Stigma Dilution and Over-Criminalization

Murat C. Mungan[*]

March 10, 2015

Abstract

Criminalizing an act that provides weak signals about a person's productivity and character can dilute the stigma attached to having a criminal record. This reduces the deterrence of serious crimes that do provide strong signals regarding the offender's character. Over-criminalization occurs when the costs associated with reduced deterrence due to stigma dilution off-set potential benefits associated with criminalizing the less harmful act. Identifying conditions under which stigma dilution is likely and comparatively costly allows the determination of factors that affect the desirability of (de-)criminalizing various acts. These factors are discussed in the context of marijuana possession offenses to illustrate how over-criminalization may reduce social welfare. The normative desirability of various practices in criminal law (e.g. the felony murder rule, conspiracy liabilities, scienters, mens rea requirements, expungements) are also discussed vis-à-vis their impacts on stigma dilution.

Keywords: Stigma, Over-criminalization, Deterrence, Crime and Punishment, Criminal Records, Expungements, Scienters, Mens Rea, Conspiracy Liability, Felony Murder Rule, Marijuana Possession.

JEL classification: K00, K14, K42.

1. Introduction

In some places in the United States practicing interior design without a license,[1] lying in a fishing tournament,[2] dredging for oysters at night,[3] and possessing marijuana are criminal offenses.[4] Given this enormous breadth of criminal law, it is not very surprising that more than 10 million crimes are committed per year,[5] which naturally leads to more than 4.5 million people being under supervision[6] and more than 1.5 million people in prison in a year.[7]

---

[*]Assistant Professor of Law, Florida State University College of Law; Courtesy Professor of Economics, Florida State University College of Social Sciences.

[1] *See, e.g.,* Fla. Stat. §481.223.

[2] *See, e.g.,* Tex. Parks & Wild. §66.023.

[3] *See, e.g.,* Tex. Parks & Wild. §76.109.

[4] *See, e.g.,* Fla. Stat. §398.03.

[5] More than 10 million crimes were committed per year between 1994-2012 (United States Department of Justice, Federal Bureau of Investigation. *Crime in the United States*, 2012).

[6] In 2012, 4,781,300 people were under supervision (Office of Justice Programs, Bureau of Justice Statistics, http://www.bjs.gov/index.cfm?ty=pbdetail&iid=4844).

[7] There were 1,571,013 people in prison in 2012 (Office of Justice Programs, Bureau of Justice Statistics, http://www.bjs.gov/content/pub/pdf/p12ac.pdf).

There are many concerns raised by the broadening of criminal law, including, the loss of individual liberties; increased costs of law enforcement; increased corruption and rent-seeking; chilling of desirable behavior not intended to be regulated by criminal law; and the distorting of law enforcers' incentives. Some of these concerns are particularly relevant to the emerging academic literature on the likely benefits associated with the decriminalization of some drug possession offenses. Most recently, Adda et al. (2014) have attributed a drop in crime rates in a borough in London to the depenalization of marijuana possession, which they associated with the reallocation of law enforcement resources to combat other crimes. A less apparent potential loss associated with the criminalization of minor offenses, including marijuana possession, is a reduction in the stigma attached to being convicted for a serious crime, such as murder, rape or theft. This effect, which I call *stigma dilution*, and its relation to the optimal scope of criminal law, are the topics of this paper.

Stigma dilution takes place when the signal provided by a serious conviction becomes noisier, and therefore less informative, due to some legal change that pools in relatively highly productive individuals with people who have high criminal tendencies. Increasing the scope of criminal law, in particular, can cause stigma dilution by broadening the pool of criminals, and therefore reducing their average deviation from the average citizen. Although this potential consequence has been noted by some scholars no formalization of it exists in the economics of law enforcement literature. Perhaps due to this reason, there exists no notable analysis that identifies factors affecting the magnitude of this dilution, and no study that provides normative comments as to whether, why, and when, such dilution may be socially undesirable.

In this article, I provide a Beckerian law enforcement model[8] that incorporates stigma through an adverse selection model. In this model, criminal activity is a proxy for lower productivity, and therefore employers, out of self-interest, offer lower wages to ex-convicts. I consider the effects of adding an additional act into the domain of criminal law under two sets of assumptions. First, I consider a situation where employers distinguish between different types of criminal records, *e.g.* convicted for marijuana possession versus murder, to gain more precise information about the offender. Then, I consider the case of *record confusion* where employers either cannot distinguish between criminal records (because they are literally confusing) or they choose not to distinguish between records (because it is prohibitively costly). My intention in focusing only on the labor market consequences of being stigmatized is to simplify the analysis, and the implications of my analysis can only be magnified if negative social effects are also factored in.

Perhaps counter-intuitively, even under the less restrictive set of assumptions (where there is no record confusion) criminalization of acts that provide weak signals regarding a person's productivity lead to stigma dilution. The reason is the following. Although a conviction for a serious offense accurately reflects the

---

[8] Becker (1968). *See also*, Polinsky and Shavell (2007) for a collection of extensions and applications of Beckerian law enforcement models.

low productivity of the convict regardless of whether or not the lesser offense is criminalized, the signal provided by other, less serious, records is affected by the criminalization of the lesser offense. In particular, decriminalizing the lesser harmful act increases the number of relatively productive individuals who do not have criminal records. Hence, it pays off more for people with high criminal tendencies to refrain from committing crime, and mimicking the behavior of people with lower criminal tendencies. This allows them to be pooled in with individuals who are more productive and therefore earn more in the labor market. Hence, when there is no record confusion, criminalizing the lesser offense leads to stigma dilution by reducing the value associated with refraining from crime, i.e. criminalization produces a less valuable *carrot*.

Stigma dilution is exacerbated when there is record confusion: criminalization not only reduces the value of the carrot, it reduces the impact of the *stick*. When employers do not distinguish between different types of criminal records for purposes of hiring, they pool together people who were convicted for minor offenses and major offenses. Therefore, the average productivity of these individuals is higher than the productivity of a person with high criminal tendencies. Hence, the legal earning of a person with a serious criminal record is increased when the less harmful act is criminalized.

These observations naturally imply that broadening the scope of criminal law reduces the deterrence of wrongful acts that were already designated as crimes. This is bad from a utilitarian perspective, if such reductions in deterrence translate into harms that more than off-set the harms avoided by the creation of new criminal categories. In other words, the marginal social benefits associated with criminalizing an act can be off-set by the marginal social costs associated with the dilution of criminal stigma. I use the term over-criminalization to refer to these instances.

I identify four factors for purposes of detecting over-criminalization: the harm associated with the act, whether the commission of the act sends a strong signal about the productivity and/or destructiveness of the individual who committed it, the likelihood with which the act can be confused with another crime, and the number of people who would like to commit the lesser harm generating act. By focusing on these factors I discuss the benefits of using strong scienters in defining crimes, the harms associated with liability imputation (e.g. through conspiracy liability or the felony-murder rule), the central role that *mens rea* plays in choosing which acts to criminalize, and how expungements can be used to reduce stigma dilution.

In the next section I briefly review the relevant literature. In section 3, I present an economic model of law enforcement which I use in section 4 to illustrate that criminalizing minor acts lead to stigma dilution. Section 5 discusses conditions under which criminalizing an act is likely to reduce welfare, i.e. conditions of over-criminalization. The same section also discusses policy implications. Section 6 briefly discusses a couple of alternatives to assumptions used in the model and their likely impact on the model's implications, and section 7 concludes.

2. Literature Review

The economics of criminal stigmatization is not a new topic. Lott (1992a) and (1992b) empirically document the reduction in criminals' post-conviction earnings, and show that these reductions are greater for higher income people. One mechanism through which criminals' earnings can be reduced is proposed in Rasmusen (1996). Rasmusen formalizes the idea that criminal convictions can signal lower labor productivity. Rational employers who make hiring decisions in a competitive labor market pay individuals their expected marginal contributions, and therefore the wages for convicts are lower than the wages for people with clean records.

