11-14-2013

Unintended Consequences of Cigarette Taxation and Regulation

Angela Hawken
Pepperdine University, angela.hawken@pepperdine.edu

Jonathan D. Kulick
Pepperdine University, jonathan.kulick@pepperdine.edu

James Prieger
Pepperdine University, james.prieger@pepperdine.edu

Follow this and additional works at: http://digitalcommons.pepperdine.edu/sppworkingpapers

Part of the Public Affairs, Public Policy and Public Administration Commons

Recommended Citation
Hawken, Angela; Kulick, Jonathan D.; and Prieger, James, "Unintended Consequences of Cigarette Taxation and Regulation" (2013). Pepperdine University, School of Public Policy Working Papers. Paper 47.
http://digitalcommons.pepperdine.edu/sppworkingpapers/47

This Article is brought to you for free and open access by the School of Public Policy at Pepperdine Digital Commons. It has been accepted for inclusion in School of Public Policy Working Papers by an authorized administrator of Pepperdine Digital Commons. For more information, please contact Kevin.Miller3@pepperdine.edu.
UNINTENDED CONSEQUENCES
OF CIGARETTE TAXATION AND REGULATION

Angela Hawken, Pepperdine University (corresponding author: ahawken@pepperdine.edu)
Jonathan Kulick, Pepperdine University
James Prieger, Pepperdine University

November 14, 2013
Abstract

Tobacco smoking harms health. Taxes and regulations can reduce that harm. But evasion reduces the efficacy of taxes and regulations and creates harms of its own in the form of illicit markets. Enforcement can reduce evasion but creates additional harms, including incarceration and violence. Peter Reuter has pointed out that a flat ban on cigarettes would be likely to generate illicit-market harms similar to the harms of existing illicit drug markets. Taxes and regulations can be thought of as “lesser prohibitions,” subject to the same sorts of risks. Minimizing total harm means minimizing the sum of abuse harms and control harms.

Tighter regulations and higher taxes on cigarettes risk increasing the size of the existing illicit tobacco markets, which are already substantial. That risk can be somewhat blunted by increasing enforcement effort, but doing so can be costly on several dimensions and might, under plausible assumptions, lead to an increase in violence. Tobacco policymaking should therefore consider illicit markets and the need for enforcement; some of the health benefits of regulation and taxation may be offset by increased illicit-market side effects and enforcement costs. The presence of licit substitutes, such as e-cigarettes, can greatly reduce the size of the problem; the regulation of e-cigarettes should take this effect into account. If enforcement is to be increased to counterbalance tightened controls, positive-feedback dynamics suggest that the enforcement increase should precede, rather than follow, the tightening.

1. Introduction

Cigarette smoking and other uses of tobacco damage health, and restrictions (taxes and regulations, or, in the extreme, prohibition) are intended to reduce that damage As Peter Reuter has pointed out (Reuter, 2013), full prohibition would be expected to have some of the bad effects of the prohibitions of other drugs: illicit markets and the costs of enforcement.

Regulations and taxes can be thought of as lesser prohibitions, and create to some extent similar opportunities for profitable evasion. Markets already exist for cigarettes produced and transported in violation of laws, including cigarettes licitly produced and sold in low-tax jurisdictions and smuggled into higher-tax jurisdictions.

Illicit transactions reduce what would otherwise be the efficacy of taxes and regulations in reducing consumption and thus protecting health, in addi-
Unintended Consequences of Cigarette Taxation and Regulation

tion to generating damage to health and other values on their own account: violence, corruption, incarceration, and impairment of collective social capital.

As stricter controls on cigarettes are implemented, we should expect to see an expansion of supply-side smuggling activity, counterfeiting, and tax avoidance, in addition to a wide variety of other demand-side price-minimization strategies to avoid the effects of any regulatory change. Optimal decisionmaking about taxes and regulations would weigh the health and other harms from illicit markets against the health gains from reduced smoking. In addition to choosing taxes and regulations, public authorities can control the consequences of violating those rules both by setting penalty levels and by choosing enforcement strategies and resource levels.

To the extent that other options are widely available to smokers, the public-health benefits of higher taxes and stricter regulations will be reduced, given that at least some smokers will turn to these strategies in order to minimize the impact of the tax increase on the price they pay for cigarettes. In the event of bans on particular types of cigarettes (e.g., flavored cigarettes of various kinds), some users of those types will turn to contraband product rather than (as intended) giving up smoking entirely. This paper will develop a model of the demand for contraband product in the event of a ban and the consequent level of crime that might follow.

Thus, in deciding whether to tighten controls on cigarettes in various ways, the question facing policymakers becomes: How much health benefit will a tighter rule in fact create, once the effects of evasion are considered, and would that health gain justify the increase in damage from criminal activity and enforcement?

2. Consequences of Stricter Regulation and Taxes

2.1. Smuggling, Tax Evasion, and Violence

Tobacco-tax evasion and smuggling is widespread. Joossens and Raw (2008, 2012) estimate that illicit tobacco transactions lead to $40 to $50 billion in lost tax revenue globally, while the U.S. Bureau of Alcohol, Tobacco, Firearms and Explosives (ATF) (2009) estimates that tobacco diversion costs over $5 billion in revenue from unpaid excise taxes annually in the United States. LaFaive and Nesbit (2013) report that the five smuggling-destination states with the highest cigarette-smuggling rates are New York (60.9 percent of the state’s total consumption), Arizona (54.4 percent), New Mexico (53.0 percent), Washington (48.5 percent), and Rhode Island (39.8 percent). Five source states had estimated net total smuggling exports that exceeded ten percent of total state consumption: Virginia, Delaware, West Virginia, Missouri and Wyoming.

The impact of evasion on cigarette consumption, and thus on health, is substantial, but its extent is controversial. Stehr (2005) estimates that up to 85
percent of the “tax paid sales response” to increases in cigarette excise taxes in the United States may be due to tax avoidance rather than reduced consumption. DeCicca, Kenkel, and Liu (2010) find that many U.S. cigarette taxes are above their revenue-maximizing level, resulting in an increase in tax avoidance, smuggling, and black-market selling. And Henchman and Drenkard (2013) find that smuggling rates generally rise in states after they adopt large cigarette-tax increases. However, in studies of global tobacco markets, Joossens, Ross, Merriman, and Raw (2009) claim that cigarette taxes have relatively little impact on the level of illicit trade, and Merriman, Yurekli, and Chaloupka (2000) conclude that higher cigarette taxes achieve both objectives of higher tax revenues and lower consumption.

LaFaive, Fleenor, and Nesbit (2008) divide cigarette smuggling into two types: “casual” smuggling, where consumers save money by buying cigarettes in lower-tax states or countries,1 and “commercial” smuggling, where larger-scale operators buy cigarettes in bulk in a lower-tax area and sell them tax-free in higher-tax areas. The U.S. Department of the Treasury (2010) finds that a significant component of illicit tobacco trade in the United States is the shipment of tobacco products from low-tax states to high-tax states to evade state taxes. Due to widely varying state taxes on wholesale tobacco products, ample opportunity exists to transport tobacco from a low-tax jurisdiction to a high-tax jurisdiction for retail sale or consumption to evade state taxes. Lovenheim (2008) finds that 13 to 25 percent of U.S. consumers purchase cigarettes in border locations, greatly reducing the potential health and revenue gains from cigarette taxation.

With black markets come organized crime and violence; New York City’s Finance Administrator described bootleg cigarettes as the “principal stoking facility of the engine of organized crime” (Fleenor, 2003, p. 7) and as constantly confronting workers with personal violence. LaFaive and Nesbit (2013) link bootleg cigarettes to an uptick in prison crime, while the Campaign for Tobacco-Free Kids (2008) details the impact illicit cigarette trafficking has on supporting terrorist organizations worldwide.

2.2. Smoker Responses

Smokers employ a variety of legal and illegal strategies to reduce the price they pay for cigarettes (Chaloupka, 2013; Xu et al., 2013). According to an analysis of the International Tobacco Control Policy Evaluation Four-Country Survey (Licht et al., 2011a), in response to increased cigarette taxes 8 percent of smokers reported buying from low or untaxed sources, 36 percent switched to discount or generic brands, and 14 percent used loose tobacco (“roll-your-own”). Licht et al. (2011b) report that smokers who use these price-

1 Such behavior is not necessarily illegal: small quantities of cigarettes imported for personal use are exempt from federal-excise-tax requirements, and most states exempt small quantities from use-tax requirements.
minimizing strategies are less likely to make attempts to quit smoking and less likely overall to succeed when they do attempt to quit; should tax increases further raise the retail price of cigarettes, these smokers would constitute a potential market for illicit product, depending on their comfort with engaging with a black market and on the level of enforcement.

Cigarette smokers can also switch from cigarettes to other tobacco products because of the lower taxes on the latter. Connolly and Alpert (2008) attributed nearly a third of the decline in legal cigarette sales in the United States over the previous decade to switching to tobacco products such as small cigars, moist snuff, and loose tobacco. Pomeroy (2012) reports that, due to the increase in federal taxes on cigarettes in 2009, monthly sales of pipe tobacco and large cigars skyrocketed over the next three years. Tynan, McAfee, Promoff, and Pechacek (2012) conclude that large tax differentials between cigarettes and other tobacco products lessen the impact of taxes on smoker behavior and public health.

3. Enforcement and Violence: What Do We Know?

Illicit drug markets are prone to violence (Andreas & Wallman, 2009). The value of illegal goods coupled with the lack of recourse in the legal system to settle conflicts creates inherent instabilities, uncertainties, and distrust in the market. This is exacerbated by illicit-market participants’ pre-existing experiences with violence; participants tend to be recruited from communities with above-average rates of violence (Moeller & Hesse, 2013). A number of explanations is offered for why stricter enforcement of illicit drug markets might increase violence, but two stand out in the literature. First, increased enforcement disrupts the market. Destabilizing established hierarchies renews competition and violence can follow as participants jostle for turf and market share (Costa Storti & De Grauwe, 2008; Papachristos, 2009; Rasmussen & Benson, 1994). Second, stricter enforcement increases the risk of detection and punishment, which in turn increases the risk premium and therefore profitability of sales (Kuziemko & Levitt, 2004). Profit margins become worth fighting for: Increasing the share of total cost attributed to enforcement risk can increase the incentive for violence insofar as violence deters enforcement agencies and potential informants (Caulkings, Kleiman, & Kulick, 2010; Kleiman, 2011).

