference-in-difference approach. This methodology compares the consumption before and after the introduction of the tax in Philadelphia. However, this is not sufficient since consumption might have changed for other reasons unrelated to the tax increase, for example a trend towards more healthy behavior. The authors therefore compare the Philadelphia consumption response to the response of stores located at least 6 miles outside of Philadelphia. The (standard) assumption is that the health trend is the same in Philadelphia as in its suburbs and that the tax change is the only relevant difference between the two regions over time.[4]

The Philadelphia tax covers all sugar-sweetened beverages except for sweetened milk products and 100 percent fruit juices. In particular, diet drinks are also taxed. For the interpretation of the results it is important to note that there is not full pass-through, meaning that the tax incidence is not fully on customers.

The main finding in Seiler et al. (2019) is that the sugar-sweetened beverages sales dropped by 46% after levying the tax. However, there is substantial tax avoidance as purchases outside the city of Philadelphia increased at the same time, implying a net decrease in Philadelphia and its surrounding area of 22%. Roberto et al. (2019) estimate a 38% quantity reduction after taking into account cross-shopping, which is larger and statistically different from Seiler et al. (2019). The main reason seems to be that Roberto et al. (2019) use a non-representative sample. Indeed, when Seiler et al. (2019) estimate a 36% quantity reduction when restricting to the non-representative sample in Roberto et al. (2019).

While a 22% drop due to the tax increase appears to be a success, taking a closer look at the data reveals some unpleasant details. Consumers tended to cut down on drinks at the lower end of the calorie spectrum, such as sports drinks, but did not change their consumption of bottled water. Sales of sports drinks, which are taxed but relatively low-calorie, decreased by about 36%. However, the sales of "natural" juices increased by 9%. These drinks have a high sugar content but are exempted from the tax since the sugar is not added and the drinks do not contain any sweeteners.

The aim of the tax is to reduce the number of calories consumed. The success in this respect is limited. Counting the total change in calories shows a fall by 16%, but this change is not statistically significant. In other words the hypothesis that the tax had no effect on

[4]Note that this assumption does not preclude permanent differences between the surrounding counties and Philadelphia.

calorie consumption can not be rejected.

What does this mean in light of our theory? We interpret the results as supporting our idea that people respond to tax incentives and change their behavior accordingly. As predicted by our theory, they substitute away from taxed goods towards non-taxed goods. Or if both goods are taxed, households substitute towards the good with higher sugar content. The lesson of this natural tax experiment is that changing customers' behavior requires a carefully designed tax system. Otherwise the result is a convolution of desired and undesired substitution effects. Instead of taxing drinks with high and low amounts of added sugar at the same rate and exempting other sugary drinks, sugar should be taxed independently of the type of consumption good. This would avoid the undesired substitution behavior observed in Philadelphia; consumers would substitute from high to low sugar goods and the calorie reduction objective is more likely to be met.

3.2.2 Taxation of Novel Tobacco/Nicotine Products

The nascent literature on the taxation and regulation of traditional cigarettes and e-cigarettes is growing fast.[5] This is hardly surprising since the cigarette market provides a good laboratory to test the effects of taxes on substitution, tax incidence and tax evasion simply because traditional cigarettes and e-cigarettes are quite close substitutes, particularly because they both provide nicotine to the consumers. We are, however, not assessing how less harmful e-cigarettes and other nicotine products are compared to than traditional cigarettes. We are only interested in the substitution effects due to tax changes.

Minnesota was the first state in the USA to introduce a tax on e-cigarettes. On August 1, 2010 Minnesota introduced an initial tax rate of 35 percent, followed by another 60 percentage points on July 1, 2013, increasing the total tax rate by 95 percent. Saffer et al. (2019) follow Abadie et al. (2010) and apply a synthetic control design to measure the causal effect of this introduction of a tax on e-cigarettes on smoking behavior. The idea is to compare

[5] For example Cotti et al. (2020), DeCicca et al. (2020), Pesko et al. (2019), Buckell et al. (2017), Dave et al. (2017) Kenkel et al. (2017)

