

**Preprints of the
Max Planck Institute for
Research on Collective Goods
Bonn 2014/17**



**Legal Uncertainty as a
Welfare Enhancing Screen**

Matthias Lang



MAX PLANCK SOCIETY



Legal Uncertainty as a Welfare Enhancing Screen

Matthias Lang

October 2016, initial version February 2010

Legal Uncertainty as a Welfare Enhancing Screen

Matthias Lang*

October 2016

Abstract

Consider legal uncertainty as uncertainty about the legality of a specific action. In particular, suppose that the threshold of legality is uncertain. I show that this legal uncertainty raises welfare. Legal uncertainty changes deterrence in opposite directions. The probability of conviction increases for firms below the threshold, while the probability of conviction decreases for firms above the threshold. Hence, legal uncertainty acts as a welfare enhancing screen and increases welfare. Legal uncertainty discourages some actions with low private benefits, while it encourages other actions with high private benefits.

JEL classifications: D8, K2, K4, L5

Keywords: Enforcement, Deterrence, Legal Uncertainty, Regulation, Asymmetric Information, Incomplete Contracts

1 Introduction

Legal uncertainty is prevalent given the complexity of many legal procedures. With legal uncertainty, I refer here to situations in which it is unclear *ex ante* whether a specific action is legal.¹ Hence, when taking an action, firms do not know with certainty whether courts or enforcement authorities judge this action to be legal. For example, assessments of efficiency defenses differ or it is uncertain which evidence will be allowed. Alternatively, enforcement authorities make measurement errors or there are different, possibly contradicting procedures applying to a case. Previous literature has shown that legal uncertainty might deter the wrong actions – over-detering socially beneficial actions, while under-detering socially detrimental ones.² Thus, legal uncertainty decreases welfare and should be avoided whenever possible.

My main result shows that legal uncertainty can increase welfare – contradicting conventional wisdom. Legal uncertainty allows mitigating restrictions of enforcement author-

I thank Christoph Engel, Tim Friehe, Thomas Gall, Georg von Graevenitz, Dominik Grafenhofer, Martin Hellwig, Jos Jansen, Johannes Koenen, Giorgio Monti, Alexander Morell, Urs Schweizer, Christian Traxler, David Ulph, Achim Wambach, two anonymous referees, and an associate editor for very helpful suggestions and discussions, and the audiences at the University of Copenhagen, Cologne, Berlin (BERA), Bonn, ETH Zurich, EARIE 2016, CLEEN Workshop 2016 and 2011, CESifo Applied Micro 2015, VfS 2014, PET 2013, EEA 2012, CRESSE 2012, and ETH Zurich/IMPRS-CI Law & Economics Workshop 2010 for comments. Financial support from the Max Planck Institute for Research on Collective Goods and the German Research Foundation (DFG) through the Bonn Graduate School of Economics (BGSE), SFB/TR 15, and GRK 629 is gratefully acknowledged.

*Free University Berlin, Economics, Boltzmannstr. 20, 14195 Berlin, Germany, lang@uni-bonn.de.

¹This is similar to the notion of D'Amato (1983).

²See, e.g., Craswell and Calfee (1986), Polinsky and Shavell (1989) or Schinkel and Tuinstra (2006).

ities, in particular, ignorance of firms' private information. Enforcement authorities can use legal uncertainty as a welfare enhancing screen. Consequently, legal uncertainty could make a norm more selective and increase social welfare.

This reasoning applies to many settings. As it is impossible to list all of these settings, I provide just some examples here. Think about pollution thresholds in environmental law or choosing the right transfer pricing in tax law. In privacy law, there are uncertain privacy thresholds for social networks and other internet businesses. Copyright and patent law also contain uncertain thresholds of originality for works to be eligible for protection. It also extends to excessive risk loading by financial institutions or accounting violations. In antitrust, price reductions might reflect lower costs or an attempt at predatory pricing. Finally, deals of patent-holders with generic drug makers to avoid "Paragraph IV" challenges, bidding patterns in procurement contests or standardization efforts might have beneficial effects or be part of some collusive agreement to harm other market participants.

In all these settings, externalities and private benefits of a specific action vary depending on circumstances. Enforcement authorities cannot perfectly distinguish between these circumstances and observe only a noisy signal about circumstances. Enforcement authorities choose optimal policies by setting a threshold of legality. Then firms decide whether or not to pursue the controversial action. Firms know circumstances and private benefits, but not the enforcement authorities' signal. Finally, enforcement authorities impose penalties on firms that are above the policy threshold according to its signal.

For the intuition, suppose there were legal certainty and no noise in the enforcement authorities' signals about circumstances. Then all firms below the policy threshold were active – no matter how small their private benefits. All firms above the policy threshold were deterred for sufficiently high penalties – no matter how large their private benefits. With legal uncertainty, some firms cannot fully anticipate whether they are above or below the policy threshold according to the enforcement authority's signal. In particular, if a firm is close to, but below the policy threshold, legal uncertainty implies that the firm is penalized with some probability. Therefore firms with low private benefits do not take the action. Legal uncertainty deters them. If a firm is close to, but above the policy threshold, legal uncertainty implies that the firm is not penalized with some probability. Hence, firms with high private benefits take the action. Legal uncertainty encourages them to take the action. Therefore, legal uncertainty increases deterrence for firms below the threshold and decreases deterrence for firms above the threshold. An optimal policy threshold implies that both effects of legal uncertainty on deterrence increase welfare. Consequently, this kind of legal uncertainty acts as a welfare enhancing screen.

Consider a real-world example that will guide us through this paper. Vertical restraints, like exclusive dealings, are prohibited in the European Union under Article 101 (TFEU), formerly Article 81 (EC).³ Due to a Block Exemption Regulation, however, this rule does not apply if the market shares of the involved firms are below 30%. Thus, circumstances

³See European Commission (2010) and Regulation No. 330/2010 for details.

could equal market shares.⁴ Then, the policy threshold is 30% market shares. The European Commission can impose penalties on firms in violation of this article. Although the European Commission gives guidelines how market shares are determined, it is extremely difficult to predict correctly the market share determined by competition authorities. The causes are discrepancies in the definition of the relevant market, information asymmetries or imprecision in the measurement of sales, and other factors. Therefore it is plausible to assume that firms know their market shares and private benefits, but not the Commission's signal, i.e., its estimate of their market shares. This creates the kind of uncertainty analyzed in the model. Thus, a firm with market shares of 25% anticipates with some probability an estimate above 30% and to pay a penalty. On the contrary, a firm with market shares of 35% anticipates with some probability an estimate below 30% and not to pay a penalty. According to my model, this legal uncertainty could be socially beneficial.

The remainder of the paper is organized as follows. Section 2 discusses the relevant literature. Section 3 sets up the model. Section 4 analyzes the welfare effects of legal uncertainty. Section 5 provides a numerical example to demonstrate the significance of legal uncertainty. Finally, Section 6 discusses possible limitations and concludes. All proofs are gathered in the appendix.

2 Related Literature

The literature has considered both costs and benefits of legal uncertainty. Legal uncertainty reduces deterrence and makes it more difficult or impossible to achieve optimal deterrence. For example, Polinsky and Shavell (1989) demonstrate that legal uncertainty lowers deterrence, because expected sanctions are reduced and less suits are brought to court. Schinkel and Tuinstra (2006) analyze firms' strategic responses to legal uncertainty in competition law. They show that both type I and type II errors lower deterrence. Png (1986) and Polinsky and Shavell (2000, Section 8) show how to adjust sanctions for type I and type II errors. Lando (2006) discusses whether type I errors influence deterrence.

These five papers assume that the same errors apply to everyone ruling out any effects on differential deterrence. Differential deterrence means that more harmful actions are more likely to be deterred than less harmful actions. In addition, the conclusions of these five papers are valid in the absence of enforcement costs. On the contrary, Besanko and Spulber (1989) determine optimal sanctions and auditing probabilities that both depend on quantities produced by firms. Enforcement authorities cannot observe firms' costs. Then it can be optimal to tolerate collusion by low-costs firms if auditing costs are sufficiently large. Hence, a (deterministic) type II error can be optimal, as auditing costs make lower deterrence optimal given an upper bound on sanctions. Between these two extremes of constant errors as in the first five papers and fully variable errors as in Besanko and Spulber (1989), Calfee and Craswell (1984) and Craswell and Calfee (1986) consider a

⁴An alternative source of legal uncertainty for vertical restraints are assessments of efficiency defenses. Then circumstances would measure the amount of efficiencies of an action.

given distribution of errors around the policy threshold. This legal uncertainty causes type I and type II errors independent of enforcement costs. In addition, there are effects on differential deterrence. Calfee and Craswell assume that firms' private benefits are constant for each action and that policy thresholds are exogenous. Therefore, legal uncertainty causes too much or too little deterrence and reduces welfare. I follow their approach to legal uncertainty adding an endogenous policy threshold, a distribution of private benefits, and information asymmetries. I show that legal uncertainty is beneficial, because it acts as a welfare enhancing screen. This screening effect is new to the literature and does not depend on auditing costs.