Since Rasmusen (1996), there have been attempts to analyze the effects of criminal stigmatization, but these are limited in number. Funk (2004) notes that although stigma can increase general deterrence, it can have the perverse effect of reducing specific deterrence, i.e. convicts can be tempted to become recidivists, because they no longer have as much to lose by adding another crime on to their criminal records. Consistent with Funk's analysis, Prescott and Rockoff (2011) present evidence that sex offender registration requirements increase general deterrence while reducing specific deterrence.

Harel and Klement (2007) and Galbiati and Garoupa (2007) are the two articles in the literature that are most closely related to stigma dilution. Harel and Klement (2007) analyze the effects of search costs on the stigmatization process and find that under some parametric conditions increasing the probability of detection can reduce deterrence by increasing search costs. This effect is conditional on, among other things, the existence of third party pressures on people who deal with ex-convicts. Galbiati and Garoupa (2007), on the other hand, provide an explanation for the widely accepted presumption that administrative penalties generate lesser stigma than criminal sanctions. They note that because the information provided by criminal convictions are more reliable (due to stronger pro-defendant rules) they result in greater stigma.

There also exist works on the optimal scope of criminal law. Shavell (1993), Bowles et al. (2008), Mungan (2012), and Kim (2013) focus on what distinguishes criminal law from other modes of regulation, such as civil law and administrative law. By studying these distinctions they identify factors that affect the normative desirability of including an act within the reach of criminal law.

Articles reviewed so far include those that use economics to analyze stigma, and some of these works touch on the topic of stigma dilution. Moreover, there is a body of work that focuses on the optimal scope of criminal law. However, there is no work that analyzes the effect of stigma dilution on the optimal scope of criminal law. The current article fills this gap in the literature.

It is also worth noting that there is a sizeable literature that focuses on the effect of decriminalizing or depenalizing marijuana possession on the deterrence of other crimes. One of the most recent contributions, Adda et al. (2014), presents evidence from an experiment from the London borough of Lambeth on the positive effects of depenalization of some cannabis possession offenses on the deterrence of other offenses. Other studies exist, e.g. Kuziemko and Levitt (2004), that focus on the link between other crimes and the incarceration of

drug offenders. These articles do not focus on stigma dilution as the potential source of reductions in the deterrence of other crimes.[9] Hence, the instant article presents a theory that links the rates of other crimes to the legal status of drug possession crimes, which can potentially be tested empirically.

Finally, there are many commonalities between criminal stigmatization and reputational sanctions imposed on producers. Iacobucci (2014), for instance, uses a model that is very similar to Rasmusen (1996) to analyze reputational sanctions. As demonstrated in Mungan (2014a), Iacobucci's model and Rasmusen's adverse selection model emerge as special cases of a generalized model of informal sanctions. This suggests that the implications of the current article may carry on to other contexts besides the field of criminal law and criminal stigmatization. This possibility is explored further in section 6.

3. Model:

Individuals weigh the costs and benefits of committing a crime in a Beckerian framework. There are two groups of individuals; those with high criminal propensities (type H) and those with low criminal propensities (type L). The ratio of type L to type H individuals is $\frac{\lambda}{1-\lambda}$. Similar to Rasmusen (1996), Iacobucci (2014), and Mungan (2014a), it is assumed that type H individuals have lower productivity (q_H) in the legal labor market compared to type L individuals (q_L), such that $q_L - q_H \equiv v > 0$. Two types of harmful acts can be committed by these individuals; one that inflicts great harms and another that causes small negative externalities. For easing references, I call the high harm act a 'felony' and the low harm act an 'infraction', although the government may or may not choose to criminalize the latter act. I denote felonies by ϕ and infractions by ι.

Felonies as well as infractions are harmful. Therefore, the government has an interest in deterring these acts, potentially by declaring them crimes, and punishing them. The government has two options regarding criminalization: although it always criminalizes felonies, it decides whether or not to criminalize infractions. The sanction for felonies is s^ϕ, and the sanction for infractions, if criminalized, is s^ι. The probability of detection and punishment is p.[10] Punishment also produces a signal, in the form of a criminal record, that is used by employers in the labor market to determine the wages offered to individuals. The reduction in one's wages due to having a criminal record is called 'stigma', and the precise relationship between the magnitude of stigma, criminalization of infractions, and crime rates is derived in the next section. Individuals compare expected formal and informal sanctions to their benefits from crime to decide whether to commit crime.

Individuals' benefits from committing ϕ and ι are denoted as B and b, respec-

tively, and people vary in their benefits from committing crimes. It is assumed that type L individuals either have no interest in committing ϕ or that the expected sanctions for committing ϕ are sufficient to deter these individuals from committing ϕ, i.e. $B < ps^\phi$ for all type L individuals. Type H individuals, on the other hand, are undeterred from committing infractions in equilibrium, i.e. $b > p(s^\iota + \sigma^\iota)$ for all type H individuals, where σ^ι is the maximum stigma associated with having an infraction on one's criminal record. This assumption is mainly to simplify the exposition of results, and relaxing it is likely to bolster the results presented, as is argued in section 6.

Employers are assumed to be self-interested profit-maximizers who rationally offer employees wages (w) that equal their expected productivity, which they estimate based on their criminal records. If infractions are not criminalized, only two criminal records are possible; a clean record and a felony record, which lead to wages of w^ν and w^ϕ, respectively. When infractions are criminalized, an additional wage is offered to those who only have an infraction record, namely w^ι. Wages in the case where there is record confusion are w^ψ for all individuals with criminal records, and w^ν for people with clean records. All wages are determined by calculating the equilibrium number of type H and type L individuals in each group and the corresponding average productivity, which always lies in between q_H and q_L.

The precise interactions between the labor market, the government's choice of criminalization, and crime rates are analyzed in the next section with references to the notation introduced in this section.

4. Analysis:

The analysis proceeds by first determining equilibrium wages and crime rates when infractions are not criminalized. The second sub-section determines the same equilibrium values when infractions are criminalized and formalizes the stigma dilution effect of criminalization. The third sub-section considers the case of record-confusion, and shows that stigma dilution is exacerbated when record-confusion exists.

4.1. Non-Criminalization of Infractions (Case L)

When the infraction is not made a crime, all individuals commit infractions, because they face no negative consequences, and as assumed, type L individuals are either deterred from, or have no interest in, committing felonies. Type H individuals, on the other hand, must decide whether or not to commit felonies. Committing a felony (in addition to committing the infraction which is not criminalized) leads to expected net-benefits of:

$$\Pi_L^\phi = B + b + p(w_L^\phi - s^\phi) + (1-p)w_L^\nu \tag{1}$$

where the sub-scripts for wages reflect the fact that infractions are not criminalized (i.e. L for legal).

Refraining from committing a felony, on the other hand, leads to the higher wage of w^ν with certainty, i.e.

$$\Pi_L^\nu = b + w_L^\nu \tag{2}$$

Hence, a type H individual commits a felony, if $\Pi_L^\phi > \Pi_L^\nu$, i.e. if:

$$B > p(s^\phi + w_L^\nu - w_L^\phi) = p(s^\phi + \sigma_L^\phi) \equiv B^L \tag{3}$$

where σ_L^ϕ is the stigma associated with having a felony record. Therefore, the proportion of type H individuals committing felonies is given by: $\left(1 - G(B^L)\right) \equiv \Theta$, where G is the cumulative distribution associated with B, with $G(0) = 0$, $\lim_{B \to \infty} G(B) = 1$, and $G'(B) > 0$ for all $B \geq 0$.

This implies that $(1 - p)\Theta$ type H individuals have clean records despite committing crime, and that $(1 - \Theta)$ individuals do not have records, because they have refrained from committing crime. Hence, there are a total of $(1 - \lambda)(1 - p\Theta)$ type H individuals and λ type L individuals, who do not have criminal records, where λ, as stated in the previous section, is the measure of type L individuals. Thus, the average productivity of people with clean records, and therefore the wage for people with clean records, is

$$w_L^\nu = \frac{\lambda q_L + (1 - \lambda)(1 - p\Theta) q_H}{\lambda + (1 - \lambda)(1 - p\Theta)} = q_H + v \frac{\lambda}{\lambda + (1 - \lambda)(1 - p\Theta)} \tag{4}$$

where $v = q_L - q_H$.