Moeller and Hesse (2013) assert that the relationship between the intensity of drug-law enforcement and violence in drug markets follows from how enforcement affects the market’s structure. Low levels of enforcement tend to result in monopolistic markets, dominated by a few well-organized suppliers. They tend to be nonviolent, because violence attracts the attention of authorities. Reuter (1983) argues that well-behaved, concentrated industries such as these are unusual in illegal drug markets because pressure from law enforcement tends to stimulate competition among rivals.
The most extensive systematic review of the relationship between drug-law enforcement and violence (excluding police violence) is by Werb et al. (2011). We expand and update this review, but our inclusion criteria deviate from theirs in several ways:

1. Period of study: Our review includes studies published in the last 25 years (August 1988 to August 2013).
2. Type of publication: We restrict our review to studies published in peer-reviewed journals.
3. Qualifying methods: We include only studies that present the results of quantitative research (cross sectional or longitudinal). We exclude studies that exclusively report a theoretical model without analysis of empirical data on enforcement and violence, and we exclude studies that are limited to qualitative analysis.

Drug-market violence is taken to include homicides, assaults, and shootings that stem from the illegal market. Law-enforcement intensity is taken to include any reasonable proxy measure; these included number of drug arrests, expenditures on enforcement, drug-seizure rates, and number of police officers.

Fourteen studies meet our inclusion criteria, with considerable variation in methods, reporting periods, and proxy measures of enforcement and violence. These definitional differences preclude their enforcement-effect estimates being combined using traditional meta-analysis techniques. Even so, the weight of the evidence falls squarely on the side of a significant positive association between enforcement intensity and violence. Two of the studies reviewed (14 percent) found no association between levels of enforcement and violent crime; the remaining 12 studies (86 percent) found a positive association between enforcement and crime. A summary of the studies reviewed and a description of their findings are provided in the Appendix.

4. Theoretical Relationship between Enforcement and Violence

4.1. Overview

When considering a hypothetical illicit market, it is reasonable to expect there to be an inverted-U-shaped relationship between enforcement and violence. For lower levels of enforcement, increases in enforcement activity can be expected to lead to more violence. Systematic reviews of the relevant literature

---

2 We follow the search methods used by Werb and colleagues, which relied on conventional search techniques of English-language articles (see Werb et al. (2011) for an elaboration of search tools), and we use the same search terms (“violence,” “drug-related violence,” “drug-market violence,” “enforcement,” “drug gangs,” and “drug crime”). Our search yielded four studies meeting inclusion criteria that were not included in the earlier review.

3 Compared with fifteen that met the criteria of Werb et al. (2011).
on the relationship between enforcement and violence conclude that “disrupting drug markets [through enforcement] can paradoxically increase violence” (Werb et al., 2011, p. 87).

The expectation that enforcement leads to violence follows from what Goldstein (2003) terms the *systemic* aspect of the drug/violence nexus. Systemic violence stems from the “traditionally aggressive patterns of interaction within the system of [illicit] drug distribution and use” (p. 100). The causes of such violence include (inter alia) territorial disputes, homicides committed to enforce norms among the distribution hierarchy, the elimination of informers, punishment for selling fake or adulterated goods, and violence associated with robberies in areas known for drug activity. As enforcement increases, the stakes become higher for market participants and violence increases (Caulkins, Reuter, & Taylor, 2006). We formalize and explain how this occurs in our model below, but the main idea is that, as prices and revenues rise in the industry in response to increased enforcement, the attendant violence in the market can be expected to rise.

Reinforcing the systemic contribution to violence, the *economic compulsive* aspect of the drug/violence nexus (Goldstein, 2003) may reinforce the positive association between enforcement and violence. As enforcement increases and street prices for the illegal substance rise, users may turn to robbery or other crimes to raise cash to support their habit. This contributing factor may not be as strong for cigarettes as for currently illegal drugs, however. The third of Goldstein’s violence factors, the *psychopharmacological* aspect, is unlikely to contribute much. Irritability associated with withdrawal symptoms, which may be expected to increase at least temporarily due to higher prices from increased enforcement, has been associated with violent crime among opiate users (Goldstein, 1979). Again, however, this factor is likely to be negligible for tobacco users—indeed, for most drug markets, the violence attendant to illicit sales swamps drug-induced violence.

At levels of enforcement high enough to substantially shrink market volumes and total revenues—assuming that such levels are practically attainable—it seems clear that violence should begin to fall. With police presence on every street, intense scrutiny of every shipment of goods into the country, and complete control of the borders, the quantity of illicit goods transacted must surely fall to low levels. Apart from the deterrence and incapacitation effects on the market of high levels of enforcement, the economic fundamentals of the market also lead to the same conclusion. We show below in our model that, under reasonable assumptions, revenue must eventually begin to decrease with more enforcement. Whether this can be effected at realistic and tolerable enforcement levels is another question, and is likely to depend on the details of the illicit-substance market under study (Reuter & Kleiman, 1986); a possible ban on menthol cigarettes is illustrative.
Here we offer an economic analysis of cigarette taxes and regulations, black markets, and violence. We acknowledge that this form of analysis misses important sociological, legal, and other factors of the market. The research provided here is intended to complement the insights from other fields.

4.2. **Background on the Economics of Smoking**

Crucial to the modeling below will be some sense of how elastic is the demand for smoking cigarettes, and how that elasticity compares with the supply elasticity. The excellent review article by Chaloupka and Warner (2000) covers the twentieth-century literature on the econometric estimation of demand for smoking cigarettes in general. They summarize the literature as finding results in the broad range of $-0.14$ to $-1.23$ for demand elasticity. Importantly, however, once a few outlying studies are excluded, they conclude that most studies find results in the relatively narrow band of demand elasticities between $-0.3$ and $-0.5$. Chiou and Muehlegger (2008) report demand elasticities (including both the decision to smoke and the intensity of smoking among those who smoke) of $-0.29$ to $-0.56$. Thus, an important conclusion for the present work is that we can assume that cigarette smoking is highly inelastic.\(^4\)

What about the demand for menthol cigarettes in particular? We may suspect that the elasticity of demand is higher the more narrowly we define the product category, because, when the price of one type of cigarette rises, consumers can switch to other types. However, what little evidence there is does not appear to bear this out. One of the very few econometric studies estimating any part of the demand decision for menthol cigarettes, Tauras et al. (2010), concluded that “menthol and non-menthol cigarettes are not close substitutes” (p. 121). Their estimated switching elasticity of $-0.24$, which they calculate conditional on being a smoker, means that a 10 percent increase in the price of menthol cigarettes is associated with a 2.4 percent decrease in the probability of being a smoker of menthol cigarettes. These few individuals are the only ones who switch to smoking non-menthol cigarettes. While Tauras et al. (2010) do not estimate the other relevant parts of the inclusive demand elasticity—the extensive margin of the smoking-participation decision and the intensive margin of how many menthol cigarettes to smoke—there is no reason to believe that the inclusive elasticity is greater than one. Most of the studies cited by Chaloupka and Warner (2000) that find elasticity in the range $-0.3$ to $-0.5$ include both the extensive and intensive margins. Since the elasticities from

---

\(^4\) As there is already a substantial black market in cigarettes in the United States, estimates based solely on official counts of legal sales would tend to overstate the demand elasticity (i.e., with a price increase that drives some smokers to the black market the reduction in consumption would be overestimated.) However, most of these studies are from survey of individuals about all use, legal or otherwise. Furthermore, to the extent that people still lie and underreport their illicit consumption, the log-log specification in most studies will still get at the true elasticity, as all changes are in percentage terms and not levels.
each part of the demand process are additive, even adding the switching elasticity of \( -0.24 \) (Tauras et al., 2010) keeps the elasticity range well within the inelastic region.

Other studies looking specifically at demand for menthol cigarettes also come to the conclusion that demand is relatively insensitive to price. Compass Lexecon (2011) finds in an econometric study that a 10 percent increase in the effective price of illegal menthol cigarettes would be associated with a 1 percent decline in overall smoking, and that a 50 percent increase in the price would reduce smoking by only 3.5 percent.

Furthermore, the demand for cigarettes is likely to be much less elastic than the supply in the long run. Supply elasticities are properties of production technology and costs in a competitive market. The more slowly costs rise as quantity produced increases, the higher the supply elasticity. Given the economies of scale in the tobacco-products industry and the easy conversion of producing and packaging menthol cigarettes to doing the same for non-menthol cigarettes, it is quite likely that the supply function for menthol cigarettes is highly elastic. Lending credence to our supposition, Sumner and Wohlgenant (1985) found in a rigorous study of supply and demand conditions for cigarettes in the United States that “the derived supply curve for cigarettes is nearly horizontal” (p. 241) (even when the domestic supply of raw tobacco is inelastic due to agricultural quotas). The ratio of demand to supply elasticity will show up in the modeling work below, and we argue that it is likely to be small in magnitude, given the inelasticity of demand and the likely high elasticity of supply.

### 4.3. The Model

Consider a retail market for menthol cigarettes that would be competitive in the absence of taxes or a ban. While some studies examine noncompetitive illicit-drug markets (e.g., Caulkins, Reuter, & Taylor, 2006), assuming a competitive market allows intuitive analysis with the standard tools of supply and demand. The other assumptions and notation are these:

- The legal market (before a ban) is taxed with a revenue tax of rate \( t \). We will treat the tax as if it were levied on consumers. The tax is depicted graphically in

---

5 See Chiou and Muehlegger (2008) for a proof.
6 This study was funded by a tobacco company and was not peer reviewed.
7 The effective price includes not only the purchase price, but also the costs to the consumer of participating in the black market.
8 This is just for analytic convenience. Basic microeconomic theory shows that the legal incidence of the tax does not affect the actual economic incidence of the tax, and so our assumption is immaterial even if in actuality tobacco taxes are levied on the retailer.
• Figure 1 as shifting the demand curve (marked $D$) down, creating a wedge of size $tp$ between the actual demand curve and the demand curve modified to show how much of the total price per unit (including the tax) goes to the seller (marked $D'$). The figure is the standard textbook case of a revenue tax. In the figure, the price including the tax at $Q_1$ is $(1 + t)p_1$, of which amount $p_1$ per unit is received by the seller as revenue and $tp_1$ per unit is collected as tax revenue.\(^9\)

Figure 1. The impact of revenue taxation on demand.

![Diagram of demand curve with tax](image)

• The ban is denoted with a binary variable $b$. Without a ban, $b = 0$. With a ban, $b = 1$.

• Enforcement is treated as a continuous variable $e \geq 0$. For example, $e$ might be expenditure on enforcement (or, to set aside questions of inefficient spending, the minimum expenditure necessary to achieve a given real outcome related to enforcement, the number of police officers devoted to enforcing the ban, or the number of arrests). It is assumed that when $b = 0$, we also have $e = 0$.\(^{10}\)

• On the supply side, enforcement of the ban raises the effective marginal cost of doing business by a fraction $\varphi$. These additional costs include the perceived risk of arrest, sanction, fine, or incarceration, as well as any supply-disruptive activity following from enforcement, such as product seizure. We assume the cost rises with enforcement: $\varphi(e)$ is a smoothly increasing con-

---

\(^9\) For readers less familiar with the historical quirks of economic pedagogy, note that the demand function expresses the quantity demanded as a function of the good’s price, while it is traditional to graphically depict the relationship in inverse form with price on the vertical axis. Thus the demand curve is drawn as if price were a function of quantity. The same inverse relationship holds between the supply function and the supply curve.