In contrast to my general setting, beneficial effects of legal uncertainty have appeared in several specific contexts. Choné and Linnemer (2008) study the effects of uncertain efficiency gains on merger control. They characterize market structures and demand elasticities that make such uncertainty beneficial. Strausz (2011) examines market structures that make regulatory risk advantageous. Lang and Wambach (2013) consider insurance fraud. They show that uncertainty about enforcement might have beneficial deterrence effects. The uncertainty concerns auditing probabilities which are the same for everyone. Therefore there is no screening and no effects on differential deterrence. In particular, this uncertainty has no effects on ambiguity-neutral firms. Osofsky (2011) and Gergen (2011) discuss different aspects of legal uncertainty in tax enforcement.

Differential deterrence can result from legal uncertainty, but also from evasion efforts or information acquisition. In Malik (1990), probabilities of conviction depend on firms' effort to avoid being caught. He shows that there is differential deterrence: firms with low private benefits are deterred, while firms with high private benefits take the action and invest into avoidance. Alternatively, in a series of papers, Kaplow assumes one-sided or mutual ignorance about externalities, because firms do not know the exact rules and enforcement authorities do not know the firm's specific circumstances. Therefore firms and enforcement authorities decide whether to acquire information. Kaplow (1990, 1994) analyses incentives to acquire information about externalities before and after taking the action. Kaplow (1990) also studies when it is optimal to set different sanctions for informed and uninformed firms. Kaplow (1995) models the trade-off between compliance costs and selectivity of rules. He shows that compliance costs are often low, even for complex rules. Finally, Kaplow and Shavell (1992) study the care exercised under different liability rules in two cases. In the first case, firms can learn the externalities they create. In the second case, firms know their externalities, but they can also learn the noisy signal of enforcement authorities about their externalities. The precision of this signal is constant. Kaplow and Shavell (1992) show that private incentives to acquire information are often excessive.

I conclude this section by discussing three concepts that are closely related to legal uncertainty: random contracts, rules vs. standards and per-se rules vs. rules of reason. Begin with random contracts. Legal uncertainty can be interpreted as a particular form of a random contract. The literature on random contracts shows that randomization is

sometimes optimal to screen risk-averse agents. Seminal papers include Gauthier and Laroque (2014), Strausz (2006), and Arnott and Stiglitz (1988). They characterize when randomization is optimal depending on the curvature of utilities. Randomization creates welfare losses by increasing risk for risk-averse agents, but welfare gains from mitigating incentive-compatibility constraints. In these settings randomization is never desirable for risk-neutral agents.

Legal uncertainty is also an important issue in the comparison of rules and standards. Rules are ex-ante specified conditions, while standards are conditions to be specified ex-post. For example, a rule can prohibit driving above 50 miles per hour. A standard, instead, would leave it to the adjudicator to determine whether the speed was excessive under the circumstances. Ehrlich and Posner (1974) compare rules and standards. They conclude that standards better distinguish beneficial from harmful actions, but provide less guidance for the concerned parties. Kaplow (1992) sees the main difference between rules and standards whether the law is given content ex ante or ex post. Hence, rules have higher initial specification costs, but lower enforcement and compliance costs than standards. In particular, the amount of legal uncertainty need not be higher for standards than for rules. Finally, Friedman and Wickelgren (2014) argue that standards allow firms to better signal their stakes in a case than rules. Therefore, standards can be more efficient than rules if litigations costs are sufficiently small.

My paper also relates to the comparison of per-se rules and rules of reason. With per-se rules, some clearly specified actions, like, e.g., certain rebates or resale price maintenance, are prohibited. A rule of reason, on the other hand, judges an action as illegal whenever the action is used in an anticompetitive way. Thus, the test of legality is whether competition was promoted or hindered.⁵ Therefore an action may be legal in some cases, but not in others, depending on its consequences. Hence, rules of reason typically imply a certain amount of legal uncertainty. The ‘more economic approach’ in the European Union also focuses more on market effects of the action. Previous literature has argued that rules of reason allow differentiating competition law in a more selective way at the price of some inherent legal uncertainty, because firms sometimes do not know whether their conduct is legal. Katsoulacos and Ulph (2009) characterize conditions, such that rules of reason are welfare-enhancing compared to per-se rules. They find that the selectivity of a rule of reason often outweighs the losses due to legal uncertainty. Katsoulacos and Ulph (2016a,b) continue this comparison. Focusing on competition law, they assume a consumer welfare standard. They distinguish two kinds of legal uncertainty, namely firms not knowing externalities or firms not knowing the competition authority’s signals about externalities. This modeling follows the two cases in Kaplow and Shavell (1992). Legal uncertainty of the first kind decreases welfare. Legal uncertainty of the second kind decreases or increases welfare. Katsoulacos and Ulph (2016a,b) consider a different concept of legal uncertainty and different welfare standards compared to my paper.

⁵Kaplow and Shapiro (2007, p. 54ff) provide a good discussion of rules of reason in antitrust.

3 Enforcement Model

An enforcement authority faces a risk-neutral firm. The authority's objective is to maximize welfare, i.e., externalities and firm's private benefits, weighting private benefits by $\alpha \in [0, 1]$.⁶ The enforcement authority sets its policy by choosing the threshold of legality $\hat{\sigma} \in \mathbb{R}$. In the example of vertical restraints, this policy is a market share of 30% for the block exemption. Every firm has the binary choice whether to take an action, like, e.g., bundling, rebates, or selective distribution systems, or to abstain from it. Depending on its choice, I refer to a firm as active or deterred.

The firm's type consists of circumstances σ , negative externalities x , and private benefits b created by the action. These three components are jointly distributed and known to the firm. Circumstances σ reflect size, scale, magnitude or severity of the action. Circumstances capture aspects of the action that are potentially observable to the enforcement authority. In the example of vertical restraints, circumstances could equal market shares. Denote the circumstances' marginal distribution by F on \mathbb{R} with a twice differentiable and positive density $f(\sigma)$. Conditionally on circumstances, externalities and private benefits are independently distributed. The conditional expectations of externalities $\mathbb{E}(x|\sigma)$ and private benefits $\mathbb{E}(b|\sigma)$ are twice differentiable. In particular, negative externalities and private benefits are on average higher for higher circumstances, i.e.,

$$\frac{\partial \mathbb{E}(x|\sigma)}{\partial \sigma} > \frac{\partial \mathbb{E}(b|\sigma)}{\partial \sigma} \geq 0$$

for all circumstances $\sigma \in \mathbb{R}$. In addition, private benefits b follow a conditional distribution $G(b|\sigma)$ on $[0, \bar{b}]$ with $\bar{b} > 0$, a differentiable and positive density $g(b|\sigma)$, and a non-decreasing hazard rate $g(b|\sigma)/(1 - G(b|\sigma))$ for all $\sigma \in \mathbb{R}$.⁷ A non-decreasing hazard rate is a common assumption in screening settings. Many familiar distributions, like the uniform, the normal or the exponential distribution, satisfy this property.

Now consider as a benchmark the constrained first-best, if circumstances and private benefits are observable and verifiable. Welfare remains unchanged if the firm takes no action. If the firm takes the action, it generates private benefits b and externalities $-x$. Thus, weighted welfare changes by $\alpha b - x$. Hence, a firm with circumstances σ and private benefits b should be active, whenever $\alpha b - \mathbb{E}(x|\sigma) \geq 0$ or, equivalently, $b \geq \mathbb{E}(x|\sigma)/\alpha$. Then the firms depicted in Figure 2 are active. Constrained first-best is attainable if the enforcement authority observes circumstances and private benefits. Alternatively, constrained first-best is attainable if the enforcement authority observes circumstances and sets a penalty of $\mathbb{E}(x|\sigma)/\alpha$. Yet, the enforcement authority cannot observe and verify firms' type. The enforcement authority only learns whether a firm is suspicious. A firm is suspicious if and only if it is active and the noisy signal $\sigma - \Delta\delta$ exceeds the policy $\hat{\sigma}$ with

⁶Besanko and Spulber (1993) and Neven and Roller (2005) compare different welfare standards.

⁷The assumption of full support rules out perfect correlation between circumstances and private benefits. With perfect correlation the enforcement authority could infer private benefits from circumstances and therefore does not need legal uncertainty as a screening device.

- At $t = 0$, the enforcement authority chooses the policy $\hat{\sigma}$.
- At $t = 1$, the firm's type with circumstances σ , negative externalities x , and private benefits b is realized and revealed to the firm.
- At $t = 2$, the firm chooses whether to take the action.
- At $t = 3$, the enforcement authority learns whether a firm is suspicious and imposes the penalty π at costs κ on suspicious firms.