The wage for individuals with felony records, on the other hand, is equal to the productivity of type H individuals, since only they commit felonies.

$$w_L^\phi = q_H \tag{5}$$

Therefore, the effect of stigma (associated with having a felony record) is given by:

$$\sigma_L^\phi = w_L^\nu - w_L^\phi = v \frac{\lambda}{\lambda + (1 - \lambda)(1 - p\Theta)} \tag{6}$$

As this expression demonstrates, stigma is a function of the crime rate, which as is reflected in (3) affects type H individuals' decisions to commit a felony. Therefore, an equilibrium exists only when the observed crime rate is equal to type H individuals' anticipated crime rate. Equivalently, it is required that the person who is indifferent between committing and not committing crime given his accurate expectation of the crime rate have the lowest B among all type H individuals who commit the felony, i.e. the equilibrium condition is:

$$B_L^* = p(s^\phi + \sigma_L^\phi(\Theta)) = G^{-1}(1 - \Theta) = B^s(\Theta) \tag{7}$$

The next proposition establishes that there is always a felony rate Θ^* that satisfies this condition.

Proposition 1: *There exists a stable equilibrium where $(1 - \lambda)\Theta_L^*$ felonies are committed when infractions are not punished.*

Proof: As is evident from (6) and (7) $B_L^*(\Theta)$ and $B^s(\Theta)$ are both functions that are decreasing in Θ, and $\lim_{\Theta \to 0} B^s(\Theta) > B_L^*(0) > B_L^*(1) > B^s(1) = 0$. Hence, the intermediate value theorem implies that there exists Θ^* such that

$B_L^*(\Theta^*) = B^s(\Theta^*)$, which is the equilibrium condition. That there exists a stable equilibrium follows from the fact that $\lim_{\Theta \to 0} B^s(\Theta) > B_L^*(0)$, which implies that the first intersection (i.e. that intersection that occurs at the smallest Θ among all intersections) of B^s and B_L^* occurs when $\frac{dB^s}{d\Theta} < \frac{dB_L^*}{d\Theta}$, which is the requirement for stability.■

In the remaining parts, to simplify the analysis, I will assume that there is a single equilibrium. However, if there were multiple equilibria, all results presented in this section would be equally valid around any stable equilibrium.

4.2. Criminalization of Infractions: No Record Confusion (Case I)

When infractions are criminalized, type L individuals commit these acts only if their net-benefits from doing so exceed the value of the legal option, i.e. when

$$\pi_I^\iota = b + p(w_I^\iota - s) + (1-p)w_I^\nu > w_I^\nu = \pi_I^\nu \tag{8}$$

which is equivalent to:

$$b > p(s^\iota + w_I^\nu - w_I^\iota) = p(s^\iota + \sigma_I^\iota) \equiv b_I^* \tag{9}$$

where σ_I^ι is the stigma associated with having an infraction record.

The pay-offs of type H individuals from committing and refraining from committing felonies, on the other hand, are given by:

$$\Pi_I^\phi = B + b + p(1-p)(w_I^\phi - s^\phi + w_I^\iota - s^\iota) + p^2(w_I^\phi - s^\phi - s^\iota) + (1-p)^2 w_I^\nu \tag{10}$$

and

$$\Pi_I^\nu = b + p(w_I^\iota - s^\iota) + (1-p)w_I^\nu \tag{11}$$

Therefore, type H individuals commit felonies, if

$$B > p[s^\phi + p w_I^\iota + (1-p)w_I^\nu - w_I^\phi] \equiv B_I^* \tag{12}$$

which implies that the expected stigma associated with committing a felony, conditional on being detected, is

$$\sigma_I^\phi = p w_I^\iota + (1-p)w_I^\nu - w_I^\phi \tag{13}$$

Wages are given by the average productivity of individuals with a felony record, an infraction record, and no record. θ denotes the proportion of type L individuals who commit infractions. Hence, $\lambda p\theta$ type L individuals have infraction records. In addition p proportion of the $(1-\lambda)(1-\Theta)$ type H individuals who do not commit felonies are nevertheless sanctioned for infractions, and therefore receive a wage of w_I^ι. Finally, p proportion of the $(1-\lambda)(1-p)\Theta$ type H individuals who commit felonies, but are only caught for infractions, have infraction records only. p^2 of the Θ type H individuals who commit felonies, and are sanctioned for both infractions and felonies, are not included in the pool of individuals who are paid wages of w_I^ι, because they reveal their types with certainty, and are paid w_I^ϕ. Hence,

$$w_I^\iota = \frac{\lambda p\theta q_L + (1-\lambda)p(1-p\Theta)q_H}{\lambda p\theta + (1-\lambda)p(1-p\Theta)} = q_H + v\frac{\lambda p\theta}{\lambda p\theta + (1-\lambda)p(1-p\Theta)} \tag{14}$$

8

Among people who have no criminal records, $\lambda(1-p\theta)$ are type L individuals, since $\lambda(1-\theta)$ of them do not commit infractions, and $\lambda(1-p)\theta$ of them are not caught despite committing infractions. On the other hand, $(1-\lambda)(1-p)(1-\Theta)$ type H individuals commit infractions only but are not caught, and $(1-\lambda)(1-p)^2\Theta$ type H individuals have clean records despite committing a felony and an infraction. Hence, the average productivity of individuals who have clean records is given by:

$$w_I^\nu = \frac{\lambda(1-p\theta)q_L + (1-\lambda)(1-p)(1-p\Theta)q_H}{\lambda(1-p\theta) + (1-\lambda)(1-p)(1-p\Theta)} =$$
$$q_H + v\frac{\lambda(1-p\theta)}{\lambda(1-p\theta) + (1-\lambda)(1-p)(1-p\Theta)} \quad (15)$$

Because only type H individuals commit felonies, the wage for people with a felony record is given by:

$$w_I^\phi = q_H \quad (16)$$

The next lemma formalizes the intuitive result that people with felony records earn less in the labor market than people with infraction records, and that people with clean records earn the most.

Lemma 1: $w_I^\nu > w_I^\iota > w_I^\phi$.

Proof: $w_I^\phi = q_H$, and therefore $w_I^\phi < \min\{w_I^\nu; w_I^\iota\} = q_H + v\rho^{j\in\{\nu,\iota\}}$, where $\rho > 0$ is the proportion of type L individuals among people with a j record. That $w_I^\nu > w_I^\iota$ can be seen by comparing (14) and (15):

$$w_I^\nu = q_H + v\frac{\lambda(1-p\theta)}{\lambda(1-p\theta) + (1-\lambda)(1-p)(1-p\Theta)}$$
$$> q_H + v\frac{\lambda p\theta}{\lambda p\theta + (1-\lambda)p(1-p\Theta)} = w_I^\iota \quad (17)$$

iff

$$(1-p\theta)\lambda\theta + (1-p\theta)(1-\lambda)(1-p\Theta) > \lambda(1-p\theta)\theta + (1-\lambda)(1-p)(1-p\Theta)\theta \quad (18)$$

which holds iff

$$1 > \theta \quad (19)$$

∎

Equilibrium wages, which, as demonstrated by lemma 1, satisfy some intuitive properties, depend on crime rates, and the crime rates depend on wages. This interdependency generates two equilibrium conditions similar to the analogous equilibrium requirement expressed by (7) when infractions are not illegalized:

$$B_I^* = p(s^\phi + \sigma_I^\phi(\theta, \Theta)) = G^{-1}(1-\Theta) = B^s(\Theta) \quad (20)$$

and

$$b_I^* = p(s^\iota + \sigma_I^\iota(\theta, \Theta)) = F^{-1}(1-\theta) \equiv b^s(\theta) \quad (21)$$

where F is the cumulative distribution function associated with type L individuals' benefits from infractions, with $F(0) = 0$, $\lim_{b \to \infty} F(b) = 1$, and $F'(b) > 0$ for all $b \geq 0$. The next proposition proves that a pair θ_I^*, Θ_I^* exists that satisfies both conditions.