\(^{10}\) As part of assuming that $e = 0$ before a ban, we set aside any considerations involving the possibility of tax evasion in the legal market.
cave function, with \( \varphi(0) = 0, d\varphi/de > 0, \) and \( d^2\varphi/de^2 < 0. \) The concavity of the additional cost reflects diminishing returns to enforcement effort (as would be the case, for example, if the most productive enforcement efforts directed toward suppliers are taken first). However, for mathematical convenience \( \varphi \) is unbounded. This reflects the assumption that the marginal product of enforcement never falls to zero. Figure 2 depicts the function \( \varphi, \) and Figure 3 depicts its effect (for a given level of enforcement) on the supply curve.

**Figure 2. The level and shape of function \( \varphi \) (or \( \rho \)).**

![Figure 2](image)

**Figure 3. The impact of enforcement on the supply curve.**

![Figure 3](image)

- A fraction \( \ell \in [0,1] \) of consumers are law-abiding, in the sense that, regardless of the street price of banned cigarettes, they would never purchase any after a ban. These consumers either switch to legal cigarettes or quit smok-

---

11 Just because costs rise with enforcement does not mean that enforcement need be very effective at raising costs for suppliers, just that it does at least a little. Reuter and Kleiman (1986) argue that enforcement in marijuana and cocaine markets had little effect on dealer and distributor costs.

12 For an explanation of the quantity \( S^{-1}(\varphi p_0) \) in Figure 3, see Equation 2’ and fn 14 below.
ing. We assume that the smoking intensity (i.e., the quantity of cigarettes consumed) of these consumers is completely typical, so that after a ban the demand curve for menthol cigarettes is only \((1 - \ell)\) what it was before the ban, all else equal. Figure 4 depicts the impact of \(\ell\) on demand (drawn for \(\ell = 0.5\)).

**Figure 4. Shift in the demand curve due to \(\ell\).**

- Enforcement may also impact demand directly, apart from its indirect impacts through the price mechanism. If enforcement targets users, then the perceived inclusive price of the product rises by fraction \(\rho(e)\). (User-targeted enforcement is included here as a general feature of drug-enforcement models. In the case of illicit cigarettes, enforcement is unlikely to target smokers, except for minors.) The consumers treat this like a “risk tax.” We assume the risk tax rises with enforcement: \(\rho(e)\) is a smoothly increasing, concave function, with \(\rho(0) = 0, d\rho/de > 0\), and \(d^2\rho/de^2 < 0\). The concavity of the risk tax reflects diminishing returns to enforcement effort (as would be the case, for example, if the most productive enforcement efforts directed toward consumers are taken first). The shape of function \(\rho\) is similar to that of \(\varphi\) in Figure 2, although we do not assume they are the same function. Note that we have simplified the model by not defining separate variables for enforcement levels on the supply and demand sides. Instead, we assume any differential impact of enforcement on the two sides of the market caused by different allocations of enforcement effort split between the two is captured by the shapes of the functions \(\varphi\) and \(\rho\) themselves. The impact of demand-side enforcement on the demand curve is as for the revenue tax in Figure 1, with \(\rho\) replacing \(t\).

- The quantity demanded in the absence of a ban is a downward-sloping function \(q_D(p)\), where \(p\) is the price consumers must pay. We assume \(q_D\) is a differentiable function, and denote the demand elasticity with \(\varepsilon = -q_D'(p) p/Q\) (so that \(\varepsilon\) is a positive number). Given the empirical evidence reviewed above, we assume that demand is inelastic in the relevant range of prices: \(\varepsilon < 1\) (although we do not assume that elasticity is constant).
• The quantity supplied in the absence of a ban is an upward sloping function \( q_S(p) \), where \( p \) is the price sellers receive. We assume \( q_S \) is a differentiable function, and denote the supply elasticity with \( \eta = -q_S'(p) p/Q \).

• We assume that violence rises with illicit revenue earned in the market. Violence is a negative externality in the market for the illicit substance. At least some of the factors associated with drug violence identified by Goldstein (2003) are closely related to revenue instead of profit. For example, robberies of drug dealers (and the violent responses in retribution) are triggered by expectations that dealers carry large amounts of cash on their person. Similarly, areas known for traffic in illicit substances are attractive targets for robberies in general (not just of known dealers), since more people on the street will be buyers carrying substantial amounts of money. Other than the monotone increasing relationship between illicit revenue and violence (and the level assumption in the next bullet) we need make no other assumptions about the function relating revenue to violence.

• We normalize violence to zero at the revenue level associated with a ban with no enforcement. Given the abstract nature of the model, this just means that we set aside any violence associated with the legal tobacco market to focus on marginal changes in violence from changing the market to an illicit one. This assumption also implies that there are no risks due to violence when there are no risks associated with running afoul of the law, either on the demand or the supply side.

• We set aside any direct beneficial effect of enforcement effort on reduction in violence. In part this is to focus on the economic effects on violence through the price mechanism. However, our treatment here is also in line with the large literature noting that enforcement activity often is associated with more, rather than less, drug violence.\(^{13}\)

With a slight redefinition of notation, \( p \) from now on will refer to the monetary price exchanging hands in the market, not including taxes or any risk adjustments. Given the assumptions above, the quantity demanded accounting for a possibility of taxes, a ban, and enforcement is given by the expression

\[
Q_D(p, b, e) = (1 - \ell)^b q_D([1 + (1 - b)t + \rho(e)]p). \tag{1}
\]

In the expression, the inclusive price the demander must pay is marked up by the tax rate \( t \) (if there is no ban) and the risk tax. The first term on the right side of (1) accounts for the proportional reduction in demand when the law-abiding citizens exit the market.

The quantity supplied, likewise, is given by the expression

\(^{13}\) See Zahn (1975) for an early study reviewing historical evidence from the Prohibition era onward pointing out that prohibition and enforcement often lead to higher homicide rates.
The expression discounts the monetary price to adjust for the risk incurred or supply disruption when there is enforcement.\textsuperscript{14} To simplify the expression, we define a function $\psi = [1 + \varphi(e)]^{-1}$, so supply can be written as

$$Q_s(p, e) = q_s \left( \frac{p}{1 + \varphi(e)} \right). \quad (2)$$

Scaling by function $\psi$ in the supply function thus converts market price $p$ to the actual average revenue per unit suppliers receive after netting out enforcement costs (see Figure 3). It will be useful for what follows to note that $\psi(e) > 0$, $\psi(0) = 1$, $d\psi/de < 0$, and $d^2\psi/de^2 > 0$, and $\lim_{e \rightarrow \infty} \psi(e) = 0$. The function $\psi$ is depicted in Figure 5. In equilibrium, supply matches demand and we have $Q_S = Q_D$.

**Figure 5. The level and shape of function $\psi$.**

\[\varphi(e)\]

1

\[Q\]

4.4. **The Impact of a Ban\textsuperscript{15} and Enforcement on Price and Revenue**

Even though enforcement in actuality will likely begin immediately upon imposition of a ban, it is useful conceptually to think of a ban and enforcement as occurring in two steps. We emphasize that this exercise is only for analytic purposes and does not reflect any assumptions or requirements in our model. When the ban is imposed in the model, but with no enforcement in place, the

\textsuperscript{14} The derivation of (2) follows from the fact that, in a competitive market, the supply curve is the (inverse) marginal cost curve of the industry. If production costs as a function of $Q$ are $MC(Q)$ in the absence of a ban, then marginal costs are $MC(Q_0)[1 + \varphi(e)]$ with a ban. Thus the inverse supply curve is given by $p = MC(Q_0)[1 + \varphi(e)]$, and inversion yields $Q_s = MC^{-1}(p[1 + \varphi(e)]^{-1})$. After identifying the function $q_s$ with inverse marginal cost, equation (2) follows. This also explains the notation for the leftmost point on the horizontal axis in Figure 3.

\textsuperscript{15} Although, for all cigarettes, tax increases are more likely than bans, here we consider a ban, which is the more likely case for menthol cigarettes. The analytical approaches to the two are similar—a ban is the limiting case of a tax increase, or, conversely, a tax increase is a lesser included case of a ban—but a ban is simpler to present.
outcomes of the model—price, quantity, and revenue—change discontinuously as \( b \) switches from zero to one. This is illustrated in the following figures.

Without the ban, the transacted price in the market, \( p_1 \), is the price implicitly defined by \( Q_D(p,0,0) = Q_S(p,0) \), or

\[
q_D((1 + t)p_1) = q_S(p_1). \tag{3}
\]

See Figure 6. In the figure, the tax has the effect of lowering the price that sellers receive from the no-tax price \( p_0 \) to \( p_1 \). The price paid by the consumer is \( p_1 \) plus the revenue tax, which adds another \( tp_1 \) per unit purchased.

**Figure 6. Equilibrium price in a taxed, legal market.**

![Graph showing equilibrium price in a taxed, legal market.](image)

With the ban but without enforcement, the following defines the new price \( p_2 \):

\[
(1 - \ell)q_D(p_2) = q_S(p_2). \tag{3'}
\]

Note that no taxes are collected in the illegal market. Figure 7 shows the impact of law-abiding consumers leaving the market: demand shifts from \( D \) to \( D' \) as in Figure 4. The shift has the effect of lowering the price that buyers pay and sellers receive from the no-tax price \( p_0 \) to new price \( p_2 \). Unless \( \ell = 0 \), so that the original and post-ban demand curves coincide (implying that \( D \) would be the same as \( D' \) in Figure 7), it cannot be determined in general which of the two prices is higher, \( p_1 \) or \( p_2 \) (nor for \( (1 + t)p_1 \) and \( p_2 \)).\(^{16}\) If \( \ell \) is zero, then we would have \( p_2 = p_0 \), and clearly then it would follow that \( p_2 > p_1 \). However, when

\(^{16}\) In the licit cigarette market the sales price reflects production costs, taxes, and brand rent (the difference in production costs between premium brands and discount brands is minimal compared with the difference in their retail prices). Eliminating brand rents would drive smokers to “cheap whites,” already prevalent in Europe (Denner, 2013).
\( \ell > 0 \), the reduction in demand from the exit of law-abiding consumers attenuates and may even reverse the price increase from the avoided taxation.\(^{17} \)

**Figure 7. Equilibrium price in an illicit market with no enforcement.**

Regardless of which case applies, it is important to note that it does not matter for our analysis of the link between enforcement and violence. Our assumption above was that increases in *illicit* revenue are associated with more violence, and that there is no violence under conditions \( b = 1, e = 0 \). Thus, even though revenue changes as we move from equation (3) to equation (3'), there is no change in violence yet.