Figure 1: Timing of the Model

a constant $\Delta \in [0, \infty)$ and a random $\delta \in [-1, 1]$. The noise δ follows a distribution H with a differentiable and positive density h . In the case of vertical restraints this captures the difficulty in determining, whether the market share is 29% or 31%. This uncertainty about the policy threshold or this measurement error is implied by the legal norms and is exogenous to enforcement authorities.

If the enforcement authority detects a suspicious firm, it makes the firm pay a penalty $\pi > 0$. The value of the penalty π is determined by law and, hence, exogenous. Here externalities are unobservable to the enforcement authority. In reality, there might be different legal and organizational reasons why it is impossible to differentiate the penalty according to occurred harm. Therefore, a constant and exogenous penalty is a common assumption in the literature on deterrence. See, e.g., Kaplow (2011). One theoretical reason for discontinuous penalties are costs κ for imposing a penalty π as in Polinsky and Shavell (2000). I assume $\kappa \leq (1 - \alpha)\pi$ and $\kappa < \pi$. The costs κ capture experts' testimonies, reports and other expenses to prove the enforcement authority's case. In Section 4.5, I discuss alternative penalties. To avoid technical constraints on the distribution functions, I assume that welfare is concave in the policy $\hat{\sigma}$ and that there are no corner solutions for the optimal policy. Figure 1 summarizes the timing.

To sum up, weighted welfare is

$$\begin{cases} 0 & \text{if the firm is deterred} \\ \alpha b - x & \text{if the firm is active and not fined} \\ \alpha b - x + \pi(1 - \alpha) - \kappa & \text{if the firm is active and fined} \end{cases} \quad (1)$$

with weighted private benefits αb , externalities x , welfare effects $\pi(1 - \alpha)$ of imposing a penalty, and enforcement costs κ . For a total welfare standard and $\alpha = 1$, imposing a penalty π does not affect welfare except for enforcement costs. Otherwise, there is a change in welfare.⁸ The firm's pay-offs are

$$\begin{cases} 0 & \text{if the firm is deterred} \\ b & \text{if the firm is active and not fined} \\ b - \pi & \text{if the firm is active and fined.} \end{cases}$$

⁸This assumption is more common in some fields than in other fields. Distortionary taxation is a common assumption in, e.g., public finance or for regulatory risk as in Strausz (2011). My main results, in particular, Theorem 1, are also valid without this assumption, if α is not too low.

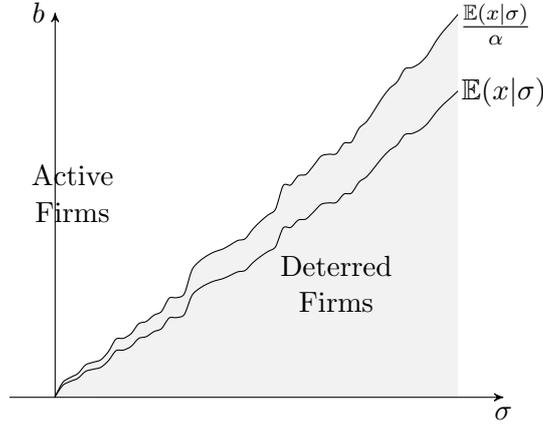


Figure 2: Active Firms in the Constrained First-Best

Given the policy constraints, the enforcement authority usually cannot enforce the constrained first-best. The next section shows that legal uncertainty allows mitigating these limitations of the enforcement authority.

4 Effects of Legal Uncertainty

4.1 Deterrence Effects

To characterize deterrence effects of legal uncertainty, consider the firm's decision. The firm faces the penalty π if the enforcement authority judges the firm to be suspicious. Therefore the firm only takes the action if its private benefits b are high enough. Thus, only firms above a cut-off $\hat{b}(\sigma)$ are active. This deterrence cut-off varies with circumstances σ . If circumstances are low, the firm knows that there are no penalties. Hence, the firm takes the action, as long as private benefits are positive. Beginning at $\sigma = \hat{\sigma} - \Delta$, the firm expects a penalty with some probability. Thus, deterrence increases in circumstances and equals the expected penalty, i.e., probabilities of conviction multiplied by the penalty. Finally, for high circumstances, $\sigma \geq \hat{\sigma} + \Delta$, the firm anticipates to be penalized for sure. Hence, deterrence is constant and equals the penalty π . Consequently, deterrence is

$$\hat{b}(\sigma) = \begin{cases} 0 & \text{if } \sigma < \hat{\sigma} - \Delta \\ \pi H\left(\frac{\sigma - \hat{\sigma}}{\Delta}\right) & \text{if } \hat{\sigma} - \Delta \leq \sigma < \hat{\sigma} + \Delta \\ \pi & \text{if } \sigma \geq \hat{\sigma} + \Delta. \end{cases} \quad (2)$$

Therefore legal uncertainty allows screening of firms. If they are close to the policy $\hat{\sigma}$, firms with low private benefits abstain from taking the action for lower circumstances σ than firms with high private benefits. Figure 3 depicts these deterrence effects. If legal norms provide legal certainty, the signal of the enforcement authority would be perfect, i.e., $\Delta = 0$, and deterrence would be sharp. Then below the policy $\hat{\sigma}$, all firms take the action. Above $\hat{\sigma}$, only those firms with private benefits above the penalty π will implement the action. The following lemma describes these deterrence effects of legal uncertainty.

Lemma 1. *Below the policy $\hat{\sigma}$, legal uncertainty (weakly) increases deterrence; above $\hat{\sigma}$, legal uncertainty (weakly) decreases deterrence.*

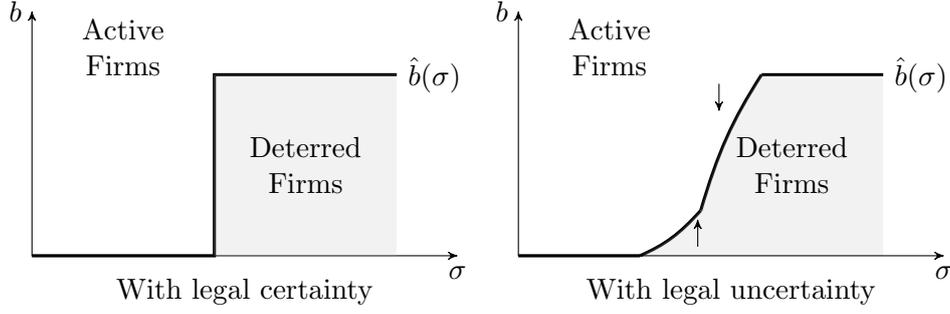


Figure 3: Deterrence Effects of Legal Uncertainty

In conclusion, legal uncertainty changes deterrence in opposite directions.

4.2 Optimal Policies

The next step considers optimal policies. The enforcement authority chooses the policy $\hat{\sigma}$ to maximize welfare which equals

$$\int \left(1 - G(\hat{b}(\sigma)|\sigma)\right) \left(\alpha \mathbb{E}(b|\sigma, b \geq \hat{b}(\sigma)) - \mathbb{E}(x|\sigma) + \hat{b}(\sigma)((1 - \alpha)\pi - \kappa)/\pi\right) dF(\sigma). \quad (3)$$

Deterrence $\hat{b}(\sigma)$ depends on the policy and influences which firms take the action. A fraction $1 - G(\hat{b}(\sigma)|\sigma)$ of firms is active. The enforcement authority expects firms' private benefits to be $\mathbb{E}(b|\sigma, b \geq \hat{b}(\sigma))$ and externalities to equal $-\mathbb{E}(x|\sigma)$. The definition of deterrence $\hat{b}(\sigma)$ in (2) implies that $\hat{b}(\sigma)/\pi = H(\frac{\sigma - \hat{\sigma}}{\Delta})$ is the probability of conviction. This probability depends on the distribution H of the error term. For low circumstances, $\sigma < \hat{\sigma} - \Delta$, there are no convictions and every firm is active. Hence, welfare equals $\mathbb{E}(\alpha b - x|\sigma)$ according to Eq. (1). For high circumstances, $\sigma > \hat{\sigma} + \Delta$, conviction is certain and only firms with the highest benefits are active. Therefore, welfare equals $(1 - G(\pi|\sigma))(\alpha \mathbb{E}(b|\sigma, b \geq \pi) - \mathbb{E}(x|\sigma) + \pi(1 - \alpha) - \kappa)$. Around the policy, legal uncertainty influences which firms take the action. Thus, probabilities of conviction increase and the fraction of active firms decreases in the circumstances.

Raising the policy makes more firms become active and raises negative externalities. The optimal policy balances these additional externalities with private benefits and enforcement costs.

Lemma 2. *With legal certainty, $\Delta = 0$, the optimal policy $\hat{\sigma}^C$ is determined by*

$$\mathbb{E}(x|\hat{\sigma}^C) = \alpha \mathbb{E}(b|\hat{\sigma}^C, b \leq \pi) - \frac{1 - G(\pi|\hat{\sigma}^C)}{G(\pi|\hat{\sigma}^C)}((1 - \alpha)\pi - \kappa). \quad (4)$$

With legal uncertainty, the same considerations apply in expectations. Again additional externalities are balanced with private benefits and enforcement costs.