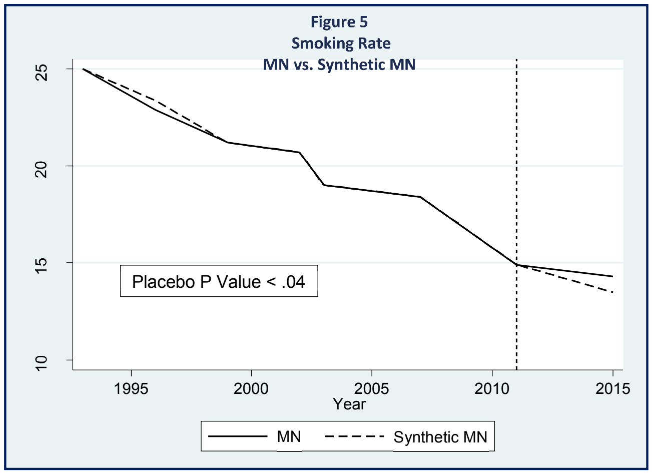

Figure 1: Figure 5 from Saffer et al. (2019): Smoking behavior in Minnesota and the rest of the US

the smoking behavior in Minnesota and the rest of the USA using a difference-in-difference approach. The key assumption of the difference-in-difference approach is that Minnesota and the rest of the USA behaved similarly before the tax hike and would have continued to be similar in the absence of a tax increase (counterfactual scenario). This means one has to ensure to select the control group of states such that they represent a valid counterfactual for Minnesota.

Saffer et al. (2019) follow Abadie et al. (2010) and use the synthetic control design to ensure that Minnesota and the control group of states feature the same pre-trend, that is they look similar before the tax introduction. The assumption is that the similarity of the pre-trend would have continued in the absence of the tax. The effect of the tax is then the difference in the behavior of Minnesota and the control group after the tax increase. Saffer et al. (2019) also make the reasonable assumption that the introduction of the tax on e-cigarettes was exogenous to the evolution of smoking behavior in Minnesota.

Figure 1 illustrates the approach. Before the tax increase, the treatment group - Minnesota - and the control group look identical in terms of smoking rate. After the introduction of the tax on e-cigarettes the smoking rate decreased more in the control group than in Minnesota.

To relate this result to the degree of substitution, note that the relative price of e-cigs versus cigarettes in 2012 was almost identical in Minnesota and in the control group (0.55 vs. 0.56). After the tax increase the relative price in Minnesota increased to 0.61 while it fell to 0.52 in the control group, that is the relative price rose by 17 percent. The result of this change in relative prices was about an 0.9 percentage points increase in smoking prevalence, equivalent to a 5.4 percent prevalence increase relative to the pre-tax increase level in Minnesota.

The findings in Saffer et al. (2019) also suggest that almost all of the increase in smoking prevalence in Minnesota is accounted for by a decrease in successful quits and there is no change in the number of cigarettes consumed by everyday smokers. The effects on adolescents are unclear and the gateway hypothesis is not explored in this paper.

To summarize, Saffer et al. (2019) use state-of-the-art methodology and find strong support for substitution between e-cigarettes and traditional cigarettes in response to a relative price increase induced by a tax increase in e-cigarettes. The substitution effects are presumably strong since both products provide the essential good desired by the consumer, nicotine, the ingredient that generates the addiction.

Cotti et al. (2020) reaches a similar conclusion, albeit using a somewhat weaker methodology. While these authors add a rich set of fixed effects which allows controlling for a large amount of heterogeneity, their methodology does not consider potential differences in pre-trends. Although controlling for fixed effects is an important step in avoiding biases, differences in pre-trends might lead to biased results. The benefit of this study is that it uses the large Nielsen Retail Scanner data from 2011 to 2017, which comprises approximately 35,000 retailers. Cotti et al. (2020) evaluates all tax changes between 2011 and 2017 whereas Saffer et al. (2019) only uses a single state, Minnesota.

They estimate the e-cigarette own-price elasticity to be -1.5, that is the demand for e-cigarettes falls if the price of e-cigarettes increases. The estimated cross-price elasticities of demand between e-cigarettes and traditional cigarettes shows that the two products are substitutes. An increase in the price of e-cigarettes due to an increase in the tax on e-cigarettes leads to an increase in the demand for traditional cigarettes. And vice versa, an increase in the price of traditional cigarettes increases the demand for e-cigarettes.

Both papers - Saffer et al. (2019) and Cotti et al. (2020) - thus agree on traditional and e-cigarettes being (quite strong) substitutes.

3.2.3 Taxation of CO_2 emissions

For the third type of product, CO_2 taxation, the available set of substitutes is limited, implying that the scope for taxes to induce substitution is also limited.