Now that we have our conceptual baseline of a ban but no enforcement yet, consider how prices, revenue, and violence change as enforcement ramps up. Define excess demand as

\[
G(p, e) = Q_D(p, b, e) - Q_S(p, e).
\]

Then applying the implicit function theorem to the equilibrium condition \( G(p, e) = 0 \) when there is a ban \((b = 1 \text{ and } e \geq 0)\) gives an expression for how the equilibrium (monetary) price changes with enforcement on the margin:

\[
\frac{dp}{de} = -\frac{\partial G/\partial e}{\partial G/\partial p} = -\frac{[(1-\ell)(\varepsilon/\eta)\rho'(e)+\psi'(e)]p}{(1-\ell)(\varepsilon/\eta)(1+\rho(e))+\psi(e)}.
\]  

\(^{17}\) For simplicity we take \( \ell \) to be an exogenous parameter. A more realistic (albeit more complicated) approach may be to model \( \ell \) as a decreasing function of the quantity consumed in an illegal market. This would introduce the notion that there may be bandwagon effects among potential scofflaws: If there is an adequate supply of illicit goods *and little or no enforcement*, some fraction of initially law-abiding consumers will see that the illegal behavior is normalized and decide to engage in it. Witness the growth of piracy in recorded music and the near absence of use-tax payments on out-of-state purchases.
In the expression, \( b \) has been set to 1 since the ban is in place.\(^{18}\) The denominator of (4) is unambiguously positive, and so the sign of the expression depends only on the numerator. Thus market price will increase on the margin, with increases in enforcement, if and only if

\[
\varepsilon \rho'(\varepsilon) < -\eta \psi'(\varepsilon).
\]

Both sides of inequality (5) are positive, and thus whether the condition is satisfied depends on the relative magnitudes involved. We analyze various cases below to draw insight from the model.

To illustrate why the price effect of increased enforcement is ambiguous in general, and to help interpret condition (5), consider Figure 8. In the left panel, the shifts in the supply and demand curves lead to a higher price (compared to the starting point at \((p', Q')\) with the ban in place but no enforcement). In the right panel, the shifts in the curves lead to a lower price. The left side of inequality (5) arises from the impact on the equilibrium price from the demand curve shifting in, which causes excess demand to fall and the price to fall. The right side of inequality (5) arises from the impact on the equilibrium price from the supply curve shifting in, which causes excess demand to rise and the price to rise. Only when the latter effect outweighs the former effect does price rise on net.

**Figure 8. Changes in equilibrium price in an illicit market due to enforcement.**

If the demand curve did not shift, and demand is inelastic as argued above, then knowing the direction of the price change also lets us know how revenue changes. When the demand curve itself shifts downward in response to enforcement, however, there is a countervailing effect on revenue, as we now

\(^{18}\) The simplified expression in (4) uses the fact that, by definition of the elasticities, \( q'_D/q'_S = -\varepsilon/[\eta(1 - \ell)] \) when both derivatives are evaluated at the market-clearing price and quantity \( Q = Q_D = Q_S \).
demonstrate. Equilibrium revenue with a ban can be written as $R = p^*Q_D(p^*, 1, e)$, where $p^*$ is the equilibrium price given enforcement level $e$. Thus the total impact on revenue of a marginal change in enforcement level is

$$\frac{dR}{de} = \frac{dp^*}{de} Q + p^* \left( \frac{\partial Q_D}{\partial p} \frac{dp^*}{de} + \frac{\partial Q_D}{\partial e} \right) = \left( Q + p^* \frac{\partial Q_D}{\partial p} \right) \frac{dp^*}{de} + p^* \frac{\partial Q_D}{\partial e}$$

(6)

= price effect + demand-shifting effect.

The first term on the far right side is the price effect. When enforcement rises, the price changes in accord with equation (4), which has a marginal impact on revenue. The term in parentheses on the far right side of equation (6) is the usual marginal-revenue (MR) term for the impact on revenue from a marginal increase in price. For inelastic demand, for which $\varepsilon < 1$, this term is always positive:

$$Q + p^* \frac{\partial Q_D}{\partial p} = Q(1 - \varepsilon) > 0.$$ 

Thus, when the demand curve sits still, revenue change has the same sign as $dp^*/de$. However, the second term on the far right side of equation (6), the direct impact of enforcement on the demand curve (the “demand-shifting” effect), is negative, since demand slopes down. From equation (1), we have:

$$p^* \frac{\partial Q_D}{\partial e} = -R\varepsilon \rho'(e) < 0.$$ 

This term is the offsetting effect on revenue of the demand curve shifting down in response to enforcement risk. Revenue falls through this channel because, with lower demand, for any given price the quantity is lower.

Thus, even when the sign on the price change is clear from inequality (5), the total effect on revenue and therefore violence is still unclear. If the demand-shifting effect is large enough, there might be less revenue in response to a price increase. However, given the relatively small elasticity values found in the literature for smoking in general and for menthol cigarettes in particular, we expect that the demand-shifting effect will attenuate but not reverse the primary impact on revenue of the price change. We therefore maintain this assumption in what follows. In the limiting case with completely inelastic demand, $\varepsilon = 0$ and the demand-shifting effect disappears.

4.4.1. Case 1: No direct enforcement on consumers

Consider what happens if enforcement is entirely focused on suppliers and consumers are left alone. If enforcement does not affect the demand side at all, then $\rho'(e) = 0$. Then condition (5) is satisfied and the price rises with enforcement effort. If demand is inelastic, then revenue rises with the price unambiguously (since there is no countervailing demand-shifting effect in
Recall that we have assumed that violence rises with revenue. Thus, in the case where enforcement targets the production, importation, and distribution side of the illicit market, there will be greater violence with stronger enforcement (conditional on the ban already being in place).

This is illustrated in Figure 9. The demand curve, adjusted for the exit of law-abiding consumers, is as in the previous figure. The supply curves with and without the impact of enforcement are as in Figure 3. Compared to the outcome under the ban with no enforcement at all (price $p_2$ and quantity $Q_2$), the price unambiguously rises to $p_3$ with enforcement only on the supply side.

**Figure 9. Case 1—Equilibrium price in an illicit market with only supply-side enforcement.**

In the long run in a competitive market, each producer is forced to produce at its minimum efficient scale (the cost-minimizing production quantity). In a constant-cost industry, minimum efficient scale leads to similar costs for all producers, and so the long-run supply curve is horizontal (Pindyck & Rubinfeld, 2009). In other words, the supply function is perfectly elastic. In terms of our model, this is the limiting case in which $\eta$ is infinite and therefore inequality (5) is satisfied. Thus, in this important “textbook” case, greater enforcement of a ban will unambiguously lead to higher prices. If, as argued above, the demand-shifting effect in equation (6) does not outweigh the price effect, there will be more revenue and therefore violence.

This case is illustrated in Figure 10, which clearly shows why there is no ambiguity in the direction of the price change. After the ban but before enforcement, demand curve $D$ and supply curve $S$ lead to price $p_4$ and quantity $Q_4$. The upward shift in the supply curve (from $S$ to $S'$) due to the enforcement risk raises the price (from $p_4$ to $p_5$) as suppliers require more compensation for any given level of production. The demand curve shifting down (from $D$ to $D'$)
does not further change the price, which is still $p_5$ due to the elastic supply curve, but decreases revenue a bit by reducing quantity. While quantity had been at the intersection of $S'$ and $D'$, it now moves left to $Q_5$. This offsetting decrease in revenue is the demand-shifting effect.

**Figure 10. Case 2—Equilibrium price with enforcement and perfectly elastic supply.**

4.4.3. Case 3: No direct enforcement on suppliers

If all enforcement effort were directed at penalizing consumption of the illicit good, so that $\psi'(e) = 0$, then condition (5) would never hold for any demand and supply functions and $dp/de < 0$. In this case, both the price effect and the demand-shifting effect work in the same direction to reduce revenue. Therefore a greater crackdown on consumers alone will unambiguously lead to less violence, by reducing demand. This is illustrated in Figure 11, with price and quantity falling from $(p_6, Q_6)$ to $(p_7, Q_7)$. In practice, however, enforcement efforts are concentrated on suppliers rather than consumers, as their smaller numbers make them easier to identify, their greater potential penalties more worth the fixed costs of targeting a suspect, and their public disfavor politically more feasible. The federal government, in particular, rarely concerns itself with simple possession cases.

---

19 So long as the supply curve is not perfectly elastic, that is.
Figure 11. Case 3—Equilibrium price in an illicit market with only demand-side enforcement.

4.4.4. Case 4: Completely inelastic demand

If there is no responsiveness to price at all on the demand side, perhaps because all but the hardest-core addicts have left the market due to high prices, then we have $\varepsilon = 0$. In this case, as in Case 2, inequality (5) is satisfied. Price rises unambiguously. Furthermore, there is no demand-shifting effect, so the impact on revenue is clear. More enforcement of a ban leads to more violence. This is illustrated in Figure 12, with the price rising from $p_8$ to $p_9$.

Figure 12. Case 4—Equilibrium price in an illicit market with perfectly inelastic demand.

4.4.5. Case 5: Higher levels of enforcement effort

Rearranging terms, condition (5) can be expressed as:

$$\frac{\varepsilon}{\eta} < -\frac{\psi'(e)}{\rho'(e)}.$$  \hfill (7)
Since $\psi$ is positive and bounded below by zero, and $\rho$ is unbounded, the right side of condition (7) approaches zero as $e$ becomes large. Thus, so long as nothing extreme happens to the elasticity ratio on the left side of inequality (7) (and we are not in any of the limiting cases above), there will be a threshold level of enforcement $\hat{\epsilon}$ beyond which additional enforcement effort will surely reduce revenue and hence violence. Eventually, at high enough enforcement levels, violence will begin to fall as the quantity sold gets choked off. Nothing in this analysis, however, suggests what that threshold would be for any particular illicit market, or whether the resultant enforcement regime (perhaps a police state) would be tolerable.

4.4.6. Case 6: Lower levels of enforcement effort

Continuing with the analysis in the previous case, what about when $e < \hat{\epsilon}$? Then the right side of (7) is large enough to satisfy the inequality for the price to rise. The only question, therefore, is whether $\hat{\epsilon} > 0$; that is, for some functions $\psi$ and $\rho$ it may be that inequality (7) is never satisfied for any positive level of enforcement. However, we argue that it is likely that $\hat{\epsilon} > 0$ for the following reason. As Reuter and Kleiman (1986)\textsuperscript{20} point out, street prices are very high for illicit drugs like marijuana, cocaine, and heroin compared to what cost-based prices in a legal market would be. Caulkins and Reuter (1998) estimate that, for the cocaine market, compensation required for the risks involved with producing and distributing in the illicit market accounts for one-third of industry cost. Thus the mere fact that there is some enforcement in the arena of currently illegal drugs has had huge impacts on prices. Given that much of the enforcement effort is targeted toward the supply side, this argues that $\psi'(e)$ (or $\varphi'(e)$) is very large in magnitude at least for the first efforts at enforcement. Thus, there is some range of enforcement levels in general (apart from some of the special cases above) where more enforcement leads to more violence. In particular, when enforcement is ramping up and still at a low level (depending on the policy decision, in anticipation of, or immediately after, an increase in regulation), the system is most vulnerable to spikes in illicit revenue and violence.