Proposition 1. *With legal uncertainty, $\Delta > 0$, the optimal policy $\hat{\sigma}^\circ$ is determined by*

$$\int_{\hat{\sigma}^\circ - \Delta}^{\hat{\sigma}^\circ + \Delta} g(\hat{b}(\sigma)|\sigma) \frac{\partial \hat{b}(\sigma)}{\partial \hat{\sigma}} \left(\mathbb{E}(x|\sigma) - \hat{b}(\sigma) \frac{\pi - \kappa}{\pi} + \frac{1 - G(\hat{b}(\sigma)|\sigma)}{g(\hat{b}(\sigma)|\sigma)} \frac{(1 - \alpha)\pi - \kappa}{\pi} \right) dF(\sigma) = 0$$

Notice that only actions $(\hat{\sigma}^\circ - \Delta, \hat{\sigma}^\circ + \Delta)$ around the policy matter for determining the optimal policy. Deterrence in this region is determined by the distribution G of private benefits and the cut-off $\hat{b}(\sigma)$. The cut-off captures expected penalties as discussed above. While the equations in Lemma 2 and Proposition 1 differ, optimal policies for small legal uncertainty and for legal certainty are closely related.

Proposition 2. *If legal uncertainty vanishes, the optimal policy $\hat{\sigma}^\circ$ converges to the optimal policy $\hat{\sigma}^C$ with legal certainty, $\lim_{\Delta \searrow 0} \hat{\sigma}^\circ = \hat{\sigma}^C$.*

Hence, optimal policies balance externalities with private benefits adjusted for firms' behavioral responses.

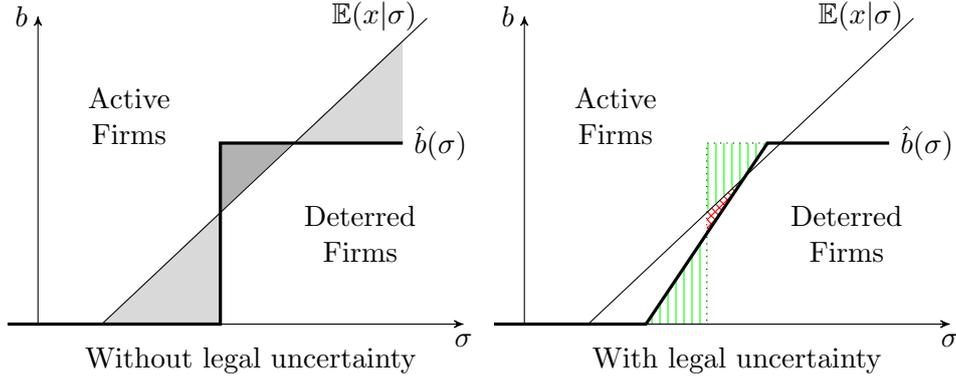
4.3 Welfare Effects

Now turn to the main result of this paper, the welfare effects of legal uncertainty. Legal uncertainty changes how many and which firms are active as depicted in Figure 3. These changes in deterrence increase welfare if legal uncertainty is not too large.

Theorem 1. *There is a $\bar{\Delta} > 0$, such that welfare increases in legal uncertainty Δ for all $\Delta \in [0, \bar{\Delta})$. The threshold $\bar{\Delta}$ is determined in the proof.*

Legal uncertainty increases welfare if the uncertainty is not too large. For the intuition, consider the case $\alpha = 1$ here, while postponing the general setting to the appendix. Remember that expected externalities $\mathbb{E}(x|\sigma)$ in Figures 2 and 4 characterize the constrained first-best. The left-hand side of Figure 4 shows that without legal uncertainty there is over- and under-deterrence compared to the constrained first-best – as indicated by the gray areas. In particular, firms close to the threshold are over- and under-deterred. According to Lemma 1, legal uncertainty mitigates these deterrence problems. Legal uncertainty implies that firms close to, but below the policy threshold expect the penalty with some probability. Hence, legal uncertainty deters these firms from taking the action if their private benefits are low. In addition, legal uncertainty implies that firms close to, but above the policy threshold are not penalized with some probability. Hence, legal uncertainty makes these firms take the action if their private benefits are high. These are precisely the firms that are over- and under-deterred without legal uncertainty. Therefore both changes increase welfare as indicated by the striped areas in Figure 4. Any potential losses in welfare – as indicated by the checked areas – are small compared to the welfare gains. This comparison follows from the definition of the optimal policy $\hat{\sigma}^\circ$. The enforcement agency chooses this policy such that externalities are balanced by private benefits for the average firm with circumstances $\hat{\sigma}^\circ$. This balance ensures that the overall effect of legal uncertainty on welfare is positive if the uncertainty is not too large. Consequently, some legal uncertainty increases welfare.

At this level of generality, the threshold $\bar{\Delta}$ for legal uncertainty to be beneficial is determined only implicitly. For many specifications, however, it is possible to determine



On the left-hand side, actions in the absence of legal uncertainty are depicted and compared to constrained first-best. Gray areas indicate over- and under-deterred firms. Introducing legal uncertainty on the right-hand side changes the implemented actions. The striped areas show increases in welfare. The checked area shows decreases in welfare. The total effect on welfare is positive.

Figure 4: Intuition of Theorem 1

this threshold explicitly. For example, the threshold equals

$$\bar{\Delta} = \frac{\pi - \kappa}{2\gamma}$$

if H is uniform, $\mathbb{E}(x|\sigma) = \gamma\sigma$ with a $\gamma > 0$, and $G(b|\sigma)$ is constant in σ . Alternatively, the threshold is

$$\bar{\Delta} = \frac{\pi}{\xi} \min_{\delta \in [-1,1]} h(\delta)$$

if $\alpha = 1$ and $\partial\mathbb{E}(x|\sigma)/\partial\sigma \leq \xi$ with a $\xi > 0$. These thresholds are intuitive. The threshold $\bar{\Delta}$ increases in the penalty π . The larger the penalty, the higher the step in deterrence $\hat{b}(\sigma)$ without legal uncertainty on the left-hand side of Figure 3. Therefore there are more under-deterred firms below the policy $\hat{\sigma}^\circ$ and more over-deterred firms above the policy $\hat{\sigma}^\circ$. Hence, there is more room for legal uncertainty to increase welfare and the threshold $\bar{\Delta}$ increases in the penalty π . To the contrary, the threshold $\bar{\Delta}$ decreases in the marginal externalities captured by γ and ξ . The higher the marginal externalities, the more externalities react to changes in circumstances and the steeper the constrained first-best in Figure 4. Therefore there are less under-deterred firms below the policy $\hat{\sigma}^\circ$ and less over-deterred firms above the policy $\hat{\sigma}^\circ$. Hence, there is less room for legal uncertainty to increase welfare and the threshold $\bar{\Delta}$ decreases in the marginal externalities. In addition, higher marginal externalities increase welfare losses due to noise in the enforcement authority's signal about circumstances reinforcing the first effect. Finally, the threshold $\bar{\Delta}$ decreases in the costs κ , because legal uncertainty increases the expected number of convictions. Hence, higher costs κ make legal uncertainty less appealing.

Return to the example of vertical restraints. Theorem 1 says that some uncertainty about the commission's market share estimates might be welfare enhancing. Consider two firms. Firm A has a market share of 25% and low private benefits. Firm B has a market share of 35% and high private benefits. Suppose in the constrained first-best firm A should be deterred, while firm B should be active. Without legal uncertainty, the

commission's market share estimates are perfect and equal 25% and 35%. Firm A knows that it fits the block exemption and does not expect any penalty. Therefore firm A is active, while firm B is deterred without legal uncertainty. Yet, discrepancies in the definition of the relevant market or imprecision in the measurement of sales create uncertainty about the commission's market share estimates. Hence, firm A faces some probability that the commission's estimate is above 30%. Then firm A becomes deterred. Vice versa for firm B. With legal uncertainty, firm B faces some probability that the commission's estimate is below 30%. Therefore, the expected penalty for firm B decreases and firm B becomes active. Both effects of legal uncertainty raise welfare.

4.4 Effects on Optimal Policies

Next, consider the effects of legal uncertainty on the optimal policy $\hat{\sigma}^\circ$. To focus on the effects of uncertainty, I neglect second-order effects of distributions changing in the circumstances. Whether legal uncertainty increases or decreases the optimal policy depends on the mean of the image measure $J(\delta) := G(\pi H(\delta)|\hat{\sigma}^\circ)/G(\pi|\hat{\sigma}^\circ)$. The image measure captures the error distribution H and the distribution G of the private benefits. This image measure determines the optimal policy in Proposition 1.