One stark example is fuel taxation in Germany. One liter of fuel leads to 2.37 kilogram of CO_2 so that a ton of CO_2 corresponds to the burning of 422 liters. This current taxation of fuel in Germany thus equates to a tax of about 275 Euros per ton of emitted CO_2. Apparently, this high CO_2 taxation has not led to a strong substitution from combustion engines to emission free engines. Apparently because emission free cars are not considered close enough substitutes to conventional combustion engines cars. We discuss in Section 4 a potential solution for overcoming the lack of a close enough substitute.

The electricity market looks more promising albeit far from being ideal in terms of available substitutes. In 2013, Great Britain implemented a carbon tax—the Carbon Price Support (CPS)— without coordinating with its Europeans neighbors who did not implement such a tax. The UCL report on international electricity trading (2019) investigated the impact of the CPS on Great Britain's carbon emissions and the potential impact on cross-border electricity trading with France and the Netherlands. The report finds that the implementation of a carbon tax caused a reduction in carbon emissions in the electricity sector, due to a large reduction in coal generation. The numbers are quite impressive. Coal usage in electricity generation in Great Britain fell from 40% in 2013 to 3% in 2019. This raises the question about substitutes. The UCL report provides an answer that appears less pleasant than the large reduction in coal usage suggests. The coal generated electricity was replaced by imported (non-carbon free) electricity from France and the Netherlands. This supports our theory that this taxation works because a substitute exists. The tax change was large enough to induce the substitution away from coal to foreign produced electricity. However, carbon taxation in the UK is far from what our theory suggests a superior strategy to be since carbon taxes are not equalized across different carbon emission sources.

Since the US does not have a CO_2 tax, Cullen and Mansur (2017) exploit a quasi-natural experiment in the US. There was a significant variation in natural gas prices as a result of the shale revolution which rendered natural gas extraction cheaper and thus increased the supply of gas which could not be exported. Cullen and Mansur (2017) assess how the price decrease affects the electricity sector's carbon emissions and use these results to infer how a price on CO_2 would affect emissions. Cullen and Mansur (2017) find that carbon prices have a small effect on emissions. A price of \$60 per ton of carbon dioxide will reduce emissions only by 10%.

4 Fiscal Forward Guidance

Differential regulation is successful if consumers perceive the Y good to be a close substitute for the X good. Without such a substitute taxing for example the X good will just reduce consumption of this good without leading to a higher Y consumption, simply because households do not assign a high value to good Y. In this Section we show how fiscal policy can be designed to overcome this problem. We show that a smart regulation can induce innovation in the new good Y today so that future consumers can choose between the X good and the new close substitute Y. With innovation, differential taxation in the future will lead to the desired consumer switching from X to Y goods.

The idea to spur innovation is quite simple and is related to what central banks do all the time worlwide (Plosser, 2013). It has been recognized for forward guidance the communication of future policy actions is a key ingredient of a successful monetary policy (Plosser, 2013). This is hardly surprising since monetary policy aims (among other things) at increasing current investment and central banks understand very well that current investment decisions depend on future profits. The same logic not only applies to monetary but also to fiscal policy. Using forward guidance in fiscal policy gives clear guidance for policy to foster innovation. Differential taxation and regulation in the future (not necessarily now) such that the future market size for Y increases, will lead firms today to engage in more innovative activity. The idea is the same as the one used by all central banks. Firms invest today since they expect profits in the future. Here, for firms to invest in Y, they must expect the demand for Y to be sufficiently high in the future. An important insight in the innovation literature is that profits are proportional to market size since the profit per unit sold is bounded. Thus, the larger the market is, the higher profits are, and the higher current investment in the new technology will be. The new role of fiscal policy envisaged here is thus twofold. First, use differential regulation to induce substitution towards Y to increase market size of Y. Second, use this policy in the future, not necessarily now. Using differential taxation right away is only sensible if a substitute as a result of past innovations already exists.

A straightforward implementation is increasing τ_x while decreasing τ_y to keep aggregate quality adjusted demand Z unchanged. This result follows from our derivations of firms' optimal innovation decision,

$$I^* = i_I = \alpha + d_I \left(1 + \tau_y\right)^{-\epsilon} Z^{1-\sigma\epsilon}.$$

Increasing τ_x decreases Z while decreasing τ_y lowers Z to keep it constant such that the optimal innovation decision I^* increases because it is decreasing in τ_y. The intuition is easy. Market size Y increases since first aggregate quality adjusted demand Z is constant and the differential policy induces a switch from X to Y.