Proposition 2: *There exists a stable equilibrium where the infraction rate among type Ls is θ_I^*, and the felony rate among type Hs is Θ_I^* when infractions are criminalized.*

Proof: Define $\theta^m(\Theta)$ as the infraction rate that satisfies $b_I^* = b^s$ when the felony level is Θ, and define $\Theta^m(\theta)$ as the felony level that satisfies $B_I^* = B^s$ when the infraction rate is θ. A simple application of Brouwer's fixed point theorem implies that there exists Θ_I^* such that $\Theta^m(\theta^m(\Theta_I^*)) = \Theta_I^*$. It follows that Θ_I^* and $\theta_I^* \equiv \theta^m(\Theta_I^*)$ constitute an equilibrium pair, since they satisfy conditions (20) and (21). That the equilibrium is stable is guaranteed by the same arguments as in the proof of proposition 1 separately applied to conditions (20) and (21).∎

Given that a stable equilibrium exits, the effect of criminalizing infractions on the equilibrium felony level can be determined by comparing Θ_L^* and Θ_I^*. The next proposition makes this comparison and concludes that $\Theta_L^* < \Theta_I^*$, due to the stigma diluting effect of criminalization.

Proposition 3: *(i) Criminalizing infractions dilutes the stigma associated with committing a felony, i.e. $\sigma_I^\phi < \sigma_L^\phi$. (ii) Criminalizing infractions increases the number of felonies committed $\Theta_L^* < \Theta_I^*$.*

Proof: (i)

$$\sigma_L^\phi = w_L^\nu - w_L^\phi > pw_I^\iota + (1-p)w_I^\nu - w_I^\phi = \sigma_I^\phi \tag{22}$$

iff

$$w_L^\nu > pw_I^\iota + (1-p)w_I^\nu \tag{23}$$

since $w_L^\phi = w_I^\phi$. Plugging in the expressions for w_L^ν, w_I^ι and w_I^ν implies that this condition is equivalent to:

$$\frac{1}{\lambda + (1-\lambda)(1-p\Theta)} > \frac{p\theta}{\lambda\theta + (1-\lambda)(1-p\Theta)} + \frac{(1-p)(1-p\theta)}{\lambda(1-p\theta) + (1-\lambda)(1-p)(1-p\Theta)} \tag{24}$$

letting $A \equiv (1-p\Theta)$, this expression becomes:

$$\frac{1}{\lambda + (1-\lambda)A} > \frac{p\theta}{\lambda\theta + (1-\lambda)A} + \frac{(1-p)(1-p\theta)}{\lambda(1-p\theta) + (1-\lambda)(1-p)A} \tag{25}$$

which is equivalent to:

$$\frac{1}{\lambda + (1-\lambda)A} > \tag{26}$$

$$\frac{\lambda(1-p\theta)\theta + (1-\lambda)(1-p)A}{(\lambda + (1-\lambda)A)(\lambda(1-p\theta) + (1-\lambda)(1-p)A) - \lambda(1-\theta)(\lambda(1-p\theta) + (1-\lambda)(1-p)A)}$$

letting $K \equiv \lambda + (1-\lambda)A$; $M \equiv \lambda(1-p\theta) + (1-\lambda)(1-p)A$; and $N \equiv \lambda(1-p\theta)\theta + (1-\lambda)(1-p)A$, (26) becomes:

$$\frac{1}{K} > \frac{N}{KM - \lambda(1-\theta)M} \tag{27}$$

which holds when:

$$K(M-N) > \lambda(1-\theta)M \tag{28}$$

simplifying $M-N$ and re-plugging in the expressions for K and M, we have that:

$$(\lambda + (1-\lambda)A)\lambda(1-p\theta)(1-\theta) > \lambda(1-\theta)(\lambda(1-p\theta) + (1-\lambda)(1-p)A) \tag{29}$$

which when simplified yields:

$$1 > \theta \tag{30}$$

(ii) Part (i) implies that $B_L^*(\Theta) > B_I^*(\Theta, \theta)$, and thus $\Theta_I^* > \Theta_L^*$. ∎

Proposition 3 formalizes the intuition conveyed in the introduction regarding stigma dilution when there is no record confusion. Because there is no record confusion, employers are able to accurately assess the average productivity of individuals with felony records and distinguish them from people who only have infraction records. Hence, criminalization has no effect on the stick component of stigma: it is equally bad to have a felony record under both regimes, i.e. $w_L^\phi = w_I^\phi$. Rather, stigma dilution occurs due to a devaluation of the carrot component: the expected benefit associated with refraining from committing a felony is smaller when infractions are criminalized, i.e. $w_L^\nu > pw_I^\iota + (1-p)w_I^\psi$.

4.3. Criminalization of Infractions: The Case of Record Confusion (Case C)

When there is record confusion, the employer does not distinguish between felonies and infractions while hiring, and instead makes a decision based on whether or not the person has *any* criminal record. In this case there is a single wage for people with records, namely w_C^ψ, and a wage for people with clean records (w_C^ν).

Type H individuals commit felonies if the following condition holds:

$$\begin{aligned}\Pi_C^\phi &= B + b + p(1-p)(2w_C^\psi - s^\iota - s^\phi) + p^2(w_C^\psi - s^\iota - s^\phi) + (1-p)^2(w_C^\nu) \\ &> \Pi_C^\nu = b + p(w_C^\psi - s^\iota) + (1-p)w_C^\nu\end{aligned} \tag{31}$$

which corresponds to

$$B > p(s^\phi + (1-p)(w_C^\nu - w_C^\psi)) \equiv B_C^* \tag{32}$$

and therefore the expected stigma (conditional on being detected) is given by:

$$\sigma_C^\phi = (1-p)(w_C^\nu - w_C^\psi) \tag{33}$$

Hence, the felony rate among type H individuals must satisfy:

$$B_C^* = p(s^\phi + \sigma_C^\phi) = G^{-1}(1 - \Theta) = B^s(\Theta) \tag{34}$$

Type L individuals, on the other hand, commit infractions if:

$$b > p(s^\iota + (w_C^\nu - w_C^\psi)) \equiv b_C^* \tag{35}$$

and therefore the equilibrium infraction rate among type L individuals is given by:

$$b_C^* = p(s^\iota + (w_C^\nu - w_C^\psi)) = F^{-1}(1 - \Theta) \equiv b^s(\theta) \tag{36}$$

Wages are determined by the average productivity of individuals who have records, and the average productivity of individuals with clean records. These are given by:

$$w_C^\psi = q_H + v \frac{\lambda p\theta}{\lambda p\theta + (1 - \lambda)p(1 + (1 - p)\Theta)} \tag{37}$$

and

$$w_C^\nu = q_H + v \frac{\lambda(1 - p\theta)}{\lambda(1 - p\theta) + (1 - \lambda)(1 - p)(1 - p\Theta)} \tag{38}$$

That there exists an equilibrium satisfying conditions (34)-(38) follows from a reasoning very similar to that in the proof of proposition 2, and therefore an existence proof for this case is omitted. The equilibrium crime rates Θ_C^* and θ_C^* and equilibrium magnitudes of stigma σ_C^ϕ and σ_C^ι can be compared to the corresponding values in the previous cases to discuss the effects of record confusion. Proposition 4, below, shows that criminalization of infractions dilute the stigma attached to felonies when there is record confusion, and that this, unsurprisingly, causes a reduction in the deterrence of felonies.