Proposition 3. *Suppose G is constant in σ , $\mathbb{E}(x|\sigma) = \sigma$, and $f'(\sigma) = 0$ for all $\sigma \in (\min\{\hat{\sigma}^C, \hat{\sigma}^\circ\} - 2\Delta, \max\{\hat{\sigma}^C, \hat{\sigma}^\circ\} + 2\Delta)$. If $\mathbb{E}_{J(\delta)}(\delta) \geq 0$, legal uncertainty decreases the optimal policy. Conversely, if $\mathbb{E}_{J(\delta)}(\delta) \leq 0$, legal uncertainty raises the optimal policy.*

Whenever the mean of the image measure $J(\delta)$ is distorted, legal uncertainty changes the optimal policy. For the first part of the intuition, suppose that private benefits are uniformly distributed. Then the image measure depends only on the error distribution. If the error distribution H is distorted upwards and $\mathbb{E}_{H(\delta)}(\delta) > 0$, the enforcement authority knows that circumstances are on average higher than the noisy signal. To make up for this distortion, the enforcement authority chooses a lower policy. Hence, legal uncertainty decreases optimal policies. If the error distribution H is distorted downwards and $\mathbb{E}_{H(\delta)}(\delta) < 0$, the enforcement authority knows that circumstances are on average lower than the noisy signal. To make up for this distortion, the enforcement authority chooses a higher policy. Therefore, legal uncertainty raises optimal policies.

For the second part of the intuition, suppose that errors are uniformly distributed. If private benefits are left-skewed in the relevant range, there are more firms with higher private benefits than with lower private benefits. Hence, there is more over-deterrence than under-deterrence close to the policy $\hat{\sigma}^\circ$. Therefore without legal uncertainty it is optimal to set a rather high policy. As explained in Theorem 1, legal uncertainty mitigates these deterrence problems and a lower policy becomes optimal. If private benefits are right-skewed in the relevant range, the argument is reversed. The image measure J combines both effects. Thus, the enforcement authority adapts its policy to the uncertainty.

4.5 Alternative Penalties

Finally, consider the question whether the enforcement authority can replicate the beneficial effects of legal uncertainty by setting penalties appropriately. So far penalties were determined exogenously by the law. Now assume that the enforcement authority sets penalties π at $t = 0$. If the enforcement authority can only observe whether a firm is suspicious, the optimal (endogenous) penalty usually is $\pi = \bar{b}$ deterring all firms above $\hat{\sigma}^\circ + \Delta$ from taking the action. Theorem 1 remains valid in this case and legal uncertainty is beneficial.

In reality, there might be different legal and organizational reasons why it is impossible to differentiate penalties finely. In the model this corresponds to the restricted observability of firm's type by the enforcement authority. Neglect these restrictions for a moment and assume that penalties for suspicious firms depend on private benefits and equal $\pi(b) = \gamma b + \lambda$ with $\gamma, \lambda > 0$. To avoid additional case distinctions and keep the analysis tractable, assume a total welfare standard with $\alpha = 1$. Deterrence $\hat{b}(\sigma)$ changes, but remains qualitatively similar to initial deterrence in Eq. (2).⁹ Therefore Lemmas 1, 2, and 3 remain valid. The intuition for my main result fully carries over to this setting.

Corollary 1. *Suppose penalties depend on private benefits, $\pi(b) = \gamma b + \lambda$ with $\gamma, \lambda > 0$. Then legal uncertainty increases welfare if the uncertainty is not too large.*

Finally, assume that the enforcement authority chooses penalties and observes a noisy signal $\sigma^S = \sigma - \Delta\delta$ of the circumstances. Then penalties are a function of the signal σ^S . For ease of exposition, assume a total welfare standard with $\alpha = 1$.

Proposition 4. *Suppose penalties depend on a signal σ^S and there is legal certainty. Then optimal penalties equal*

$$\pi(\sigma^S) = \begin{cases} 0 & \text{for } \sigma^S < \hat{\sigma}^\circ \\ \mathbb{E}(x|\sigma^S) + \kappa & \text{for } \sigma^S \geq \hat{\sigma}^\circ. \end{cases}$$

For socially beneficial types with low signals σ^S , the enforcement authority tolerates active firms by setting penalties to zero. As the policy $\hat{\sigma}^\circ$ is passed, strictly positive penalties are imposed, because lower penalties do not justify spending the costs κ to enforce the penalties. Beginning with $\hat{\sigma}^\circ$, penalties π increase in the signal σ^S . Notice that there is a step in the penalties at $\hat{\sigma}^\circ$, because it is not optimal to use very low penalties. This discontinuity is sufficient for legal uncertainty to increase welfare. The discontinuity ensures a similar deterrence pattern around the policy $\hat{\sigma}^\circ$ as before with a exogenously given penalty. Therefore the intuition carries over and there are positive welfare effects to legal uncertainty.

Proposition 5. *Suppose penalties depend on a signal σ^S and $\mathbb{E}(x|\sigma) = \sigma$ for all σ . Then legal uncertainty increases welfare if the uncertainty is not too large.*

⁹Exact deterrence is stated in the proof of Corollary 1.

Costs κ for imposing a penalty are just one of several possible theoretical reasons for discontinuous penalties. In summary, my main message is valid quite generally: Some legal uncertainty acts as a welfare enhancing screen and has positive welfare effects.

5 Numerical Example of the Effects of Legal Uncertainty

This numerical example of my setting as set up in Section 3 emphasizes the welfare effects of legal uncertainty. Suppose that circumstances σ are normally distributed with mean $\mu = 200$ and variance $\sigma^2 = 10000$. Expected externalities equal the circumstances. The private benefits b are uniformly distributed on the interval $[0, 600]$. I consider welfare weights $\alpha = 1/3$ ($2/3$ and 1 , resp.). For simplicity, I assume that the penalty is sufficiently high to deter all firms, i.e., $\pi = 600$. The costs of imposing penalties are $\kappa = (1 - \alpha)\pi$. The error term in the noisy signal is uniformly distributed.

According to Lemma 2, the optimal policy is $\hat{\sigma}^C = \alpha\mathbb{E}(b) = 300\alpha$ without legal uncertainty. All firms below the policy threshold $\hat{\sigma}^C$ will be active, while all firms above $\hat{\sigma}^C$ will be deterred. Hence, at the optimal policy externalities are on average offset by private benefits weighted with the welfare weight α . Yet, firms above the policy with high private benefits are over-deterred, while firms below the policy with low private benefits are under-deterred. Now, add legal uncertainty. Recall my main result in Theorem 1: Some legal uncertainty increases welfare, because legal uncertainty acts as a welfare enhancing screen. Figure 5 shows how welfare depends on legal uncertainty. Beginning at $\Delta = 0$ with legal certainty, legal uncertainty increases. Crucially, welfare increases for small amounts of legal uncertainty, in particular for $\Delta \leq 300\alpha$. For $\Delta = 300\alpha$, welfare with legal uncertainty approaches even constrained first-best welfare levels. The reason is that the uniform distribution of private benefits and error terms allows to match constrained first-best deterrence as depicted in Figure 2. This does not hold in general. Nevertheless, it indicates substantial increases in welfare by legal uncertainty, even if the amount of uncertainty is not optimal. Table 1 summarizes maximal increases in welfare. Appendix B discusses optimal policies and additional benchmarks.

For the intuition, remember the asymmetric information between firms and the enforcement authority. The enforcement authority cannot observe the firm's private benefits. Without legal uncertainty, the firm's decision is therefore independent of its private benefits. Legal uncertainty influences deterrence in opposite directions. If a firm is close

	$\mu = 250,$ $\sigma^2 = 10000$	$\mu = 200,$ $\sigma^2 = 10000$	$\mu = 250,$ $\sigma^2 = 5000$	$\mu = 200,$ $\sigma^2 = 5000$
$\alpha = 1/3$	77%	48%	305%	148%
$\alpha = 2/3$	103%	56%	244%	99%
$\alpha = 1$	58%	30%	77%	34%

Table 1: Relative increase in welfare by legal uncertainty compared to legal certainty

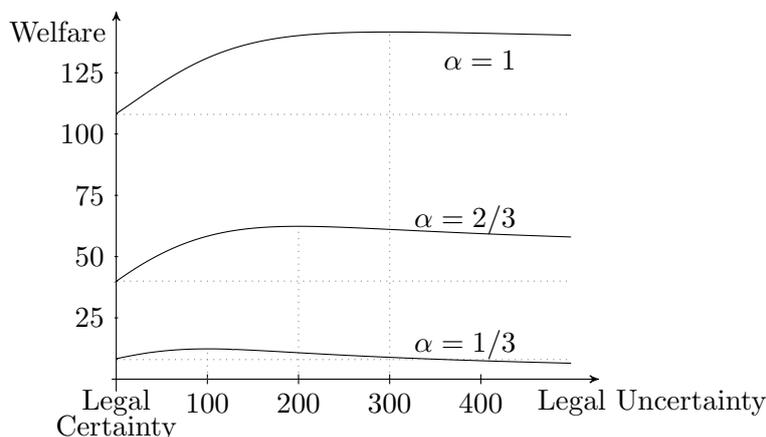


Figure 5: Welfare depending on Legal Uncertainty

to, but below the policy, legal uncertainty implies that the firm gets convicted with some probability. Therefore firms with low private benefits do not take the action. Legal uncertainty deters them. If a firm is close to, but above the policy, legal uncertainty implies that the firm might get off without a penalty. Hence firms with high private benefits take the action. Legal uncertainty encourages them to take the action. Both effects increase welfare. Table 1 and Figure 5 show that this increase in welfare can be significant.