Result 5. *Increasing $\frac{\tau_x}{\tau_y}$ while keeping future aggregate quality adjusted demand Z constant leads to more innovation in the new good.*

Keeping Z constant is not necessary for fiscal forward guidance to work and other options are also available. As in the previous section we want to point out however, that some differential policies might have unintended side effects.

One example is an increase in τ_x without taking any measures to make Y more attractive. The logic is simple and builds on Result 2. An increase in τ_x always lowers the demand for X. To assess the effect on innovation, we need to know the response of the Y market. The larger the expansion of the Y market is, the larger the incentives to innovate are. The magnitude of the Y market expansion depends (among other things) on the parameter $\sigma\epsilon$, which governs how easily households will substitute between the two goods. If $\sigma\epsilon$ is small, there is little substitution and Y increases only slightly in response to a demand drop in X. As a result, innovative activity is low and may even fall. On the other hand, if $\sigma\epsilon$ is large, there is a strong substitution and a large increase in Y in response to a demand drop in X. As a result, innovative activity increases.

Result 6. - *An increase in τ_x decreases innovation I^* if $\sigma\epsilon$ is less than one but bounded away from* 0.

- *An increase in τ_x increases innovation I^* if $\sigma\epsilon$ is large.*

Another side effect of innovation inducing policies is, as discussed in Section 3.1, that such a policy might lead to an expansion or reduction of the total number of items sold, $X + Y$.

Example Innovation *A simple example illustrates that an increase in τ_x which increaeses innovations could increase or decrease $X + Y$. We again fix the parameters $\epsilon = 2$, $\alpha = 0.6$, $q = 1$, $\lambda = 2$, $\psi = 1$. Set $\tau_x = 60\%$, $\tau_y = 10\%$ and $d_f = 11.2969$. In this example we allow for an endogenous innovation response in stage 1. We consider two cases, $\sigma = 3$ and $\sigma = 4$. In both cases innovation activity increases if τ_x is increases from 60% to 65%. However in the first case, $\sigma = 3$, $X + Y$ decreases. On the other hand in the second case, $\sigma = 4$, $X + Y$ increases.*

This reveals an important trade-off for politicians. Ex-ante when no substitutes exist and innovation is needed, politicians have an incentive to promise strong differential regulation in the future as this will expand the Y market and lead to more innovation today. Ex-post after significant innovative activities, politicians might be interested in reducing the size of the market $X+Y$, because they want to reduce nicotine consumption or the number of cars . At that point in time, because innovation investments are sunk, the market size reduction is not detrimental anymore. However, innovators anticipating this behavior from politicians will respond ex-ante and will reduce innovations. The politician or planner therefore needs some commitment power and needs to decide whether a larger market, a higher $X_1 + Y_1$ with a high Y_1 and low X_1 is better than a market with a smaller size $X_2 + Y_2$ but with $X_2 > X_1$, that is more harmful goods. The options are thus between more consumption but less harmful or less consumption but more harmful.

Our analysis in this paper not only reveals these trade-offs but could also guide the decision making. We conclude that a properly designed policy works.

Result 7. *Forward guidance (properly implemented) is effective and stimulates research in less harmful substitutes.*

5 Conclusion

It is now well understood that many consumption goods have externalities, such as CO_2 emission. How can governments respond to this challenge? One of two extreme responses would be to do nothing and let the private sector and the market deal with it. This strategy is unlikely to be successful since it is the nature of externalities that they are not fully taken into account by market participants without government intervention. The other extreme is to forbid consumption of goods with externalities altogether. This is also hardly optimal since households derive utility from consuming these goods.

This paper offers a less extreme alternative that could balance households desire for consumption and the government's desire to reduce externalities. The idea is to use differential taxation and regulation to induce households to switch from the traditional to the new good while at the same time preserving their utility from consumption.

Our paper offers several lessons for differential regulation.

a) Differential regulation and taxation work in theory and practice. Taxing the traditional good more and/or the new good less induces the desired substitution towards the new good.

b) The differential regulation has to be well designed. Taxing only the traditional good might lead to an increase or a decrease in total consumption in this market. If this is undesired, a combination of taxes/subsidies on the traditional and the new good is needed.

c) A prerequisite for such a policy is however the existence of a substitute good households perceive as yielding similar utility as the traditional good. Without a reasonable substitute taxing the traditional good is a mild form of forbidding its consumption, mainly for low-income households. Households just reduce consumption of the traditional good when its tax is increased and cannot switch to another good with similar perceived properties.

d) A commitment to implement a differential taxation/regulation framework in the future - conduct forward guidance - spurs innovation today. This innovation activity will provide the substitute in the future, which will enable the traditional good to be substituted with a less harmful good. It is important that if a substitute is currently unavailable, then differential taxation does not work today but a commitment to do so in the future will.

e) The politician faces a trade-off between two scenarios. Scenario 1 involves more innovation, less harmful goods, more harmless goods but involves an expansion of the market size. Scenario 2 involves less innovation, more harmful goods, fewer harmless goods but a contraction of the market size.