Proposition 4: *When there is record confusion (i) criminalizing infractions reduces the expected stigma associated with committing a felony, i.e. $\sigma_C^\phi < \sigma_L^\phi$, and (ii) criminalizing infractions increases the number of felonies committed $\Theta_L^* < \Theta_C^*$.*

Proof: (i)

$$\sigma_C^\phi = (1 - p)(w_C^\nu - w_C^\psi) = \tag{39}$$

$$(1-p)v \left(\frac{\lambda(1 - p\theta)}{\lambda(1 - p\theta) + (1 - \lambda)(1 - p)(1 - p\Theta)} - \frac{\lambda p\theta}{\lambda p\theta + (1 - \lambda)p(1 + (1 - p)\Theta)} \right) <$$

$$(1 - p)v \left(\frac{\lambda(1 - p\theta)}{\lambda(1 - p\theta) + (1 - \lambda)(1 - p)(1 - p\Theta)} \right) < v \frac{\lambda}{\lambda + (1 - \lambda)(1 - p\Theta)}$$

$$= w_L^\nu - w_L^\phi = \sigma_L^\phi$$

since, the second inequality is equivalent to

$$(1-p)(1-p\theta)\lambda + (1-p)(1-p\theta)(1-\lambda)(1-p\Theta) < \lambda(1-p\theta) + (1-\lambda)(1-p)(1-p\Theta) \tag{40}$$

12

which holds when $p > 0$.

(ii) Part (i) implies that $B_C^* = p(s^\phi + \sigma_C^\phi(\Theta, \theta)) < p(s^\phi + \sigma_L^\phi(\Theta)) = B_L^*$ for all $\theta \in (0,1)$, including θ_C^*, which implies that $\Theta_L^* < \Theta_C^*$. $\blacksquare$

Given that criminalization dilutes stigma even absent record confusion, the results presented in proposition 4 are unsurprising. In fact, as stated in the introduction, stigma dilution is even greater when there exists record confusion. The next proposition establishes this result.

Proposition 5: *(i) Stigma dilution is enhanced by record confusion, i.e. $\sigma_C^\phi(\Theta, \theta) < \sigma_I^\phi(\Theta, \theta)$, (ii) record confusion increases the stigma associated with having a misdemeanor record, i.e. $\sigma_I^\iota < \sigma_C^\psi$, and (ii) criminalizing infractions increases the number of felonies by more than the analogous increase when there is no record confusion.*

Proof: (i)

$$(1-p)(w_C^\nu - w_C^\psi) \tag{41}$$

$$= (1-p)v\left(\frac{\lambda(1-p\theta)}{\lambda(1-p\theta) + (1-\lambda)(1-p)(1-p\Theta)} - \frac{\lambda p\theta}{\lambda p\theta + (1-\lambda)p(1+(1-p)\Theta)}\right)$$

$$< v\left(p\frac{\lambda p\theta}{\lambda p\theta + (1-\lambda)p(1-p\Theta)} + (1-p)\frac{\lambda(1-p\theta)}{\lambda(1-p\theta) + (1-\lambda)(1-p)(1-p\Theta)}\right)$$

$$= pw_I^\iota + (1-p)w_I^\nu - w_I^\phi$$

which holds, since the condition is equivalent to

$$0 < p\frac{\lambda p\theta}{\lambda p\theta + (1-\lambda)p(1-p\Theta)} + (1-p)\frac{\lambda p\theta}{\lambda p\theta + (1-\lambda)p(1+(1-p)\Theta)} \tag{42}$$

(ii)

$$\frac{\sigma_I^\iota}{v} = \frac{\lambda(1-p\theta)}{\lambda(1-p\theta) + (1-\lambda)(1-p)(1-p\Theta)} - \frac{\lambda p\theta}{\lambda p\theta + (1-\lambda)p(1-p\Theta)} \tag{43}$$

$$< \frac{\lambda(1-p\theta)}{\lambda(1-p\theta) + (1-\lambda)(1-p)(1-p\Theta)} - \frac{\lambda p\theta}{\lambda p\theta + (1-\lambda)p(1+(1-p)\Theta)} = \frac{\sigma_C^\psi}{v}$$

which holds, since the first terms of $\frac{\sigma_I^\iota}{v}$ and $\frac{\sigma_C^\iota}{v}$ are equal, and the denominator of the second term of $\frac{\sigma_I^\iota}{v}$ is smaller than the denominator of the second term of $\frac{\sigma_C^\iota}{v}$.

(iii) Define $\Theta_{j \in \{I,C\}}^m(\theta)$ as

$$B_j^* = p(s^\phi + \sigma_j^\phi(\Theta_j^m, \theta)) = G^{-1}(1 - \Theta_j^m) \equiv B^s \tag{44}$$

Because the equilibrium felony rates are given by $\Theta_j^* = \Theta_j^m(\theta_j^*)$ for $j \in \{C, L\}$, it follows that $\Theta_C^* > \Theta_I^*$ if $\Theta_C^m(\theta_C) > \Theta_I^m(\theta_I)$ for all $(\theta_C, \theta_I) \in [0,1] \times [0,1]$. This condition holds, because (a) $\frac{d\Theta_C^m}{d\theta} > 0$, (b) $\frac{d\Theta_I^m}{d\theta} < 0$ and as implied by part

(i) $\Theta_C^m(\theta) > \Theta_I^m(\theta)$ for all θ. To see that statements (a) and (b) are true it is sufficient to make the following observations:

(a) $\frac{\partial \sigma_C^\phi}{\partial \theta} < 0$ since $\partial w_C^\nu / \partial \theta < 0$ and $\partial w_C^\psi / \partial \theta > 0$ (the last two inequalities follow immediately from expressions (37) and (38)). Hence, $\frac{\partial B_C^*(\Theta_C^m, \theta)}{\partial \theta} = p \frac{\partial \sigma_C^\phi}{\partial \theta} < 0$ and therefore $\frac{d\Theta_C^m}{d\theta} = -\frac{\frac{\partial B_C^*(\Theta_C^m, \theta)}{\partial \theta}}{\frac{\partial B_C^*}{\partial \Theta} - \frac{\partial B^s}{\partial \Theta}} > 0$; and

(b)

$$\frac{\partial \sigma_I^\phi}{\partial \theta} / v \tag{45}$$

$$= \frac{\lambda (1-\lambda) p^3 (1-\Theta p)}{(\theta \lambda p + p(1-\Theta p)(1-\lambda))^2} - \frac{\lambda (1-\lambda)(1-p)^2 p(1-\Theta p)}{(\lambda(1-\theta p) + (1-\Theta p)(1-\lambda)(1-p))^2} > 0$$

which holds iff

$$\frac{p^2}{(\theta \lambda p + p(1-\Theta p)(1-\lambda))^2} > \frac{(1-p)^2}{(\lambda(1-\theta p) + (1-\Theta p)(1-\lambda)(1-p))^2} \tag{46}$$

or equivalently

$$p(\lambda(1-\theta p) + (1-\Theta p)(1-\lambda)(1-p)) > (1-p)(\theta \lambda p + p(1-\Theta p)(1-\lambda)) \tag{47}$$

which holds iff $1 > \theta$. Hence, $\frac{\partial B_I^*(\Theta_I^m, \theta)}{\partial \theta} > 0$, and therefore $\frac{d\Theta_I^m}{d\theta} = -\frac{\frac{\partial B_C^*(\Theta_I^m, \theta)}{\partial \theta}}{\frac{\partial B_I^*}{\partial \Theta} - \frac{\partial B^s}{\partial \Theta}} < 0$. ∎

Proposition 5 demonstrates the exacerbating effect of record confusion on stigma dilution. Not only is the expected rewards to not having a criminal record is reduced, but the cost of having a felony record is also reduced; both the stick and the carrot become less effective. This is simply because some relatively high productivity individuals (i.e. type L) who would otherwise not have a criminal record are included in the pool of individuals with criminal records when the infraction is criminalized. This simultaneously increases the average productivity of individuals with criminal records and reduces the average productivity of individuals with no records. Thus, stigma dilution is exacerbated when there is record confusion.

5. Overcriminalization and Policy Implications
5.1. Overcriminalization and Marijuana Possession

Overcriminalization occurs if the criminalization of an act causes more social harms than benefits. The main benefit of criminalizing an act is presumably reducing the frequency with which that harmful act is committed. On the other hand, the main cost associated with criminalization identified in this article is the reduction in the deterrence of more serious crimes through stigma dilution. Next, I identify the conditions under which stigma dilution, and therefore the costs of criminalizing minor offenses, are likely to be high. Then, I identify conditions under which the benefits from punishing an act is low. Based on these conditions I discuss when over-criminalization is likely, and suggest that

14

the criminalization of marijuana possession meets almost all criteria identified for over-criminalization.