4.4.7. Summary of analysis

In the event of a ban, the level of violence occasioned by a black market, relative to the revenue of that market, will tend to increase with enforcement intensity, because the incentives for violence change as enforcement risks rise. That revenue depends on both the price and the level of consumption of the black-market good. Consumption will decrease as prices increase, all else equal, and prices and consumption will decrease the more law-abiding (i.e., not

---

\textsuperscript{20} Reuter and Kleiman (1986) discuss this point in their conclusions, and are careful to point out that this does not contradict their main argument, which is that the marginal impact of additional enforcement on prices is likely to be low. That main part of their argument bolsters the analysis above for Case 5.
entering the black market) current smokers are. The price depends, as well, on the intensity and kind (focused on the supply side or demand side) of enforcement of the ban; all else equal, stricter enforcement raises the price—unless that enforcement is focused entirely on consumers (which is unrealistic in practice).

The current illicit tobacco market is high-dollar, low-violence; increasing enforcement might tend to increase violence over a fairly wide range. The particular circumstances under which increasing enforcement tends to increase revenue (i.e., when a small increase in enforcement raises the price by proportionally more than it reduces consumption), and thereby to increase violence, depend on the details of consumers’ sensitivity to prices and on their law abidingness. Without estimates of those consumer propensities, it is impossible to determine the optimal degree of enforcement to minimize violence—and this is without taking the monetary and other costs of enforcement into account. (The research agenda below describes an approach to estimating the required consumer propensities, for the case of a ban on menthol cigarettes.) Evidence from currently illicit drugs suggests that, at still-low-but-increasing levels of enforcement, violence increases as enforcement does.

### 4.5. Menthol

Bans on particular types of cigarettes create special complexities. Menthol cigarettes are facing a possible ban in the United States (Tavernise, 2013) and an imminent one in the European Union (Dalton & Esterl, 2013). Menthol cigarettes constitute approximately one-third of the overall U.S. cigarette market (Canterbury, 2011; O’Connor, Bansal-Travers, Carter, & Cummings, 2012) and account for approximately $25 billion in annual retail sales (Esterl, 2011). The majority of menthol smokers are African American or Latino, and Winickoff et al. (2011) and Pearson, Abrams, Niaura, Richardson, and Vallone (2012) both report strong support for a potential menthol ban in these communities. Menthol smokers exhibit particularly strong brand loyalties (Anderson, 2011).

Congress exempted menthol when it banned non-tobacco flavorings in cigarettes in 2009 (Family Smoking Prevention And Tobacco Control Act, 2009), but mandated a study of the matter by the U.S. Food and Drug Administration (FDA). FDA’s Tobacco Products Scientific Advisory Committee (2011) report concluded, “Removal of menthol cigarettes from the marketplace would benefit public health in the United States.” An industry-perspective report submitted to the FDA at the same time (Heck, Hamm, & Lauterbach, 2011) acknowledged that all cigarettes are hazardous but argued that there is no scientific basis for regulating menthol cigarettes differently. On July 23, 2013, the FDA released a report (U.S. Food and Drug Administration, 2013a) that found that menthol cigarettes are not inherently more harmful than non-menthol cigarettes on a unit basis, but that menthol makes it easier to start smoking and more difficult to cut back or quit. The report was posted to begin a 60-day public-comment
period to allow the FDA to “determine what, if any, regulatory action with respect to menthol in cigarettes is appropriate” (U.S. Food and Drug Administration, 2013b)\(^{21,22}\)

If menthol is banned, we might expect to see some level of “channeling” to other tobacco products, switching to non-menthol brands or varieties, quitting tobacco altogether, tax evasion, and smuggling. Hartman (2011), O’Connor et al. (2012), and Pearson et al. (2012) all report that 35–40 percent of menthol smokers say they would quit smoking altogether if menthol cigarettes were no longer available. O’Connor et al. (2012), however, report that 25 percent said they would “find a way to buy a menthol brand,” indicating their willingness to purchase menthol cigarettes on the black market. Tobacco-company representatives contend (Heck, Hamm, & Lauterbach, 2011; Murillo, 2011), and industry-supported studies argue (Goozner, 2011; Roland Berger Strategy Consultants, 2013) that severe negative impacts on public health, criminal activity, and tax revenues will ensue should a menthol ban pass.

An (industry-supported) analysis of proprietary data on menthol- and non-menthol- cigarette prices and sales (Compass Lexecon, 2011) concludes that current menthol smokers would largely turn to the black market for menthol cigarettes: A 10 percent increase in the effective price of black-market menthol cigarettes would, by this estimate, lead to an initial decline in overall smoking of about one percent, and black-market sales would be about 87 percent of current menthol sales; a 25 percent increase would lead to an initial decline in overall smoking of about two percent, and black-market sales would be about 72 percent of current menthol sales.

### 4.6. Electronic Cigarettes

Electronic cigarettes (“e-cigarettes”) are a rapidly growing piece of the U.S. tobacco-products market. They produce a nicotine-laced vapor, rather than smoke, and currently are allowed to be used in many circumstances where smoking is prohibited. Analysts estimate that retail and Internet sales of e-cigarettes could reach $1–2 billion in 2013 (Associated Press, 2013b; Drill, 2013) and $10 billion (approximately ten percent of the combined conventional and e-cigarette market)\(^{23}\) by 2018 (Drill, 2013). The FDA attempted to ban e-cigarettes as unapproved drug/device combination products but was overruled by the U.S. Court of Appeals in 2010 (Deyton, 2011). The court ruled that e-

---

\(^{21}\) The release came a day ahead of a deadline for the United States to respond to the finding last year by a World Trade Organization (2013) appeal board that upheld Indonesia’s claim that the flavoring ban—which prohibits Indonesian-made clove cigarettes but exempts domestically produced menthol cigarettes—constitutes an unfair trade practice. According to WTO rules, the United States can either allow clove cigarettes, ban menthols, or be subject to retaliatory action by Indonesia (Action on Smoking & Health, 2013).

\(^{22}\) The public-comment period was later extended by another 60 days, until November 23, 2013 (Menthol in Cigarettes, 2013).

\(^{23}\) Based on the U.S. tobacco-market projection of MarketLine (2013).
cigarettes and other products made or derived from tobacco can be regulated by the FDA as “tobacco products” under the Tobacco Control Act, and are not drugs/devices unless they are marketed for therapeutic purposes (Deyton, 2011). The FDA is currently scheduled to release a set of proposed rules by October 2013\textsuperscript{24} (Drill, 2013), in addition to a list of dangerous chemicals found in e-cigarettes and testing/reporting requirements for tobacco additives (Klein Moynihan Turco LLP, 2013a, 2013b, 2013c, 2013d; Mangan, 2013). These rules may entail a ban on flavors other than tobacco, with menthol flavoring included. New York City (Arroyo, 2013) and the State of California (Corbett, 2013) are both pursuing expanded regulations on e-cigarettes.

The health effects of e-cigarettes are a matter of debate due to a paucity of peer-reviewed research on the subject (Head, 2011; Henningfield & Zaatari, 2010; McQueen, Tower, & Sumner, 2011). The claim that e-cigarette smoking has only a fraction of the risks of conventional smoking is given surface plausibility by the absence of particulates, carcinogenic “tars,” and hot gases, but nothing resembling a controlled trial has been attempted, and human-subjects concerns might render any such study virtually impossible to conduct. Kozlowski (2007, 2013a, 2013b) claims that e-cigarettes are proven smoking-cessation aids, serve harm reduction in existing smoking populations due to e-cigarettes’ lack of contaminants and adulterants, and promote public health due to the absence of smoke or secondhand smoke. Warner (2013) reports that nicotine-only e-cigarettes are “dramatically” less dangerous than combusted tobacco. Burstyn (2013) finds that e-cigarette users’ exposure to contaminants in the inhaled vapor is well below workplace-safety standards for involuntary exposure. The UK Medicines and Healthcare Products Regulatory Agency (2013) currently treats e-cigarettes as medicines for treating nicotine addiction. The American Medical Association finds that, from a public-health perspective, e-cigarettes may present an effective alternative to traditional tobacco use (Head, 2011).

If e-cigarettes are much safer than traditional cigarettes and reasonably attractive to a large share of smokers, then promoting or enabling their use would allow for increased regulations on tobacco now by allowing non-quitting smokers to become “vapers” instead (and thereby diverting them from black markets). E-cigarettes, however, may present an alternative not only to smoking, but also to quitting; if appealing, inexpensive, and widely tolerated e-cigarettes made it easier for smokers not to quit (by serving as a “bridge” between smoking sessions, in circumstances where smoking is prohibited or frowned upon), then the public-health benefits of increased tobacco regulation would not be as compelling (Benowitz & Goniewicz, 2013). There is no evidence to date that indicates whether e-cigarettes act as a net substitute or complement to combustible cigarettes. (Another objection to e-cigarettes, that they induce youth to use nicotine and thereby act as a gateway to smoking (Grana,

\textsuperscript{24} This action may have been delayed by the Federal Government shutdown (Sebastian, 2013).
2013; Kelland & Hirschler, 2013) does not amount to a finding that their availability leads to greater net smoking initiation than would their prohibition, but such a possibility is not ruled out by available data; fashion effects, in particular, are always difficult to predict.)

If menthol combustible cigarettes were to be banned, switching to menthol-flavored e-cigarettes would represent an additional choice for menthol smokers, beyond switching to traditional cigarettes or quitting. If that choice were popular, and if the health risks of using e-cigarettes are as much smaller than the risks of using conventional cigarettes as has been claimed, the public-health benefits of a menthol ban might increase while the illicit-market costs would shrink. If a menthol ban led a sizable percentage of current menthol smokers to switch to e-cigarettes, and to do so rather quickly, that would constitute a substantial increase in e-cigarette market penetration. This, in turn, could yield an additional public-health benefit by making it easier (less novel) for non-menthol smokers to switch to e-cigarettes. The increase in menthol e-cigarette market penetration depends on whether the FDA moves to ban all flavorings in e-cigarettes.