The economic approach proposed in this paper to deal with reducing consumption of sin goods seems promising. Using differential regulation will induce this substitution and will induce innovative activity to provide households with attractive less harmful or harmless alternatives. However, the policies have to be designed carefully to avoid undesired side effects which in the political process might easily be interpreted as a failure of the idea of differential regulation while in actuality it is just a bad implementation of this idea.

Bibliography

Abadie, A., A. Diamond, and J. Hainmueller (2010): "Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California's Tobacco Control Program," *Journal of the American Statistical Association*, 105, 493–505.

Acemoglu, D. and F. Zilibotti (1997): "Was Prometheus unbound by chance? Risk, diversification, and growth," *Journal of Political Economy*.

Aghion, P. and P. Howitt (1992): "A Model of Growth Through Creative Destruction," *Econometrica*, 60, 323.

Buckell, J., J. Marti, and J. L. Sindelar (2017): "Should Flavors be Banned in E-cigarettes? Evidence on Adult Smokers and Recent Quitters from a Discrete Choice Experiment," Working Paper 23865, National Bureau of Economic Research.

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," Working Paper 26724, National Bureau of Economic Research.

Cullen, J. A. and E. T. Mansur (2017): "Inferring Carbon Abatement Costs in Electricity Markets: A Revealed Preference Approach Using the Shale Revolution," *American Economic Journal: Economic Policy*, 9, 106–33.

Dave, D., B. Feng, and M. F. Pesko (2017): "The Effects of E-Cigarette Minimum Legal Sale Age Laws on Youth Substance Use," Working Paper 23313, National Bureau of Economic Research.

DeCicca, P., D. S. Kenkel, and M. F. Lovenheim (2020): "The Economics of Tobacco Regulation: A Comprehensive Review," Working Paper 26923, National Bureau of Economic Research.

Farhi, E. and X. Gabaix (2020): "Optimal Taxation with Behavioral Agents," *American*

Economic Review, 110, 298–336.

Grossman, G. and E. Helpman (1991): "Trade, Knowlege Sprillovers, and Growth," *European Economic Review*, 17.

Kenkel, D. S., S. Peng, M. F. Pesko, and H. Wang (2017): "Mostly Harmless Regulation? Electronic Cigarettes, Public Policy and Consumer Welfare," Working Paper 23710, National Bureau of Economic Research.

Pesko, M. F., C. J. Courtemanche, and J. C. Maclean (2019): "The Effects of Traditional Cigarette and E-Cigarette Taxes on Adult Tobacco Product Use," Working Paper 26017, National Bureau of Economic Research.

Plosser, C. I. (2013): "Forward Guidance," Prepared for Stanford Institute for Economic Policy Research's (SIEPR) Annual Meeting, Stanford, CA, February 12, 2013.

report on international electricity trading, U. (2019): "The value of international electricity trading," Report commissioned by the office of gas and electricity markets (ofgem), UCL Institute for Sustainable Resources, g. Castagneto Gissey, B. Guo, D. Newbery, G. Lipman, L. Montoya, P. Dodds, M. Grubb, P. Ekins.

Roberto, C. A., H. G. Lawman, M. T. LeVasseur, N. Mitra, A. Peterhans, B. Herring, and S. N. Bleich (2019): "Association of a Beverage Tax on Sugar-Sweetened and Artificially Sweetened Beverages With Changes in Beverage Prices and Sales at Chain Retailers in a Large Urban Setting," *JAMA*, 321, 1799–1810.

Saffer, H., D. L. Dench, M. Grossman, and D. M. Dave (2019): "E-Cigarettes and Adult Smoking: Evidence from Minnesota," Working Paper 26589, National Bureau of Economic Research.

Seiler, S., A. Tuchman, and S. Yao (2019): "The Impact of Soda Taxes: Pass-Through, Tax Avoidance, and Nutritional Effects," Research Paper No. 19-12, Stanford University Graduate School of Business.