As stated in proposition 5, stigma dilution is maximal when there is record confusion. In this case, the amount of stigma dilution is given by $\sigma_L^\phi - \sigma_C^\phi$, and is increasing in v and λ, which confirms the intuition that stigma dilution is greatest when the act being criminalized is one that does not provide much information about how unproductive the individual is (i.e. v is large), and when the number of people interested in committing the act is large. Furthermore, as shown in the proof of proposition 5, the expected stigma cost associated with committing a felony, σ_C^ϕ, is decreasing in the rate of infractions θ. This implies that stigma dilution is greatest when many people commit infractions.

Insights from the standard theory of law enforcement also suggest that the cost of dilution is increasing in the harm inflicted through the felony (denote this as H) and the probability of detection, p. This is because the total harm caused by the felony equals $(1-\lambda)\Theta H$, and therefore the cost of reducing stigma is given by $(1-\lambda)G'(B_L^*)pH$. Although increases in p may affect the amount of stigma (Klement and Harel (2007)), a reduction in stigma is unambiguously more harmful to social welfare in a situation where p is high rather than low. If one also includes the benefit to criminals in the social welfare calculus, then the degree of under-deterrence, and the distribution of felons' benefits over crime also become relevant considerations. In these cases, the effect of stigma dilution is positively related to the degree of under-deterrence, i.e. the average difference between H and B_L^*.

The primary benefit of criminalization, on the other hand, is the reduction of the harms inflicted through the infraction, i.e. $\lambda(1-\theta)h$, where h is the social harm associated with the infraction. Hence, the benefit associated with criminalization is lowest when h is small and the number of people being deterred as a result of criminalization is small. The latter condition is most likely met when the expected cost associated with committing the infraction is low, which, in turn, occurs when the sanction (s^ι) for the infraction is low.

In sum, over-criminalization is most likely when (i) record confusion is likely, (ii) the act being criminalized does not provide much information about the person's lack of productivity, (iii) many people commit the infraction even when it is criminalized, (iv) the harm associated with the infraction relative to the felony (i.e. h/H) is small, and (v) the probability of detection for the infraction relative to the probability of detection for the felony (i.e. p^ι/p) is low.[11]

Applying these conditions to marijuana possession suggests that the criminalization of these acts are very strong candidates for producing more social costs than benefits, i.e. over-criminalization.

---

[11] This condition is subject to λp^ι not being small. Otherwise, the effect of stigma dilution is small, because not too many type L individuals receive criminal records, and therefore stigma dilution is limited. This condition most likely holds in the context of marijuana possession. Nguyen and Reuter (2012) report, for instance, that "there were an estimated 14,000,000 arrests for all offenses in 2008; of that number, 12.2% were drug abuse violations.... Marijuana possession alone accounted for almost half (45%) of all drug law violations (Federal Bureau of Investigation, 2009)."

Factor (iii) is the easiest condition to obtain information on. According to a 2012 survey reviewed by the National Institute on Drug Abuse "more than 111 million Americans over the age of 12 had tried marijuana at least once, and nearly 19 million had used the drug in the month before the survey."[12] Needless to say, the fact that this information comes from a survey causes concern. Numbers may be under-estimates if people did not want to provide information on their criminal activities. More importantly, some of this information presumably comes from states where it is no crime to consume marijuana. Nevertheless, the numbers are quite high and indicate that people use marijuana, whether or not it is illegal to do so. This may be because s^i is relatively low.

A good proxy for factor (v) is the arrest rate for various offenses. According to the FBI's Uniform Crime Reports for 2013 the clearance rate for murder and non-negligent manslaughter was around 64 percent. It is hard to find similar information for marijuana possession, because a much lower percentage of marijuana possessions are reported to law enforcers. Fortunately, some information on arrest rates for marijuana users is provided in Nguyen and Reuter (2012). The authors estimate the national probability of arrest, conditional on use within the last twelve months, for marijuana possession between 1982 and 2008. They find numbers that vary between 0.7% and 1.9% where the peak arrest rate was observed in 2000. These numbers appear sufficient for purposes of demonstrating the enormous gap between the two probabilities of conviction. One may claim that the clearance rate for murder is exceptionally high compared to other crimes, and therefore one ought to focus on other crimes. Doing this does not appear to change much of these conclusions, because the reported clearance rate for all property crimes is around 20%, which is still about 10 times higher than the highest rate reported by Nguyen and Reuter (2012).

Making statements about (iv) is relatively difficult, due to lack of data, but also due to the inherent necessity of making value judgments. Marijuana possession is (when declared as such) a crime that is quite different from crimes such as murder and rape where there is at least one very easily identifiable direct victim. Therefore, people who strongly oppose the legalization of marijuana point to social values and the protection of the youth against addictive substances as the primary goal of criminalization. When harms are broadly defined, it is difficult to ascertain their magnitude. Nevertheless, it is doubtful that one would rank the harms associated with an incident of marijuana possession as being higher than the harms associated with an incident of, for instance, breaking and entering.

Similarly, it is hard to assess whether consumption of marijuana reduces productivity, or whether less productive individuals opt in (on average) to becoming marijuana consumers. Given the number of studies on the issue, discussion of this topic is one that is best left to experts in the field. However, regardless of whether or not productivity is reduced, the fact that most Americans have at some point in their lives used marijuana suggests that the signal value of possession is very low. Hence, condition (ii) is also probably met.

---

[12] U.S. Department of Health and Human Services (2014).

Finally, whether or not record confusion is likely depends on a number of factors, including how the crime of possession is defined (e.g. drug possession versus marijuana possession), how it is reported, and whether employers devote the time necessary to weed out through various criminal records. However, I am currently unable to see any reason why this crime would cause less record confusion compared to any other minor infraction whose criminalization may cause stigma dilution. Accordingly, factor (i) does not appear to be one that would potentially weaken the candidacy of criminalization of marijuana possession for over-criminalization.

Thus, a review of the five factors identified in the context of marijuana possession suggests that, if there is any act whose criminalization is likely to cause costs associated with stigma dilution that can off-set benefits obtainable from deterring the commission of that act, marijuana possession is probably it.

5.2. Policy Implications

Stigma dilution is greater when the act being criminalized provides little information regarding the offender's productivity. There are at least three ways to mitigate potential stigma dilution. The first one is quite obvious: not criminalizing (or decriminalizing) acts that do not, on average, provide much information regarding a person's productivity. The second is the regulation of such acts through other modes of regulation, such as administrative law. The third is the use of expungements in combination with criminal sanctions imposed for these acts. The first two of these remedies are self-explanatory. The third one, which I have introduced elsewhere,[13] requires the sealing of a criminal record as soon as the punishment is imposed, i.e. the non-publicizing of the person's criminal record.

Either one of these three methods will reduce the stigma dilution caused by the regulation of the targeted act. Hence, given a list of acts whose criminalization is likely to significantly dilute stigma, there are ways to mitigate the harm caused through regulation. Another question remains: what are good methods to identify which acts ought not to be subject to traditional criminalization?

As most criminal law scholars agree, a strong mens rea element must be a prerequisite to defining an act as a crime. Being more confident regarding the offender's mental state simply increases the information that one obtains regarding the offender's intentions and his likely behavior in a working environment. Furthermore, although this is not the primary focus of this article, a conviction for a crime that requires a strong mens rea element also provides information regarding the criminal's attitude in his social circles.

Related to this point, attaching a strong scienter to all non-expungable criminal offenses is likely to reduce stigma dilution. The Model Penal Code, for instance, defines four scienters: purposely, knowingly, recklessly, and negligently. Recklessness requires a gross deviation from social norms. If deviation is correlated with the productivity and/or reliability of the individual (conditional on the act itself being one that is generally disapproved of), a crime which requires reckless conduct may produce a good signal. However, strict liability offenses

---

[13]See Mungan (2014b).

and offenses that only require negligence, presumably produce less reliable information regarding the individual. One way to mitigate harms due to stigma dilution without reducing much the deterrence of the act being regulated is to define multiple crimes for the same act that require different scienters, and allow an automatic expungement of records, as in Mungan (2014b), if the defendant's conduct cannot be shown to be at least reckless.