6 Conclusion

Legal studies frequently consider legal certainty as a value in itself. This paper takes a welfare perspective and studies the welfare effects of legal uncertainty. I show that some legal uncertainty raises social welfare. Suppose that the threshold of legality is not specified exactly. In addition, there is asymmetric information between firms and the enforcement authority. The enforcement authority cannot observe firms' private benefits of an action. With legal certainty, hence, firms decide independently of the action's private benefits whether to take the action.¹⁰ With legal *uncertainty*, the probability of conviction depends on the distance to the threshold of legality. If a firm is close to, but below the threshold, there is some probability of being convicted. Therefore firms with low private benefits do not take the action. Legal uncertainty deters them. If a firm is close to, but above the threshold, there is some probability of not being convicted. Hence firms with high private benefits take the action. Legal uncertainty encourages them to take the action. Both effects increase social welfare independently of underlying distributions. The effects are also robust to random audits, alternative penalties and different welfare standards. Consequently, some legal uncertainty about the threshold of legality increases welfare.

The case for legal uncertainty I make here is just one argument in a large debate, like, e.g., Schinkel (2010). In particular, there are limitations to the benefits of legal uncertainty. Welfare effects are certainly not monotone in the amount of uncertainty. Furthermore, there may be negative effects of legal uncertainty that are not captured in my analysis.

¹⁰Here expected penalties equal the maximum private benefits, but proofs do not rely on this assumption.

For instance, legal uncertainty may reduce the accountability of the enforcement authority making it more challenging to deter corruption or regulatory capture. In addition, legal uncertainty might result in socially wasteful expenses in safeguards and evidence production. Policymakers, however, might positively influence the effects of legal uncertainty and steer deterrence towards harmful behavior by complementing a general norm with specific exceptions, like safe harbors, or detailed information with respect to some procedural aspects.¹¹ Yet, there are also additional benefits of legal uncertainty. Precise legal norms often imply huge social costs, for example, in lengthy court cases about conceptual futilities. These costs are a deadweight loss. Legal uncertainty allows to save these costs in addition to the screening effects scrutinized in my model.

An interesting avenue for future research are the dynamic effects of legal uncertainty. Legal uncertainty might give firms incentives to experiment and therefore implement more controversial actions with negative externalities than under legal certainty. Yet, the costs of such behavior, e.g., possible penalties, are incurred by individual firms, while the benefits spill over to all firms, as they learn, e.g., court decisions reducing legal uncertainty.¹² Similar effects are discussed in the literature on strategic learning and multi-armed bandits.

Even in a static model, legal uncertainty has many causes. I consider enforcement policies carried out by administrative bodies as well as judgments taken by courts. Uncertainty about the threshold of legality might be due to missing precedents or unclear norms. In particular, the legal uncertainty is not by design, but inherent in legal norms. I do not recommend designing especially ambiguous norms. Yet, if norms contain some legal uncertainty, this uncertainty need not be a drawback, but might even increase welfare.

A Appendix

Proof of Lemma 1: The derivative of $\hat{b}(\sigma)$ with respect to legal uncertainty Δ equals

$$\frac{\partial \hat{b}(\sigma)}{\partial \Delta} = \begin{cases} \pi h\left(\frac{\sigma - \hat{\sigma}}{\Delta}\right) \frac{\hat{\sigma} - \sigma}{\Delta^2} & \text{for } \sigma \in (\hat{\sigma} - \Delta, \hat{\sigma} + \Delta) \\ 0 & \text{otherwise.} \end{cases}$$

Therefore deterrence increases for $\sigma < \hat{\sigma}$ and decreases for $\sigma > \hat{\sigma}$. \square

Proof of Lemma 2: Suppose there is no legal uncertainty and $\Delta = 0$. Then welfare equals

$$\int_{-\infty}^{\hat{\sigma}} \alpha \mathbb{E}(b|\sigma) - \mathbb{E}(x|\sigma) dF(\sigma) + \int_{\hat{\sigma}}^{\infty} (1 - G(\pi|\sigma)) \left(\alpha \mathbb{E}(b|\sigma, b \geq \pi) - \mathbb{E}(x|\sigma) + (1 - \alpha)\pi - \kappa \right) dF(\sigma).$$

The first derivative of welfare with respect to the policy $\hat{\sigma}$ yields the first-order condition:

$$\mathbb{E}(x|\hat{\sigma}^C) = \frac{1}{G(\pi|\hat{\sigma}^C)} \left(\alpha \mathbb{E}(b|\hat{\sigma}^C) - (1 - G(\pi|\hat{\sigma}^C)) \left(\alpha \mathbb{E}(b|\hat{\sigma}^C, b \geq \pi) + (1 - \alpha)\pi - \kappa \right) \right) =$$

¹¹For instance, the FTC-DOJ merger guidelines use safe harbors for small market shares. Alternatively, Christiansen and Kerber (2006) propose modified or structured rules of reason. In a different context, Demougin and Fluet (2008) suggest burden of proof guidelines.

¹²D'Amato (1983) argues that legal uncertainty increases over time. Ben-Shahar (1997) models the trade-off between magnitude and probability of punishment in a dynamic setting.

$$= \alpha \mathbb{E}(b|\hat{\sigma}^C, b \leq \pi) - \frac{1 - G(\pi|\hat{\sigma}^C)}{G(\pi|\hat{\sigma}^C)}((1 - \alpha)\pi - \kappa).$$

The first-order condition determines the unique solution, because welfare is concave in the policy $\hat{\sigma}$. \square

Lemma 3. *The derivative $\partial W(\hat{\sigma})/\partial y$ of welfare $W(\hat{\sigma})$ with respect to $y \in \{\hat{\sigma}, \Delta\}$ equals*

$$\int_{\hat{\sigma}-\Delta}^{\hat{\sigma}+\Delta} g(\hat{b}(\sigma)|\sigma) \frac{\partial \hat{b}(\sigma)}{\partial y} \left(\mathbb{E}(x|\sigma) - \hat{b}(\sigma) \frac{\pi - \kappa}{\pi} + \frac{1 - G(\hat{b}(\sigma)|\sigma)}{g(\hat{b}(\sigma)|\sigma)} \frac{(1 - \alpha)\pi - \kappa}{\pi} \right) dF(\sigma).$$

Proof: Taking the derivative of welfare with respect to $y \in \{\hat{\sigma}, \Delta\}$ results in

$$\begin{aligned} \frac{\partial W(\hat{\sigma})}{\partial y} &= \int_{\hat{\sigma}-\Delta}^{\hat{\sigma}+\Delta} -g(\hat{b}(\sigma)|\sigma) \frac{\partial \hat{b}(\sigma)}{\partial y} \left(\alpha \mathbb{E}(b|\sigma, b \geq \hat{b}(\sigma)) - \mathbb{E}(x|\sigma) + \hat{b}(\sigma) \frac{(1 - \alpha)\pi - \kappa}{\pi} \right) + \\ &\quad + \left(1 - G(\hat{b}(\sigma)|\sigma) \right) \left(\alpha \frac{\partial \mathbb{E}(b|\sigma, b \geq \hat{b}(\sigma))}{\partial y} + \frac{\partial \hat{b}(\sigma)}{\partial y} \frac{(1 - \alpha)\pi - \kappa}{\pi} \right) dF(\sigma). \end{aligned}$$

For $y = \Delta$, I use the envelope theorem here assuming $\hat{\sigma} = \hat{\sigma}^\circ$. Notice that the conditional expectation equals

$$\mathbb{E}(b|\sigma, b \geq \tilde{b}) = \frac{1}{1 - G(\tilde{b}|\sigma)} \int_{\tilde{b}}^{\bar{b}} b dG(b|\sigma).$$

Hence,

$$\begin{aligned} \frac{\partial \mathbb{E}(b|\sigma, b \geq \tilde{b})}{\partial \tilde{b}} &= \frac{1}{(1 - G(\tilde{b}|\sigma))^2} \left(-(1 - G(\tilde{b}|\sigma)) \tilde{b} g(\tilde{b}|\sigma) + g(\tilde{b}|\sigma) \int_{\tilde{b}}^{\bar{b}} b dG(b|\sigma) \right) = \\ &= g(\tilde{b}|\sigma) \frac{\mathbb{E}(b|\sigma, b \geq \tilde{b}) - \tilde{b}}{1 - G(\tilde{b}|\sigma)}. \end{aligned}$$