4.7. Policy Implications

To sum up, consider which of the various cases considered above are most likely to obtain if menthol cigarettes were banned in the United States. Federal drug-enforcement policy has always been more oriented towards pursuing high-visibility or large players on the supply side than the demand side (Reuter & Caulkins, 1986) (which has Case 1 in the limit). Standard industry considerations from microeconomics and industrial organization suggest that there is no reason the long-run supply curve of the illicit good would not be highly elastic (which has Case 2 in the limit). We have argued above that demand is highly inelastic (which has Case 4 as a limiting case). In each of these cases, at least in the limit the unambiguous conclusion is that more enforcement will lead to more violence, if there is a supply of acceptable-quality illicit menthol cigarettes.

Is policy captive to the inexorable economics of the model? No. Several parameters of the model are susceptible to manipulation by public policy, apart from the choice of enforcement level. One of the most obvious is the possibility of increasing the elasticity of demand for menthol cigarettes by public-health information campaigns, perhaps targeted at communities with above-average usage rates (e.g., African-Americans, Asian-Americans, and youth). Some previous mass-media campaigns directed at public health have been shown to have at least “small to moderate effects” on behavior (Noar, 2006), although as the pool of remaining smokers dwindles their resistance to persuasion may

---

25 Although states and localities have varied tremendously in their own targeting and punishment of users.
increase. Increasing the elasticity of demand will both shrink the market more rapidly and—if elasticity tips into the elastic region of demand—reverse the relationship between price and revenue that drives the pessimistic results above. Even when failing to raise demand elasticity above 1, any increases in magnitude will have the salutary effect of making condition (5) harder to satisfy.

Shifting more enforcement to the demand side from the supply side would move us closer to Case 3, in which enforcement does reduce violence. Such activity may be politically unpopular, however, and hugely undesirable on other grounds, since it would mean more arrests of minority-group members for trivial offenses.

### 4.8. Optimal Enforcement

An optimal regulation/enforcement policy reflects a balance of the costs and benefits of behavior under the resultant regime; the levels of both regulation and enforcement need to be determined jointly rather than taking one as given and then optimally determining the other. The optimization calculus, in turn, depends on estimates and valuations of the intended benefits (principally improved health and reduced mortality), unintended consequences (many, as discussed), and the costs of enforcing the regime. These valuations are either disputed, even if widely adhered-to conventions exist (e.g., quality-of-life-adjusted years for reduced mortality), or unknown (e.g., the retail price of menthol cigarettes in the event of a ban). A complete analysis is beyond the scope of this paper (and is, in all likelihood, untenable at this time), but an approach that allows for better decisionmaking can be sketched out.

The analytical challenge is even greater than suggested above. There is good reason to expect that—even with great confidence in the model structure and parameter values—there is not a single optimal regulation/enforcement policy that can readily be achieved from a given starting point. That is, there are multiple equilibria to which the regulation/enforcement system may settle (Caulkins & Reuter 2010; Feichtinger & Tragler, 2001). Which equilibrium is reached depends on the system’s history (i.e., it is “path dependent”), and moving from one equilibrium to another may entail substantial costs not reflected in a static analysis.

In the present context, the simplest formulation of the challenge is, “How much of a black market (and its attendant harms) are we willing to tolerate in exchange for the health benefits of higher taxes on, or stricter regulation of, cigarettes?” There are three notional equilibria:26

<table>
<thead>
<tr>
<th>Regime</th>
<th>Costs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Weak-regulation, strict-enforcement</td>
<td></td>
</tr>
</tbody>
</table>

26 The weak-regulation, strict-enforcement regime does not apply, as strict enforcement requires something substantial to enforce.
Note that enforcement is “strict” or “weak” compared to the underlying activity; a little bit of enforcement is sufficient to make moonshining unattractive. In any market, if there is a low violation rate initially, it is easy to detect and punish violators, so that a low level of enforcement is sufficient to keep crime under control, regardless of the underlying demand function.

A promising approach to determining the optimal level of cigarette regulation follows that of Poret (2009) for drug-law enforcement generally, which seeks to minimize the total social cost (social harm from distribution and use, cost of detection minus fines collected, minus surplus and profits). The actors are a supply chain, represented as wholesalers and retailers; consumers, who cannot be punished for using; and an enforcement authority, with resources to arrest and punish sellers. The enforcement decisions that bear on outcomes and costs are (a) allocation of monitoring efforts to wholesalers or retailers, (b) efforts to increase probability of detecting illegal transactions, and (c) sanctions to impose on convicted sellers.

Poret shows that three types of policies are optimal solutions under different conditions: strict sanctions and high probability of detection, no detection, or no sanctions. But remember that this analysis allows for independently varying the regulation and enforcement regimes; in the case of a specified regulatory regime an optimal solution may not be available, and a second-best solution must be sought.

Poret also shows that, contrary to the usual assumption in the literature on the economics of enforcement, the two instruments of law enforcement—detecting transactions and sanctioning—are not necessarily substitutes:

<table>
<thead>
<tr>
<th>regulation</th>
<th>enforcement</th>
<th>use</th>
<th>crime</th>
<th>enforcement</th>
</tr>
</thead>
<tbody>
<tr>
<td>strict</td>
<td>strict</td>
<td>low</td>
<td>low</td>
<td>low/high</td>
</tr>
<tr>
<td>strict</td>
<td>weak</td>
<td>high</td>
<td>high</td>
<td>low</td>
</tr>
<tr>
<td>weak</td>
<td>weak</td>
<td>high</td>
<td>low</td>
<td>low</td>
</tr>
</tbody>
</table>

27 A variety of theoretical approaches to optimal enforcement of drug laws is presented in Baveja et al. (2000); Becker, Murphy, and Grossman (2004); Garoupa (1997); and Tragler, Caulkins, and Feichtinger (2001). Efforts at empirical implementation include Chang, Coulson, and Wang (2002) and Yunker (2012).
28 In this model, consumers are assumed not to be addicts—each consumer buys one unit of drugs. Poret acknowledges this limitation, as harms incurred vary with the quantity and frequency of consumption. Application of the model to the partial prohibition of cigarettes, the vast majority of which are consumed by nicotine addicts, would require modifying this condition.
29 Poret (2009) considers only the two extreme cases of targeting only wholesalers or only retailers; a mixed, intermediate approach is developed in Poret (2002). The essential findings of the analysis are captured in the simplified approach.
30 See, e.g., Becker (1968).
When the main cause of harm is crime, the probability of detection and the sanction are complements. This apparently paradoxical result comes from the objective of the law enforcement policy. It is no longer a question of deterring the drug trade, that is, by reducing market quantities, but of minimizing the net social cost related to illicit drugs. When the harm is such that the reduction of social cost related to drugs is equivalent to the maximization of the drug trade, the instruments are complementary (p. 226).

5. Policy Options

5.1. Enforcement First, Then Regulation

Illicit markets generate positive feedback through “enforcement swamping” (Kleiman, 1993, 2009; Kleiman & Kilmer, 2009). Increased enforcement will be more efficacious and less costly, and will yield fewer pernicious side effects if applied before market growth than after a market is well established: Potential violators respond to the risk of enforcement (i.e., probability of detection and arrest), which is a function of (a) enforcement resources expended and (b) the size of the illicit market.

If a potential violator violates a ban he will suffer a punishment, $P$. The cost of the punishment can be thought of as the sum of all of the negative effects that attach to the sanction (fines, the unpleasantness of days incarcerated, lost liberty and forgone income while serving the sentence, etc.). But detection and punishment are not certain. The probability of being punishment conditional on having violated (let’s call this $\pi$), depends on the law-enforcement effort. The expected punishment is then $\pi P$.

The probability of punishment, $\pi = f(E, M)$, is an increasing function of law-enforcement effort, $E$ (some measure of enforcement capacity, such as the number of arrests or number of days of incarceration available), and a decreasing function of the size of the illegal market, $M$, over which enforcement resources must be spread. In this case, for example, a reasonable expression for the probability of arrest is $\pi = \frac{A}{N}$, where $A$ is the number of arrests, and $N$ is the number of illegal-market participants. As $N$ increases, the available punishment, $A$, is divided over a larger number of illicit-market participants, so that the risk facing any particular violator is reduced.

Now suppose that the number of illicit-cigarette-market participants increases (as a result of a menthol ban, for instance), with no concomitant increase in enforcement. The violator’s punishment risk, $\pi$, declines. When there are many potential violators and enforcement resources are constrained, violators face interdependent choices: the higher the overall violation rate, the smaller the risk of punishment for any individual violator. This relation yields a
positive-feedback loop, or vicious cycle: An exogenous shock (such as a ban on menthol) increases the number of violators. The ensuing greater violation rate reduces the individual risk, which, in turn, prompts others to join the illicit market, which further dilutes enforcement, which further reduces the risk of punishment, and so on. This feedback magnifies the effect of the initial increase in illicit activity and, depending on the circumstances of the illicit market, can “tip” the market into a high-violation equilibrium that is difficult to disrupt. This is why it is more expensive to break up an established market than to prevent its emergence.

Punishment is expensive, to both the agent delivering (and paying for) the punishment and to those being punished. A well-designed enforcement system should discourage noncompliance and impose no more sanctions than are required to motivate the desired compliance. This principle has been long recognized, from Beccaria (1764/1986) to recent work in behavioral economics (McAdams & Ulen, 2008): offenders evaluate potential losses and gains in making decisions, and seek risks when losses have moderate probabilities and gains have small probabilities.

One approach, in a case such as a menthol ban, is to increase the conditional probability of punishment in the illicit-cigarette market prior to imposition of the prohibition. This can reduce the total enforcement cost because the new illicit market never has a chance to take hold, and the public-health costs associated with enforcement are never realized. This increase in enforcement need not be permanent. Detection and punishment are both expensive. The upfront investment in enforcement prior to the ban can have an enduring effect on violation rates if it ensures that the system stays in a low-violation equilibrium (any given violation is easier to detect, and punishment resources are available to be meted out). But the tipping equilibria can work in both directions; both high and low violation rates are self-sustaining. In the absence of an upfront increase in enforcement, the system can tip into a high-violation equilibrium. Once an illicit market is well established, it is costly (in both enforcement and the likely violence that would ensue) to tip it back to a lower-violation equilibrium.

Illicit cigarettes account for a growing share of total cigarette consumption in the United States (LaFaive & Nesbit, 2013). This trend is driven by the large profit margins available: profits from interstate cigarette smuggling are high compared to current enforcement, which means that, for most smugglers, the benefits outweigh the risks. We use New York City as an illustration. Marlboros that sell for $4.50 per pack retail in Virginia can be resold wholesale in New York City for $7.50 (compared to the fully taxed retail price of about $12.50); a truck containing 50,000 cartons (500,000 packs) yields a gross mar-

---

gin of $1.5 million for a drive of less than 300 miles. The penalty in New York City, if caught, is $600 per carton. In expected-value terms, the smuggler breaks even on a carton if the expected profit from the sale of a carton (gross profits here would be $3 per pack x 10 packs per carton = $30/carton) equals the expected costs if caught. The expected cost is the size of the fine ($600/carton) multiplied by the probability of detection (let’s call this \(d\)). \(600d = 30\) yields a breakeven probability of detection of \(d = 5\%\). The volume of illicit cigarettes moving into New York City is massive (conservatively estimated at five-million cartons per year), and chasing cigarette smugglers is a relatively low priority for law enforcement. The law-enforcement resources devoted to tobacco enforcement in New York are minimal (the city devotes five sheriff’s deputies, two fraud investigators and a lieutenant to tobacco control) (Caruso, 2013). The size of the illicit market indicates that smugglers feel they are beating the odds.