Two contexts in which conviction presumably produces biased information regarding the offender's mental state are conspiracy liabilities and the felony murder rule. In both cases, the law imputes the mens rea of a more serious act to people who have completed the actus reus for a lesser harm act. Consider, for instance, two text book examples: the driver in an armed robbery conspiracy and the burglar who accidentally kills a person while retreating. There are many plausible scenarios where neither of these individuals have the stomach to actually kill a person, but their criminal records may indicate otherwise. Stigma dilution is therefore a likely by-product of these doctrines. In fact, there exists some empirical evidence that can be interpreted as providing some support for the existence of stigma dilution in the context of the felony murder rule. One would imagine that, absent stigma dilution or other frictions, imposing the felony murder rule would have a significant deterrent effect on felonies. But, Malani (2002) finds that "the felony-murder rule has a relatively small effect on criminal behavior". There may be many reasons for this finding, but at the very least, the findings are consistent with the theory of stigma dilution.

6. Assumptions and Extensions

6.1. Reputational Sanctions

The similarities between reputational sanctions imposed on producers due to past wrongdoings and stigma in the criminal context has been identified in the literature. (Iacabucci (2014), Mungan (2014a)) Because civil and administrative sanctions imposed on producers provide signals to potential consumers regarding the quality of products, one may suspect that reputational sanctions, just like stigma, may be diluted if producers are subjected to sanctions for things that do not correlate much with the quality of their products.

Although it is hard to tell absent further research whether the analogy between reputational sanctions and stigma is strong enough vis-à-vis dilution, there are at least three meaningful dimensions on which reputational sanctions and stigma can be compared.

One dimension concerns the availability of information on past wrong-doings. Average consumers of a daily consumption good may not know much about the producer's compliance with minor regulations. However, they may hear about non-compliance if it causes harms that impact many lives. In other words, the relevance of the information and its accessibility appear, at first sight, to be positively related. This stands in contrast to criminal records, which are generally equally easily available for serious crimes as well as lesser offenses. Hence, dilution appears to be less of a concern in the context of reputational sanctions.

A second important distinction, as pointed out by Galbiati and Garoupa (2007), stems from the different procedures used in administrative law and

criminal law. Because the standard proof in criminal law is much higher, the information produced through a conviction is more reliable than a sanction imposed through administrative law. Accordingly, reputational sanctions are diluted to begin with, and, therefore, are unlikely produce as large stigma as in the criminal context. This observation further undermines potential concerns associated with the dilution effect of regulating firms for acts that do not provide information about there quality.

One distinction that may cut in the opposite direction is the potentially greater 'record confusion' in the context of regulation violations. It may be very difficult, or perhaps impossible, for average consumers to tell the difference between a technical regulation that is related to the quality of a product and another technical regulation that does not.

None of these brief comments should be interpreted as suggesting that it is harmless to subject firms to random regulations. Chilling competitive and productive efforts by firms is a cost whose importance is well recognized in the literature. The only purpose is to demonstrate that costs associated with reputational sanction dilution is likely to be much smaller than in the context of criminal stigma.

6.2. Dependency of p to Legal Regime

Different arguments can be made regarding the effect of criminalizing an act on the probability of convicting people who commit other crimes, i.e. p. First, one may argue that criminalization of a new act may result in the diversion of fixed (or not perfectly elastic) law enforcement resources towards the enforcement of the newly criminalized act. This would presumably reduce the probability of conviction for other acts. Adda et al. (2014), for instance, relies on this dynamic in providing an explanation for the drop in criminal activity post depenalization of cannabis possession in the borough of Lambeth in London. This argument, absent other considerations appears valid, and its implication is simple: criminalization of new acts reduce the deterrence of previously criminalized acts not only through stigma dilution, but also *enforcement dilution.*

Others may argue that a countervailing effect may actually cause reductions in p. The investigation of infractors can give law enforcers opportunities (or excuses) to investigate other crimes they suspect were committed (or are being committed) by those individuals. Visualizing this effect does not require much imagination. There are many cases in which police officers initiate an investigation because they suspect that an individual possesses illegal drugs, e.g. marijuana, but the investigation ends with the arresting of that individual for a more serious crime. Therefore, if criminalizing minor offenses eases the initiation of investigations, it can actually increase the conviction rate for more serious crimes.[14]

The second argument may or may not be empirically sound. But, any rea-

---

[14]Another, similar, argument is that incarcerating drug offenders can reduce the crime rates of other, more serious, offenses, by incapacitating more people with high tendencies. This effect is what Kuziemko and Levitt (2004) rely on in explaining the small negative relationship that they find between incarceration of drug offenders and the commission of other crimes.

soning that relies on this argument to conclude that decriminalization of minor offenses, e.g. marijuana possession, is normatively undesirable, because it will reduce the crime rate of other offenses, is at the very least overlooking a very important point. First, even if one desires to tie the initiation of criminal investigations to marijuana possession (or some other act), one need not criminalize this act. Instead, one can simply attach a loss of rights, in the form of being subject to searches by law enforcers, to marijuana possession. In other words, one can let marijuana possession be sufficient to trigger investigation, but not sufficient (on its own) for conviction. This would achieve the same increase in p without resorting to criminalization. Some may argue that this is probably impossible in the United States due to Fourth Amendment considerations. But, regulatory authority is very frequently used to empower the police to investigate further, e.g. "if you drive, you are subject to random stops and searches". In fact, as demonstrated in Logan (2015) in great detail, even decriminalization is insufficient to stop arrests based on marijuana possession. Moreover, when one considers the way *Terry stops* are conducted, one understands quickly that law enforcers do not need much of a reason to stop and frisk individuals who they find suspicious. Some raw numbers may be convincing on this last point: only about 6% of the 4.4 million Terry stops conducted by the New York Police Department between 2004 and 2012 resulted in arrests.[15]

The objective of this brief discussion is certainly not to legitimize the frequent stops and searches unrelated to the targeted crime. In fact, such searches can generate a long list of problems, including, public choice problems; bribery; false arrests; and easing discrimination; and profiling. Rather, the objective here is to illustrate that decriminalization does not necessarily prevent these stops and searches. Hence, even if one sees benefits associated with these stops and searches, these benefits (on their own) do not raise concerns regarding decriminalization.

6.3. When Some Type H Individuals Can be Deterred From Committing Infractions (Case D)

Throughout the main analysis I assumed that all type H individuals commit the crime, because their minimum benefit from these infractions are sufficiently large. Here, I discuss the effects of relaxing this assumption. To formalize these effects, let β represent type H individuals' benefits from committing an infraction.