APPENDIX

Proof of Result 2

Substituting the expressions of x_i and y_i into Z, we obtain:

$$
\begin{aligned}
Z &= \left(\int_0^{\alpha} q_i x_i^{\frac{\epsilon-1}{\epsilon}} di + \int_{\alpha}^{1} q_i y_i^{\frac{\epsilon-1}{\epsilon}} di\right)^{\frac{\epsilon}{\epsilon-1}} \\
&= \left(\int_0^{\alpha} q^{\epsilon} \left((1+\tau_x)\frac{\epsilon}{\epsilon-1}\psi\right)^{-(\epsilon-1)} di + \int_{\alpha}^{I^*} (\lambda q)^{\epsilon} \left((1+\tau_y)\frac{\epsilon}{\epsilon-1}\psi\right)^{-(\epsilon-1)} di\right. \\
&\quad \left. + \int_{I^*}^{1} q^{\epsilon} \left((1+\tau_i)\frac{\epsilon}{\epsilon-1}\psi\right)^{-(\epsilon-1)}\right)^{\frac{\epsilon}{\epsilon-1}} Z^{\frac{(1-\sigma\epsilon)(\epsilon-1)}{\epsilon}\frac{\epsilon}{\epsilon-1}} \\
&= q^{\frac{\epsilon^2}{\epsilon-1}} \left(\frac{\epsilon}{\epsilon-1}\psi\right)^{-\epsilon} Z^{1-\sigma\epsilon} \left(\alpha(1+\tau_x)^{-(\epsilon-1)} + ((I^*-\alpha)\lambda^{\epsilon} + 1 - I^*)((1+\tau_y))^{-(\epsilon-1)}\right)^{\frac{\epsilon}{\epsilon-1}}.
\end{aligned}
$$

We can substitute $I^* = \alpha + d_I(1+\tau_y)^{-\epsilon} Z^{1-\sigma\epsilon}$ to simplify $(I^*-\alpha)\lambda^{\epsilon} + 1 - I^*$ into

$$
\begin{aligned}
(I^*-\alpha)\lambda^{\epsilon} + 1 - I^* &= d_I(1+\tau_y)^{-\epsilon} Z^{1-\sigma\epsilon}\lambda^{\epsilon} + 1 - \alpha - d_I(1+\tau_y)^{-\epsilon} Z^{1-\sigma\epsilon} \\
&= d_I(1+\tau_y)^{-\epsilon} Z^{1-\sigma\epsilon}(\lambda^{\epsilon}-1) + 1 - \alpha.
\end{aligned}
$$

Then we can further simplify Z as

$$
Z = q^{\frac{\epsilon^2}{\epsilon-1}} \left(\frac{\epsilon}{\epsilon-1}\psi\right)^{-\epsilon} Z^{1-\sigma\epsilon} \left(\alpha(1+\tau_x)^{-(\epsilon-1)} + \left(d_I(1+\tau_y)^{-\epsilon} Z^{1-\sigma\epsilon}(\lambda^{\epsilon}-1) + 1 - \alpha\right)(1+\tau_y)^{-(\epsilon-1)}\right)^{\frac{\epsilon}{\epsilon-1}}, \tag{A1}
$$

which implies

$$
Z^{\sigma\epsilon} = q^{\frac{\epsilon^2}{\epsilon-1}} \left(\frac{\epsilon}{\epsilon-1}\psi\right)^{-\epsilon} \left(\alpha(1+\tau_x)^{-(\epsilon-1)} + \left(d_I(1+\tau_y)^{-\epsilon} Z^{1-\sigma\epsilon}(\lambda^{\epsilon}-1) + 1 - \alpha\right)(1+\tau_y)^{-(\epsilon-1)}\right)^{\frac{\epsilon}{\epsilon-1}}, \tag{A2}
$$

and

$$Z = q^{\frac{\epsilon}{\sigma(\epsilon-1)}} \left(\frac{\epsilon}{\epsilon - 1}\psi\right)^{-\frac{1}{\sigma}} \left(\alpha\left(1+\tau_x\right)^{-(\epsilon-1)} + \left(d_I\left(1+\tau_y\right)^{-\epsilon} Z^{1-\sigma\epsilon}\left(\lambda^\epsilon - 1\right) + 1 - \alpha\right)\left(1+\tau_y\right)^{-(\epsilon-1)}\right)^{\frac{1}{\sigma(\epsilon-1)}}. \tag{A3}$$