Together with the chain rule, this results in $\frac{\partial \mathbb{E}(b|\sigma, b \geq \hat{b}(\sigma))}{\partial y} = g(\hat{b}(\sigma)|\sigma) \frac{\partial \hat{b}(\sigma)}{\partial y} \frac{\mathbb{E}(b|\sigma, b \geq \hat{b}(\sigma)) - \hat{b}(\sigma)}{1 - G(\hat{b}(\sigma)|\sigma)}$ for $y \in \{\hat{\sigma}, \Delta\}$. Therefore the derivative with respect to $y \in \{\hat{\sigma}, \Delta\}$ equals

$$\int_{\hat{\sigma}-\Delta}^{\hat{\sigma}+\Delta} g(\hat{b}(\sigma)|\sigma) \frac{\partial \hat{b}(\sigma)}{\partial y} \left(\mathbb{E}(x|\sigma) - \hat{b}(\sigma) \alpha + \left(\frac{1 - G(\hat{b}(\sigma)|\sigma)}{g(\hat{b}(\sigma)|\sigma)} - \hat{b}(\sigma) \right) \frac{(1 - \alpha)\pi - \kappa}{\pi} \right) dF(\sigma).$$

Rearranging yields the statement of the lemma. \square

Proof of Proposition 1: Lemma 3 on p. 17 above determines the first derivative of welfare with respect to the policy $\hat{\sigma}$. Hence, the first-order condition reads

$$\begin{aligned} \int_{\hat{\sigma}^\circ-\Delta}^{\hat{\sigma}^\circ+\Delta} g(\hat{b}(\sigma)|\sigma) h\left(\frac{\sigma - \hat{\sigma}^\circ}{\Delta}\right) \mathbb{E}(x|\sigma) dF(\sigma) = \\ \int_{\hat{\sigma}^\circ-\Delta}^{\hat{\sigma}^\circ+\Delta} g(\hat{b}(\sigma)|\sigma) h\left(\frac{\sigma - \hat{\sigma}^\circ}{\Delta}\right) \left(\hat{b}(\sigma) \frac{\pi - \kappa}{\pi} - \frac{1 - G(\hat{b}(\sigma)|\sigma)}{g(\hat{b}(\sigma)|\sigma)} \frac{(1 - \alpha)\pi - \kappa}{\pi} \right) dF(\sigma) \end{aligned} \quad (5)$$

The exact threshold of legality depends on expected externalities $\mathbb{E}(x|\sigma)$ and the distributions F , G , and H . The first-order condition (5) determines the unique solution, because welfare is concave in the policy $\hat{\sigma}$. \square

Proof of Proposition 2: Rearranging the first-order condition for the optimal policy in Proposition 1 yields

$$\int_{\hat{\sigma}^\circ - \Delta}^{\hat{\sigma}^\circ + \Delta} g(\hat{b}(\sigma)|\sigma) \frac{\partial \hat{b}(\sigma)}{\partial \hat{\sigma}} \left(\hat{b}(\sigma)\alpha - \left(\frac{1 - G(\hat{b}(\sigma)|\sigma)}{g(\hat{b}(\sigma)|\sigma)} - \hat{b}(\sigma) \right) \frac{(1 - \alpha)\pi - \kappa}{\pi} - \mathbb{E}(x|\sigma) \right) dF(\sigma) = 0 \quad (6)$$

Notice that $\partial \hat{b}(\sigma)/\partial \sigma = -\partial \hat{b}(\sigma)/\partial \hat{\sigma}$ for all σ and

$$\begin{aligned} - \int_{\hat{\sigma}^\circ - \Delta}^{\hat{\sigma}^\circ + \Delta} g(\hat{b}(\sigma)|\hat{\sigma}^\circ) \frac{\partial \hat{b}(\sigma)}{\partial \hat{\sigma}} d\sigma &= \int_{\hat{\sigma}^\circ - \Delta}^{\hat{\sigma}^\circ + \Delta} \frac{\partial G(\hat{b}(\sigma)|\hat{\sigma}^\circ)}{\partial \sigma} d\sigma = G(\hat{b}(\hat{\sigma}^\circ + \Delta)|\hat{\sigma}^\circ) - G(\hat{b}(\hat{\sigma}^\circ - \Delta)|\hat{\sigma}^\circ) = \\ &= G(\pi|\hat{\sigma}^\circ) - G(0|\hat{\sigma}^\circ) = G(\pi|\hat{\sigma}^\circ). \end{aligned}$$

Hence, expected externalities in (6) converge to

$$\lim_{\Delta \searrow 0} \int_{\hat{\sigma}^\circ - \Delta}^{\hat{\sigma}^\circ + \Delta} g(\hat{b}(\sigma)|\sigma) \frac{\partial \hat{b}(\sigma)}{\partial \hat{\sigma}} \mathbb{E}(x|\sigma) dF(\sigma) = -G(\pi|\hat{\sigma}^\circ) f(\hat{\sigma}^\circ) \mathbb{E}(x|\hat{\sigma}^\circ). \quad (7)$$

Analogously, integration by substitution of σ by b yields

$$\lim_{\Delta \searrow 0} \int_{\hat{\sigma}^\circ - \Delta}^{\hat{\sigma}^\circ + \Delta} g(\hat{b}(\sigma)|\sigma) \frac{\partial \hat{b}(\sigma)}{\partial \hat{\sigma}} \hat{b}(\sigma) dF(\sigma) = -f(\hat{\sigma}^\circ) \int_0^\pi b g(b|\hat{\sigma}^\circ) db = -f(\hat{\sigma}^\circ) G(\pi|\hat{\sigma}^\circ) \mathbb{E}(b|\hat{\sigma}^\circ, b \leq \pi).$$

Finally,

$$\begin{aligned} \int_{\hat{\sigma}^\circ - \Delta}^{\hat{\sigma}^\circ + \Delta} -g(\hat{b}(\sigma)|\hat{\sigma}^\circ) \frac{\partial \hat{b}(\sigma)}{\partial \hat{\sigma}} \left(\frac{1 - G(\hat{b}(\sigma)|\hat{\sigma}^\circ)}{g(\hat{b}(\sigma)|\hat{\sigma}^\circ)} - \hat{b}(\sigma) \right) d\sigma &= \int_{\hat{\sigma}^\circ - \Delta}^{\hat{\sigma}^\circ + \Delta} \frac{\partial(1 - G(\hat{b}(\sigma)|\hat{\sigma}^\circ)) \hat{b}(\sigma)}{\partial \sigma} d\sigma = \\ &= (1 - G(\hat{b}(\hat{\sigma}^\circ + \Delta)|\hat{\sigma}^\circ)) \hat{b}(\hat{\sigma}^\circ + \Delta) - (1 - G(\hat{b}(\hat{\sigma}^\circ - \Delta)|\hat{\sigma}^\circ)) \hat{b}(\hat{\sigma}^\circ - \Delta) = (1 - G(\pi|\hat{\sigma}^\circ)) \pi. \end{aligned}$$

Consequently, the first terms of (6) converge to

$$-f(\hat{\sigma}^\circ) G(\pi|\hat{\sigma}^\circ) \left(\alpha \mathbb{E}(b|\hat{\sigma}^\circ, b \leq \pi) - \frac{1 - G(\pi|\hat{\sigma}^\circ)}{G(\pi|\hat{\sigma}^\circ)} ((1 - \alpha)\pi - \kappa) \right). \quad (8)$$

Rewrite the first-order condition (4) in Lemma 2 with legal certainty as

$$0 = \alpha \mathbb{E}(b|\hat{\sigma}^C, b \leq \pi) - \frac{1 - G(\pi|\hat{\sigma}^C)}{G(\pi|\hat{\sigma}^C)} ((1 - \alpha)\pi - \kappa) - \mathbb{E}(x|\hat{\sigma}^C). \quad (9)$$

The previous steps have shown that the right-hand side of Eq. (6) converges to the right-hand side of Eq. (9) multiplied by $-f(\hat{\sigma}^\circ) G(\pi|\hat{\sigma}^\circ) < 0$. Therefore, the optimal policy converges and $\lim_{\Delta \searrow 0} \hat{\sigma}^\circ = \hat{\sigma}^C$, as the uncertainty vanishes. \square

Proof of Theorem 1: By Lemma 3 (p. 17), the derivative of welfare (3) with respect to the amount of legal uncertainty Δ equals

$$\int_{\hat{\sigma}^\circ - \Delta}^{\hat{\sigma}^\circ + \Delta} g(\hat{b}(\sigma)|\sigma) \frac{\partial \hat{b}(\sigma)}{\partial \Delta} \left(\mathbb{E}(x|\sigma) - \hat{b}(\sigma) \frac{\pi - \kappa}{\pi} + \frac{1 - G(\hat{b}(\sigma)|\sigma)}{g(\hat{b}(\sigma)|\sigma)} \frac{(1 - \alpha)\pi - \kappa}{\pi} \right) dF(\sigma).$$