In this case, type H individuals have four, rather than two choices. They can commit both acts, only the infraction, only the felony, or neither crime. When infractions are criminalized (and there is no record confusion), the expected pay-offs associated with these options are given, in the same order, by:

$$\Pi_D^2 = (B + \beta) + p(1-p)(w_D^\phi + w_D^\iota - s^\phi - s^\iota) + p^2(w_D^\phi - s^\phi) + (1-p)^2 w_D^\nu \quad (48)$$

$$\Pi_D^\iota = \beta + p(w_D^\iota - s^\iota) + (1-p)w_D^\nu \quad (49)$$

$$\Pi_D^\phi = B + p(w_D^\phi - s^\phi) + (1-p)w_D^\nu; \text{ and} \quad (50)$$

---

[15] Floyd v. City of New York, 959 F.Supp.2d 540, 558-559 (2013).

$$\Pi_D^0 = w_D^\nu \qquad (51)$$

A comparison of these pay-offs reveals that:

$$\Pi_D^\phi > \Pi_D^0 \text{ if } B > p(w_D^\nu - w_D^\phi + s^\phi) \equiv B_D^\phi \qquad (52)$$

$$\Pi_D^\iota > \Pi_D^0 \text{ if } \beta > p(w_D^\nu - w_D^\iota + s^\iota) \equiv \beta_D^\iota \qquad (53)$$

$$\Pi_D^2 > \Pi_D^0 \text{ if } B > p(s^\phi - w_D^\nu - w_D^\phi) + p(1-p)(s^\iota - w_D^\nu - w_D^\iota) + \beta \equiv B_D^2(\beta) \quad (54)$$

$$\Pi_D^2 > \Pi_D^\iota \text{ if } B > p(s^\phi + pw_D^\iota + (1-p)w_D^\nu - w_D^\phi) \equiv B_D^\iota; \text{ and} \qquad (55)$$

$$\Pi_D^2 > \Pi_D^\phi \text{ if } \beta > p(s^\iota + pw_D^\iota + (1-p)w_D^\nu - w_D^\iota) \equiv \beta_D^\iota \qquad (56)$$

Hence, any individual with $B > B_D^\phi$ and $\beta > \beta_D^\phi$ commits both crimes since for these individuals $\Pi_D^2 > \Pi_D^\phi > \Pi_D^0$. Similarly, individuals with $B > B_D^\iota$ and $\beta > \beta_D^\iota$ commit both crimes, because $\Pi_D^2 > \Pi_D^\iota > \Pi_D^0$ for these individuals. Moreover, individuals with $B \in (B_D^\iota, B_D^\phi)$ and $\beta \in (\beta_D^\phi, B_D^\iota)$ commit both crimes iff $B > B_D^2(\beta)$, since for these individuals $\Pi_D^2 > \Pi_D^0 > \max\{\Pi_D^\phi, \Pi_D^\phi\}$, where the last inequality follows from (48) and (49). The remaining individuals' pay-offs are maximized either by committing neither act or just one of them, since $B < B_D^2(\beta)$. Hence, those with $B > B_D^\phi$ but $\beta < \beta_D^\phi$ commit only felonies, and those with $B < B_D^\iota$ but $\beta > \beta_D^\iota$ commit only infractions. The last sub-set of these individuals, i.e. those with $B < B_D^\phi$, $\beta < \beta_D^\iota$ and $B < B_D^2(\beta)$, commit neither act. These observations are summarized by figure 1, below.

[Insert Figure 1, here]

To illustrate how this alternative assumption affects the stigma dilution effect, consider how figure 1 would be affected if all type H individuals were to commit infractions. As reflected by expression (20) all individuals with $B > B_I^\phi = B_D^\phi$ would commit both crimes, and all individuals with $B < B_I^\phi = B_D^\phi$ refrain from committing a felony, and therefore only commit the infraction. In other words, the distinction is that now people in area I commit 0 offenses rather than 1 offense; people in area II commit 0 offenses rather than 2; and people in area III commit 1 offense rather than 2. The behavior of all other individuals are invariant to the assumption employed. This implies that this alternative assumption increases the number of type H individuals with clean records, which, in turn, reduces the average productivity of people with clean records further. Hence, the value of holding a clean record, i.e. the *carrot*, is devaluated more relative to the case where all type H individuals commit infractions. Thus, absent other considerations, the stigma dilution effect is likely to be magnified by considerations related to type H individual being deterred from committing infractions.

7. Conclusion

Punishing small harm offenses can have the unintended consequence of reducing the stigma attached to other, more serious, crimes. Given that the elimination of harms associated with more serious crimes should be the priority of

the criminal justice system, it would be wise to reserve criminal punishment to those acts that send strong signals about a person's mental state and character. Otherwise, signals provided through the criminal justice system become noisy and its effectiveness in deterring individuals can be reduced, and the resulting case can be described as over-criminalization.

A question that is often asked today is whether the penalization of marijuana related offenses constitutes over-criminalization. Given the current movement among several states towards partial or complete legalization of marijuana possession and consumption, it is important to discuss and document the potential gains and losses associated with the policy choices we soon have to make. This article provides a theoretical framework to discuss the normative desirability of criminalizing various acts. It concludes that if there are any acts whose punishment are likely to cause stigma dilution, possession and consumption of marijuana should be on the top of that list.

References

[1] Adda, J., B. McConnell, and I. Rasul (2014) "Crime and the Depenalization of Cannabis Possession: Evidence from a Policing Experiment" Journal of Political Economy 122, 1130-1202.

[2] Becker, Gary S. "Crime and Punishment: An Economic Approach." Journal of Political Economy 76 (1968): 169-217.

[3] Benson, B., Rasmussen, D., (1991) "Relationship between Illicit Drug Enforcement Policy and Property Crimes" Contemporary Policy Issues 9: 106–115.

[4] Bowles, R., M. Faure, N. Garoupa (2008) "The Scope of Criminal Law and Criminal Sanctions: An Economic View and Policy Implications" Journal of Law and Society 35, 389-416.

[5] Corman, H. and N. Mocan (2000) "Time-Series Analysis of Crime, Deterrence, and Drug Abuse in New York City" American Economic Review 90, 584-604.

[6] Funk, P. (2004) "On the effective use of stigma as a crime-deterrent" European Economic Review 48, 715-728.

[7] Galbiati, R. and N. Garoupa (2007) "Keeping Stigma Out of Administrative Law: An Explanation of Consistent Beliefs" Supreme Court Economic Review 15, 273-283.

[8] Harel, A, and A. Klement (2007) "The Economics of Stigma: Why More Detection of Crime May Result in Less Stigmatization", Journal of Legal Studies 36, 355-377.

[9] Iacobucci, E. (2014) "On the Interaction between Legal and Reputational Sanctions" Journal of Legal Studies 43, 189-207.

[10] Kim, I. "Overcriminalization in Korea" in Regulatory Crimes and Overcriminalization 37-96 (I. Kim and J. K. Hyun, eds 2013)

[11] Kuziemko, I. and S. Levitt (2004) "An Empirical Analysis of Imprisoning Drug Offenders" Journal of Punlic Economics 88, 2042-2066.

[12] Logan, W. (2015) "After the Cheering Stopped: Decriminalization and Legalism's Limits" Cornell Journal of Law and Public Policy, 24, Forthcoming.

[13] Lott, J. (1992a) "An Attempt at Measuring the Total Monetary Penalty from Drug Convictions: The Importance of an Individual's Reputation", 21 Journal of Legal Studies 159.

[14] Lott, J. (1992b) "Do We Punish High Income Criminals Too Heavily?" 30 Economic Inquiry 583.

[15] Malani, A. "Does the Felony-Murder Rule Deter? Evidence from FBI Crime Data", working paper.

[16] Mungan, M. "The Scope of Criminal Law" in The Research Handbook on the Economics of Criminal Law 51-66 (A. Harel and K. Hylton, eds. 2012)

[17] Mungan, M. (2014a) "A Generalized Model of Reputational Sanctions and the (Ir)relevance of the Interactions between Legal and Reputational Sanctions" mimeo.

[18] Mungan, M. (2014b) "On the Sealing of Criminal Records, Stigma, and Deterrence" FSU College of Law, Law, Business & Economics Paper No. 14-13.

[19] U.S. Department of Health and Human Services (2014) NIH Publication No. 14-4036

[20] Nguyen, H. and Peter Reuter (2012) "How Risky Is Marijuana Possession? Considering the Role of Age, Race and Gender" Crime & Delinquency 58, 879-910.

[21] Polinsky, A. Mitchell and Steven Shavell "Public Enforcement of Law" in Handbook of Law and Economics 403-454 (A. M. Polinsky & S. Shavell, eds. 2007)

[22] Prescott, J. and J. Rockoff (2011) "Do Sex Offender Registration and Notification Laws Affect Criminal Behavior?" Journal of Law and Economics 54, 161-206.

[23] Rasmusen, E. (1996) "Stigma and Self-Fulfilling Expectations of Criminality" Journal of Law and Economics 39, 519-543.

[24] Shavell, Steven "The optimal structure of law enforcement" Journal of Law and Economics 36 (1993): 255–287.