Consider the impact of increasing τ_x on Z. Take log in equation A2, we obtain:

$$\log Z = \log\left(q^{\frac{\epsilon}{\sigma(\epsilon-1)}} \left(\frac{\epsilon}{\epsilon - 1}\psi\right)^{-\frac{1}{\sigma}}\right) + \frac{1}{\sigma\left(\epsilon-1\right)} \log\left(\alpha\left(1+\tau_x\right)^{-(\epsilon-1)} + \left(d_I\left(1+\tau_y\right)^{-\epsilon} Z^{1-\sigma\epsilon}\left(\lambda^\epsilon - 1\right) + 1 - \alpha\right.\right.$$

Denote $\alpha\left(1+\tau_x\right)^{-(\epsilon-1)} + \left(d_I\left(1+\tau_y\right)^{-\epsilon} Z^{1-\sigma\epsilon}\left(\lambda^\epsilon-1\right)+1-\alpha\right)\left(1+\tau_y\right)^{-(\epsilon-1)}$as $D\left(Z,\tau_x\right)$, and then take derivatives w.r.t. τ_x:

$$\frac{dZ/d\tau_x}{Z} = \frac{1}{\sigma\left(\epsilon-1\right)} \frac{-\left(\epsilon-1\right)\alpha\left(1+\tau_x\right)^{-\epsilon} + dZ/d\tau_x\left(1-\sigma\epsilon\right)d_I\left(1+\tau_y\right)^{-\epsilon} Z^{-\sigma\epsilon}\left(\lambda^\epsilon-1\right)\left(1+\tau_y\right)^{-(\epsilon-1)}}{D},$$

which can be further simplified as:

$$\sigma\frac{dZ}{d\tau_x} = \left(-\alpha\left(1+\tau_x\right)^{-\epsilon} + \frac{dZ}{d\tau_x}\frac{1-\sigma\epsilon}{\epsilon-1}d_I\left(1+\tau_y\right)^{-\epsilon} Z^{-\sigma\epsilon}\left(\lambda^\epsilon-1\right)\left(1+\tau_y\right)^{-(\epsilon-1)}\right)\frac{Z}{D},$$

$$\left(\sigma - \frac{1-\sigma\epsilon}{\epsilon-1}d_I\left(1+\tau_y\right)^{-\epsilon} Z^{-\sigma\epsilon}\left(\lambda^\epsilon-1\right)\left(1+\tau_y\right)^{-(\epsilon-1)}\frac{Z}{D}\right)\frac{dZ}{d\tau_x} = -\alpha\left(1+\tau_x\right)^{-\epsilon}.$$

We can see that $dZ/d\tau_x < 0$ if and only if

$$\sigma - \frac{1-\sigma\epsilon}{\epsilon-1}d_I\left(1+\tau_y\right)^{-\epsilon} Z^{-\sigma\epsilon}\left(\lambda^\epsilon-1\right)\left(1+\tau_y\right)^{-(\epsilon-1)}\frac{Z}{D} > 0, \tag{A4}$$

and if $\sigma = 1/\epsilon$, i.e., $1-\sigma\epsilon = 0$, this condition is satisfied.

First, let us consider the case that σ is small, such that $1-\sigma\epsilon > 0$. Notice that the LHS of equation A4 is continuous in σ, so we know that if $\sigma < 1/\epsilon$ but not far, it is still possible that that expression is greater than 0 and $\frac{dZ}{d\tau_x} < 0$. Formally speaking, given any set of parameters, $\exists \upsilon$ such that $\forall \sigma \in (1/\epsilon - \upsilon, 1/\epsilon)$, $dZ/d\tau_x < 0$. In this region, as Z decreases, y_i and Y, which are proportional to $Z^{1-\sigma\epsilon}$, also decrease.

Second, let us consider the case that σ is large, such that $1-\sigma\epsilon<0$. In this case, the LHS of equation A4 is always positive and we know for sure that $dZ/d\tau_x<0$. Moreover, y_i and Y increase because $1-\sigma\epsilon<0$.