5.2. Reduce Incentives to Smuggle

Incentives to evade taxes depend not only on the absolute levels of taxes and regulations but on their differences across jurisdictions, especially over easily crossed borders. Smugglers buy cigarettes in low-tax jurisdictions and transport them to high-tax jurisdictions. They can sell these cigarettes at a discount to the local retail price and make a sizable profit, even having paid the full retail price in the source jurisdiction. The greater the price differential (almost all of which reflect different tax rates), the greater the opportunities for profit, and the larger the number of players that would consider the risk of smuggling worthwhile.

In the United States, the state and local cigarette-excise-tax burden varies widely. Supporting a smoking habit in Virginia, which has the second-lowest tax rate in the country at $0.30 a pack, is relatively cheap, so there is little incentive for consumers to go outside the licit market and little opportunity to smuggle even cheaper cigarettes in from another state. It is much more expensive to be a smoker in New York, with a state tax of $4.35. And New York City tacks on an additional $1.50. The result is massive illicit cigarette trafficking between Virginia and New York. LaFaive and Nesbit (2013) estimate that nearly 61 percent of cigarettes sold in New York State are not fully taxed (the highest rate in the country, due also in part to illegal sales from Indian reservations). A 2012 study commissioned by the New York Association of Convenience Stores estimates that cigarette-tax evasion robs the state of $1.7 billion in tax revenues each year and 6200 jobs (John Dunham and Associates, 2011). Proximity to lower-priced states and efficient transport networks account for much of New York’s large illicit market, but smuggling is also a low-

---

32 ATF uses less conservative estimates than ours in calculating expected profits for smugglers taking cigarettes from Virginia to New York. They provide profit estimates of $4,080,000 for a single truckload.
risk enterprise in New York—the penalty for smuggling is a fine of $600 per carton (Povich, 2013).\textsuperscript{33,34}

Figure 13 uses state-level data to show the relationship between regulation strength (here the proxy for strength of regulation is the state excise tax per pack) and the size of the illicit market (expressed as a percentage of total consumption). The data show a clear positive relationship between the magnitude of the cigarette tax and the size of the illicit market. A simple linear function fit to these data finds that a one-dollar per pack tax increase is associated with a 15 percentage-point increase in the size of the illicit market.

\textsuperscript{33} The fine was raised from $150 only in June 2013. New York City is considering creating a city fine of $2,000 for selling untaxed cigarettes (and allowing the city to close violators for 60 days), with a $5,000 fine for a second offense and mandatory revocation of the retailer’s cigarette license if it occurs within three years of the first offense. The city Finance Commissioner said that most merchants who sell contraband cigarettes “think of our enforcement efforts as the cost of doing business” (Associated Press, 2013a).

\textsuperscript{34} With an operating profit per carton of approximately $40 a rational smuggler would continue with his trade if the probability of detection were less than about 0.07. There are no good data on this risk for cigarette smugglers, but the risks for dealers in illicit drugs are typically much lower (Bouchard & Tremblay, 2005), and there is no reason to assume that the risk for cigarette smugglers approaches this level.
We should also be concerned about a regulatory environment that puts huge numbers of otherwise law-abiding citizens on the wrong side of the law. Lovenheim (2008) estimates that 13 percent to 25 percent of cigarette smokers nationwide engage in casual smuggling as a result of cross-border differences. LaFaive and Nesbit (2011) reach similar conclusions. It is no surprise that New York is the state with the largest share of such smokers (again, due to its proximity to lower-priced jurisdictions and the large magnitude of the price differential). Although it is harder for authorities to prosecute these smaller cases than commercial-scale smuggling cases, the social implications of nudging people towards illicit behavior are troubling.

The problem of differential pricing can be attacked from both sides of the price divide. Raising taxes in Virginia can reduce smuggling as much as can reducing them in New York.\textsuperscript{35}

\textsuperscript{35} While equalizing taxes across states would eliminate interstate smuggling for tax avoidance or evasion, it would invite other sources of untaxed cigarettes. At the moment, interstate...
Some products sold in illicit markets are near-perfect substitutes for licit products (e.g., Virginia-sourced Marlboros sold in New York). Others are much less so (e.g., Chinese counterfeit Marlboros sold in New York). The greater the perceived quality gap, the smaller the resulting illicit market.

Thus the size of the market in untaxed cigarettes does not by itself demonstrate that a large market would emerge for illicitly manufactured menthol cigarettes in the face of a ban. Consumers of cigarettes smuggled interstate can rely on the quality-assurance efforts of licit manufacturers. Consumers of illicitly manufactured cigarettes cannot; Chinese manufacturers appear to be much better at reproducing packaging than at reproducing product quality, because the process of counterfeiting destroys any “brand name” incentive to provide desirable product. A manufacturer of low-quality product does not suffer reputational harm unless consumers or middlemen can learn to associate a given source with low quality. Apparently, street vendors of internationally smuggled and domestically produced counterfeit cigarettes have lost most of their market share to storefront sellers of smuggled genuine product because participants in that market have failed to solve the collective-action reputational-externality problem. Buyers seeking to evade taxation seem to prefer more expensive genuine smuggled product to the cheap and nasty smokes formerly available on the street (LaFaive & Nesbit 2013). In the long run, supply is elastic, so that someone will supply high-quality illicit product—if not Chinese counterfeiters, then Indian reservations or smugglers of Canadian menthol cigarettes.

Both licit and illicit cigarette markets face competition from non-cigarette products, including e-cigarettes. Cheaper and more consumer-acceptable e-cigarettes will tend to shrink both sorts of cigarette markets. On the other hand, e-cigarettes might also be complements to combustible cigarettes via gateway effects or by reducing the pressure to quit smoking altogether. The net health benefit or cost of taxing, regulating, or banning e-cigarettes depends on both the cross-elasticities and the relative health harms. A ban on mentholated combustible cigarettes would be less likely to boost the illicit market if mentholated e-cigarettes remained available.

6. Research agenda

6.1. Demand Functions

Key to analyzing the impact on licit and illicit markets of a ban on menthol cigarettes is the ability to predict how individual smokers will respond to a ban. A ban can have the effect of raising the price of menthol cigarettes. In the case of a ban with effective enforcement and full compliance, the price rises to infinity; in actuality one would expect instead that menthol cigarettes are still smuggling is the least costly and easiest approach; with that source foreclosed, we would expect the supply from Indian reservations and Canada to increase.

November 14, 2013  P a g e  |  34
available but at higher prices. A menthol smoker will respond to a price increase by cutting back on the number of menthol cigarettes smoked, by self-mentholating, by switching to non-menthol cigarettes, by cessation of smoking tobacco, or by turning to the illicit market.

Predicting how many individuals would choose each option requires knowledge of the demand functions for regular and menthol cigarettes; in particular, estimates of the price elasticity of demand for menthol cigarettes and the cross-price elasticity of demand for regular cigarettes with respect to changes in the price of menthol cigarettes. Such estimates can be arrived at by asking smokers what their plans would be under a ban or by econometric evaluation of actual demand relationships found in market data.

Economists and other social scientists generally treat actual market outcomes (“revealed preferences”) as more credible than surveys or opinion polls (“stated preferences”) for predicting human economic behavior. While a current smoker might state that he would cut back or quit smoking if prices reached a certain level, that stated intention might not be consistent with the smoker’s observed insensitivity to prices revealed by his purchases. Stated-preference data may be subject to systematic biases. For example, in a social atmosphere where intending to quit smoking is clearly “the right answer,” a survey respondent may be more likely to state that option than she would be to actually quit smoking. The revealed-preference approach that underlies econometric estimation of demand for cigarettes, on the other hand, rests on the notion that when the consumer has to “put her money where her mouth is” her actual preferences can be inferred by examining her actions. Thus, while the stated-preference approach allows for coming up with answers to counterfactual questions—“what would you do if menthol cigarettes were banned?”—the credibility of the answers is questionable (and to an unknown degree).

We are aware of only two econometric studies of demand for menthol cigarettes: Tauras et al. (2010) and Compass Lexecon (2011). Each study leaves ample scope for new work to arrive at more confident estimates of the own-price elasticity of menthol-cigarette demand and the cross-price elasticity of regular cigarettes. Both of these estimates are vitally important to analyze the impact of a ban on menthol cigarettes.

Tauras et al. (2010) estimate the impact of price changes for menthol cigarettes on one of the three relevant margins in the demand function: the choice for a smoker to buy menthol or to buy regular cigarettes. They use standard econometric modeling to find that smokers did not find menthol and regular cigarettes to be close substitutes for one another. They did not, however, analyze the impact of menthol-cigarette price changes on either the usage

---

36 As noted previously, a ban would also eliminate the brand rent of licit cigarettes, which would tend to lower the prices paid by consumers of cheap whites. It is not clear whether this effect would overwhelm the price-increasing effects of enforcement and compliant-smoker exit.
Unintended Consequences of Cigarette Taxation and Regulation

decision (how much to smoke) or the decision whether to quit smoking altogether. They also were not able to examine the most recent data from the Tobacco Use Survey, which is now available to researchers.

Similarly, several improvements can be made to the Compass Lexecon (2011) study. They estimate demand elasticities through two methods, a “margins analysis” and an econometric estimation. Their “margins analysis” is based on theoretical results for monopoly pricing in an isolated product market, whereas one of their own conclusions is that the markets for menthol and regular cigarettes are related. The econometric analysis is based on the limited data available and could be improved profitably. Their econometric results appear to be predisposed to find overly large price elasticities of demand. The price variation in the study comes from expiring promotional discounts for menthol cigarettes. However, given that cigarettes are durable goods, at best this approach picks up short-run price elasticities driven in part by stockpiling before the discounts end. For policy purposes when evaluating a ban, however, longer-run elasticities are required to predict non-transient behavior. Second, since they have data for only a single brand of menthol cigarettes, they are in essence estimating brand elasticities, not product elasticities: Demand for any one brand will be much more elastic than demand for the good (menthol cigarettes) itself at the market level. Using a brand elasticity for policy purposes when evaluating the impact of banning an entire product is undesirable. Similarly, their cross-price elasticities (which are calculated as the sensitivity of demand for non-menthol cigarettes to changes in the price of a single menthol brand) are likely to be underestimated for the same reason. Third, they have only eight cities in their panel for the cross-price-elasticity estimation. While a small sample does not necessarily bias the estimates, it does make the case for out-of-sample extrapolation (i.e., the external validity of the study) for purposes of policy evaluation harder to support. Fourth, they do not use the actual prices customers pay but instead must rely on wholesale-price data.