Proof of Result 3

First, a decrease in $\frac{1+\tau_y}{1+\tau_x}$ while Z staying constant, implies that τ_x increases and τ_y decreases, because from equation A3, we can see that if τ_x increases and Z stays constant, then it must be that τ_y decreases. In this case, innovation $I^*=\alpha+d_I\left(1+\tau_i\right)^{-\epsilon}Z^{1-\sigma\epsilon}$ increases. So X decreases, as we can see from

$$X:=\alpha\left(\frac{\left(1+\tau_x\right)\frac{\epsilon}{\epsilon-1}\psi}{q_x}\right)^{-\epsilon}Z^{1-\sigma\epsilon}, \tag{A5}$$

and Y increases, as we can see from

$$Y:=\left(I^*-\alpha\right)\left(\frac{\left(1+\tau_y\right)\frac{\epsilon}{\epsilon-1}\psi}{\lambda q_y}\right)^{-\epsilon}Z^{1-\sigma\epsilon}+\left(1-I^*\right)\left(\frac{\left(1+\tau_y\right)\frac{\epsilon}{\epsilon-1}\psi}{q_y}\right)^{-\epsilon}Z^{1-\sigma\epsilon}, \tag{A6}$$

given that $\lambda>1$.

Similarly, we can prove the results for $\frac{q_x}{q_y}$. If q_x decreases, to keep Z constant, then q_y must increase, if other policies stay constant, as we can see from the following expression of Z:

$$\begin{aligned} Z &= \left(\int_0^\alpha q_i x_i^{\frac{\epsilon-1}{\epsilon}}di+\int_\alpha^1 q_i y_i^{\frac{\epsilon-1}{\epsilon}}di\right)^{\frac{\epsilon}{\epsilon-1}} \\ &= \left(\alpha q_x^\epsilon\left(\left(1+\tau_x\right)\frac{\epsilon}{\epsilon-1}\psi\right)^{-(\epsilon-1)}+\left(\left(I^*-\alpha\right)\lambda^\epsilon+1-I^*\right)q_y^\epsilon\left(\left(1+\tau_y\right)\frac{\epsilon}{\epsilon-1}\psi\right)^{-(\epsilon-1)}\right)^{\frac{\epsilon}{\epsilon-1}}Z^{1-\sigma\epsilon}. \end{aligned}$$

Obviously, from equation A5 and A6, we can see that X decreases as q_x decrease, and Y increases as q_y increases.

Proof of Result 4

Denote the response of X to an increase of τ_x as $X_x\doteq dX/d\tau_x$, and similarly for Y and $X+Y$:

Y_x and XY_x. Moreover, denote the responses to τ_y as X_y, Y_y and XY_y. We can construct the policy change in the following way. First, consider increasing τ_x, and depending on whether $X+Y$ increase or decrease, choose to increase or decrease τ_y to offset the increase or decrease of $X+Y$. Then the following three cases can happen: (1) X decreases, Y increases; (2) X increases, Y decreases; (3) X and Y do not change. In case (1) we have already obtained the policy. In case (2), we can just use the opposite policy, i.e., decreasing τ_x and adjusting τ_y accordingly, and then we obtain the policy. Case (3) is not generic though and be ignored.

Proof of Result 5

First, if we increase τ_x, we can always reduce τ_y by a proper size to keep Z constant. Consider changing tax rates from τ_x, τ_y to $\tau_x + \triangle\tau_x, \tau_y - \triangle\tau_y$, then according to equation (A3), Z changes by $\triangle$:

$$\triangle = q^{\frac{\epsilon}{\sigma(\epsilon-1)}} \left(\frac{\epsilon}{\epsilon-1}\psi\right)^{-\frac{1}{\sigma}} \left(\alpha\left(1+\tau_x+\triangle\tau_x\right)^{-(\epsilon-1)} + \left(d_I\left(1+\tau_y-\triangle\tau_y\right)^{-\epsilon} Z^{1-\sigma\epsilon}\left(\lambda^{\epsilon}-1\right)+1-\alpha\right)\left(1+\tau_y-\right.\right.$$

We can see that if $\triangle\tau_y = 0$, then $\triangle < 0$, while if $\triangle\tau_y \to 1+\tau_y > 0$, then $\triangle \to +\infty$. By the intermediate value theorem, we know that $\exists\triangle\tau_y > 0$ such that $\triangle = 0$, i.e., Z stays unchanged.

From the expression of innovation cutoff

$$I^* = \alpha + d_I\left(1+\tau_y\right)^{-\epsilon} Z^{1-\sigma\epsilon}, \tag{A7}$$

we can see that when τ_y decreases while Z stays constant, I^* increases, i.e., innovation increases.

Proof of Result 6

Equation (A7) shows that innovation I^* is positively related to Z, so whether an increase in τ_x results in an increase or a decreases in I^* is the same as whether an increase in τ_x results in an increase or a decreases in Z, as stated in Result 2.