Moving forward, we suggest that researchers consider estimating demand functions for menthol and regular cigarettes using data from the Tobacco Use Supplement (TUS) to the Current Population Survey (CPS). Data from the 2006–07 and 2010–11 waves of the TUS are now available to researchers. The survey contains extensive data on usage, prices, and socio-demographic information. To allow for more variation in prices than is offered by state-level analysis, we recommend specifying models at three levels of granularity: state, county, and Core Based Statistical Area (CBSA) (OMB has defined 917 CBSAs, which notionally correspond to “cities”). The big question for county-level and CBSA-level data is: how do we know that the smoker bought cigarettes in the same unit where he lives? The issue of cross-jurisdictional purchases is sidestepped in state-level analysis, since (a) few people in the CPS report cross-state purchases, and (b) studies that account for cross-state purchases don’t come to radically different results. However, cross-county or out-of-CBSA pur-
chases would be harder to ignore. For this reason, we suggest pursuing specifications at all three levels of resolution.

Price and tax data can be supplemented with another commonly used source: the Tax Burden on Tobacco report. Unlike in Tauras et al. (2010), it is possible to estimate all of the relevant margins in the demand relationship: the intensive margin of how the quantity of cigarettes consumed varies with the relevant prices, and the two extensive margins of whether to smoke at all and, if so, whether to smoke menthol or regular cigarettes. Coupling the estimated elasticities with assumed illicit prices for menthol cigarettes will allow researchers to predict how many menthol smokers will quit altogether, how many will switch to regular cigarettes, and how many will continue to smoke (illicit) menthol cigarettes. For the latter group it is possible to predict how their usage would decrease, if at all. This proposed analytical approach would be expected to yield different estimates of a response to a ban than those suggested by stated-preference studies. For example, O'Connor et al. (2011) state that 36 percent of menthol smokers report that they would respond to a ban by trying to quit smoking. The estimated cross-price elasticity between menthol and non-menthol cigarettes may in fact lead to a substantially different prediction (in either direction). In that case, the stated- and revealed-preference approaches can add depth to the analysis of a ban by providing a reasonable range of possible outcomes.

6.2. Analysis of Illicit Markets

More research is required on the likely impact of a menthol ban on illicit markets. To the extent that a ban enlarges the illicit tobacco trade, its health-protective effects will be offset: enforcement of the ban will yield arrests, incarceration, and violence in the illicit market, all of which are known to adversely affect the health of those involved. A comparison to the existing illicit cigarette market in the United States offers only limited insight. The illicit market (principally in legitimate cigarettes trafficked across state and country borders to avoid taxes) is large but has seen a relatively low level of associated violence (although there are indications that more-violent actors are entering the market). The case of an illicitly trafficked product, the consumption of which is otherwise legal, is quite distinct from an inherently illicit product. Participants in today's cigarette black market may either not know or fully understand the illegality of their conduct, whereas a menthol user under a ban would be flagrantly skirting the law.

Better insights might be had from studying marijuana markets. Until recently, any marijuana user was aware that she was engaging in an illegal activity (liberal medical-marijuana laws in many states, and state-level legalization in Washington and Colorado, make this more complicated now). A menthol smoker, not otherwise disposed to criminal behavior, who now finds himself a criminal may be disinhibited to other criminal activity, breeding disrespect for
legitimate authority and eroding social order. A large share of law-enforcement
resources has been devoted to enforcing marijuana prohibition (at the expense
of other crime-control efforts). There are many unknowns regarding the in-
crease in violence that might follow from a menthol ban. Increasing
enforcement in illicit markets tends to foster violence within those markets.
(And, unlike in licit markets, increasing competition in illicit markets does not
necessarily reduce prices. In a vicious cycle, violence among suppliers compet-
ing for market share elevates the risk premium to participate in the market,
thereby maintaining high prices, revenues, and continued violence.)

When risk premium are high, profits become big enough to fight for. And
Mexico looms large. If Mexico were to become a significant source for illicit
menthol cigarettes it is possible that its drug-trafficking organizations (notor-
ious for their propensity to violence) would enter the trade; they have many
strategic advantages including well-established smuggling networks and an
established dominance in many border areas. A much better understanding of
the likely violence response is needed to help authorities plan for the law-
enforcement resources that will be required to keep the illegal market in check.

And, finally, many of the factors essential to the modeling laid out in this
paper are themselves poorly understood. If we are to estimate the effects of
alternative policies, we require better baseline data on the kinds and extent of
violence attendant to the illicit cigarette trade, the levels and cost of en-
forcement, the sizes of illicit cigarette markets, and the characteristics of illicit
distribution channels.

7. Conclusions

The risks of evasion and the need for enforcement complicate the prob-
lems of tobacco regulation. Higher taxes and tighter regulations, especially
those that create strong gradients across easily-crossed borders, may have
substantial unintended and unwanted consequences, and at some point the
costs of further tightening might exceed the benefits.

Evasion incentives depend not only on the absolute level of taxes and
regulations but on their variation, especially over easily-crossed borders. Low-
ering taxes in high-tax juris-dictions that consume smuggled cigarettes can
reduce the profitability of smuggling, but so can raising taxes in low-tax source
jurisdictions (which may itself invite other evasions and harms).

Banning specific product types (such as mentholated cigarettes) will tend
to ad-antage illicit substitutes. Health gains from reduced consumption due to
such bans need to be offset with health and other losses from larger illicit mar-
kets. Not applying such par-tial bans to competing nicotine vehicles such as e-
cigarettes will tend to reduce sales of both licit and illicit tobacco cigarettes. A
ban on mentholated tobacco cigarettes would be less likely to boost the illicit market if mentholated e-cigarettes remained available.

In the United States current penalties against tobacco smuggling appear to be too low to support some of the existing price differentials, suggesting the need for greater enforcement to prevent growth in illicit-market shares (already high in some places). Illicit markets have positive feedbacks through enforcement swamping and normalization. Increased enforcement will be more efficacious and less costly in terms of expenditures and side effects if applied before market growth than after market growth. It might therefore be advantageous to delay the introduction of potentially illicit-market-enhancing regulatory changes until adequate enforcement capacity has been put in place.

Too much remains unknown to predict the likely consequences of a ban, sufficient to formulate policy. We propose a research agenda that would provide better estimates than are currently available of the demand function for menthol cigarettes, which would allow for predicting the behavioral response to a ban. The violence attendant to existing markets for illicit drugs compels a better understanding of the relation between available revenues and violent competition for market share.

Acknowledgments

This study was funded by Cornerstone Research under contract to Altria Client Services. The funding source had no role in the writing of the paper and exercised no editorial control. Jeremy Ziskind provided invaluable research assistance.

References


Caraballo, Ralph S. “Menthol and Demographics.” Presentation on behalf of the CDC before the FDA Tobacco Products Scientific Advisory Committee (March 30). 2010.


Unintended Consequences of Cigarette Taxation and Regulation


### Appendix: Studies Reviewed and a Description of Their Findings

<table>
<thead>
<tr>
<th>Author, Year</th>
<th>Location</th>
<th>Study Design</th>
<th>Study Period</th>
<th>Main Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Goldstein, Brownstein, Ryan, &amp; Bellucci, 1989</td>
<td>New York City</td>
<td>Longitudinal observational study of 414 homicide events</td>
<td>8-month period in 1988</td>
<td>Positive association between enforcement and homicides. Nearly 40% of all homicides events were “systemic.” This was regarded as primarily due to prohibition/enforcement effects.</td>
</tr>
<tr>
<td>Rasmussen &amp; Benson, 1994</td>
<td>Florida</td>
<td>Longitudinal observational study of 67 Florida counties</td>
<td>1989</td>
<td>Increased drug enforcement has spillover effect; it increases the size of the market in adjoining jurisdiction, resulting in higher violent crime.</td>
</tr>
<tr>
<td>Brumm &amp; Cloninger, 1995</td>
<td>USA</td>
<td>Longitudinal observational study of 57 US cities</td>
<td>1985</td>
<td>No significant association between enforcement and violence.</td>
</tr>
<tr>
<td>Benson, Kim, &amp; Rasmussen, 1998</td>
<td>Florida</td>
<td>Longitudinal observational study of 67 Florida counties</td>
<td>1983–1987</td>
<td>Measures of enforcement were significantly and positively related with violent crime.</td>
</tr>
<tr>
<td>Riley, 1998</td>
<td>6 US cities</td>
<td>Longitudinal observational study of 6 US cities</td>
<td>1995</td>
<td>Increased enforcement was associated with increased homicide rates in 4 of the 6 cities studied.</td>
</tr>
<tr>
<td>Miron, 1999</td>
<td>USA</td>
<td>Longitudinal observational study in the US</td>
<td>1900–1995</td>
<td>Enforcement is positively and significantly related to the homicide rate over the study period.</td>
</tr>
<tr>
<td>Levitt &amp; Venkatesh, 2000</td>
<td>Chicago</td>
<td>Longitudinal observational study in Chicago</td>
<td>4-year period in the 1990s</td>
<td>Lack of formal dispute-resolution mechanism in illicit market and drug-law enforcement prompted a high level of violence.</td>
</tr>
<tr>
<td>Benson, Leburn, &amp; Rasmussen, 2001</td>
<td>Florida</td>
<td>Longitudinal observational study of 67 Florida counties</td>
<td>1994–1997</td>
<td>Increases in the rate of drug arrests were associated with an increase in violent crime.</td>
</tr>
<tr>
<td>Miron, 2001</td>
<td>USA</td>
<td>Longitudinal observational study in the US</td>
<td>1993–1996</td>
<td>6 of 9 enforcement proxies were positively and significantly related to the homicide rate.</td>
</tr>
<tr>
<td>Shepard &amp; Blackley, 2007</td>
<td>USA</td>
<td>Longitudinal observational study of 1300 US counties</td>
<td>1994–2001</td>
<td>Enforcement was positively and significantly associated with all crime, including violent crime.</td>
</tr>
<tr>
<td>Owens, 2013</td>
<td>USA</td>
<td>Longitudinal observational study using state-level variation in homicides</td>
<td>Data span passage–repeal of temperance laws</td>
<td>Greater enforcement during Prohibition was not associated with an increase in homicides.</td>
</tr>
<tr>
<td>Moeller &amp; Hesse, 2013</td>
<td>Denmark</td>
<td>Longitudinal observational study of 269 jurisdictions in Denmark</td>
<td>2000–2009</td>
<td>A significant relationship between a policy crackdown (which led to an increase in arrests) and charges for serious violent crime in the year that followed.</td>
</tr>
</tbody>
</table>