Denote the term in brackets by $\mathcal{T}(\sigma) = \mathbb{E}(x|\sigma) - \hat{b}(\sigma) \frac{\pi - \kappa}{\pi} + \frac{1 - G(\hat{b}(\sigma)|\sigma)}{g(\hat{b}(\sigma)|\sigma)} \frac{(1 - \alpha)\pi - \kappa}{\pi}$. The function $\mathcal{T}(\sigma)$ decreases in $\sigma \in (\hat{\sigma}^\circ - \Delta, \hat{\sigma}^\circ + \Delta)$ for sufficiently small $\Delta > 0$, because the derivative of $\mathcal{T}(\sigma)$ with respect to σ equals

$$\frac{\partial \mathcal{T}(\sigma)}{\partial \sigma} = \frac{\partial \mathbb{E}(x|\sigma)}{\partial \sigma} - \frac{\partial \hat{b}(\sigma)}{\partial \sigma} \frac{\pi - \kappa}{\pi} + \left(\frac{\partial \frac{1 - G(\hat{b}(\sigma)|\sigma)}{g(\hat{b}(\sigma)|\sigma)}}{\partial b} \frac{\partial \hat{b}(\sigma)}{\partial \sigma} + \frac{\partial \frac{1 - G(\hat{b}(\sigma)|\sigma)}{g(\hat{b}(\sigma)|\sigma)}}{\partial \sigma} \right) \frac{(1 - \alpha)\pi - \kappa}{\pi} =$$

$$\begin{aligned}
&= \frac{\partial \mathbb{E}(x|\sigma)}{\partial \sigma} + \frac{\partial \frac{1-G(\hat{b}(\sigma)|\sigma)}{g(\hat{b}(\sigma)|\sigma)} (1-\alpha)\pi - \kappa}{\partial \sigma} \frac{1-\alpha}{\pi} \\
&\quad - \frac{\pi}{\Delta} h\left(\frac{\sigma - \hat{\sigma}^\circ}{\Delta}\right) \underbrace{\left(\frac{\pi - \kappa}{\pi} - \frac{\partial \frac{1-G(\hat{b}(\sigma)|\sigma)}{g(\hat{b}(\sigma)|\sigma)} (1-\alpha)\pi - \kappa}{\partial b} \frac{1-\alpha}{\pi}\right)}_{>0} < 0 \tag{10}
\end{aligned}$$

for sufficiently small $\Delta > 0$. The (weakly) increasing hazard rate of the distribution G of b implies a (weakly) decreasing inverse hazard rate. Therefore the term in brackets in Eq. (10) is positive as $\kappa \leq (1-\alpha)\pi$ and $\kappa < \pi$ by assumption. As $\frac{\pi}{\Delta}$ goes to infinity for $\Delta \rightarrow 0$ and the first two terms do not depend on Δ , the derivative $\partial \mathcal{T}(\sigma)/\partial \sigma$ is negative for sufficiently small $\Delta > 0$. In particular, there is a $\bar{\Delta} > 0$ such that $\mathcal{T}(\sigma)$ is decreasing in $\sigma \in [\hat{\sigma}^\circ - \Delta, \hat{\sigma}^\circ + \Delta]$ for all $\Delta \in (0, \bar{\Delta})$.

Proposition 1 derives the first-order condition (5) for the optimal policy $\hat{\sigma}^\circ$:

$$\int_{\hat{\sigma}^\circ - \Delta}^{\hat{\sigma}^\circ + \Delta} g(\hat{b}(\sigma)|\sigma) h\left(\frac{\sigma - \hat{\sigma}^\circ}{\Delta}\right) \mathcal{T}(\sigma) dF(\sigma) = 0. \tag{11}$$

In the relevant range, $g(\cdot) \geq 0$ and $h(\cdot) > 0$ with strict inequalities for some circumstances σ . As $\mathcal{T}(\sigma)$ decreases in σ , this first-order condition (11) implies $\mathcal{T}(\hat{\sigma}^\circ - \Delta) > 0 > \mathcal{T}(\hat{\sigma}^\circ + \Delta)$ for all $\Delta \in [0, \bar{\Delta})$. In particular, for all $\Delta \in [0, \bar{\Delta})$ there is a $\sigma_0 \in (\hat{\sigma}^\circ - \Delta, \hat{\sigma}^\circ + \Delta)$ such that

$$\mathcal{T}(\sigma) \begin{cases} > 0 & \text{for all } \sigma < \sigma_0 \\ < 0 & \text{for all } \sigma > \sigma_0. \end{cases}$$

The first-order condition (11) yields

$$\int_{\hat{\sigma}^\circ - \Delta}^{\hat{\sigma}^\circ} g(\hat{b}(\sigma)|\sigma) h\left(\frac{\sigma - \hat{\sigma}^\circ}{\Delta}\right) \mathcal{T}(\sigma) dF(\sigma) = - \int_{\hat{\sigma}^\circ}^{\hat{\sigma}^\circ + \Delta} g(\hat{b}(\sigma)|\sigma) h\left(\frac{\sigma - \hat{\sigma}^\circ}{\Delta}\right) \mathcal{T}(\sigma) dF(\sigma) \geq 0. \tag{12}$$

The second term is non-negative, because $\mathcal{T}(\sigma)$ changes sign exactly once. Assume for the moment $\sigma_0 \leq \hat{\sigma}^\circ$. Then $\mathcal{T}(\sigma) < 0$ for all $\sigma \in [\hat{\sigma}^\circ, \hat{\sigma}^\circ + \Delta)$. Remember that Lemma 1 proves that the derivative $\partial \hat{b}(\sigma)/\partial \Delta$ is negative for $\sigma \in (\hat{\sigma}^\circ, \hat{\sigma}^\circ + \Delta)$. Hence,

$$\int_{\hat{\sigma}^\circ}^{\hat{\sigma}^\circ + \Delta} g(\hat{b}(\sigma)|\sigma) \frac{\partial \hat{b}(\sigma)}{\partial \Delta} \mathcal{T}(\sigma) dF(\sigma) \geq 0$$

for all $\Delta \in [0, \bar{\Delta})$, because $\frac{\partial \hat{b}(\sigma)}{\partial \Delta} \mathcal{T}(\sigma) > 0$ for all circumstances σ in the relevant range. In addition,

$$\begin{aligned}
&\int_{\hat{\sigma}^\circ - \Delta}^{\hat{\sigma}^\circ} g(\hat{b}(\sigma)|\sigma) \frac{\partial \hat{b}(\sigma)}{\partial \Delta} \mathcal{T}(\sigma) dF(\sigma) = \\
&= \int_{\hat{\sigma}^\circ - \Delta}^{\sigma_0} g(\hat{b}(\sigma)|\sigma) \pi \frac{\hat{\sigma}^\circ - \sigma}{\Delta^2} h\left(\frac{\sigma - \hat{\sigma}^\circ}{\Delta}\right) \underbrace{\mathcal{T}(\sigma)}_{>0} dF(\sigma) + \int_{\sigma_0}^{\hat{\sigma}^\circ} g(\hat{b}(\sigma)|\sigma) \pi \frac{\hat{\sigma}^\circ - \sigma}{\Delta^2} h\left(\frac{\sigma - \hat{\sigma}^\circ}{\Delta}\right) \underbrace{\mathcal{T}(\sigma)}_{<0} dF(\sigma) > \\
&> \frac{\hat{\sigma}^\circ - \sigma_0}{\Delta^2} \int_{\hat{\sigma}^\circ - \Delta}^{\sigma_0} g(\hat{b}(\sigma)|\sigma) \pi h\left(\frac{\sigma - \hat{\sigma}^\circ}{\Delta}\right) \mathcal{T}(\sigma) dF(\sigma) + \frac{\hat{\sigma}^\circ - \sigma_0}{\Delta^2} \int_{\sigma_0}^{\hat{\sigma}^\circ} g(\hat{b}(\sigma)|\sigma) \pi h\left(\frac{\sigma - \hat{\sigma}^\circ}{\Delta}\right) \mathcal{T}(\sigma) dF(\sigma) = \\
&= \frac{\hat{\sigma}^\circ - \sigma_0}{\Delta^2} \pi \int_{\hat{\sigma}^\circ - \Delta}^{\hat{\sigma}^\circ} g(\hat{b}(\sigma)|\sigma) h\left(\frac{\sigma - \hat{\sigma}^\circ}{\Delta}\right) \mathcal{T}(\sigma) dF(\sigma) \geq 0
\end{aligned}$$

for all $\Delta \in [0, \bar{\Delta}]$, because $\int_{\hat{\sigma}^\circ - \Delta}^{\hat{\sigma}^\circ} g(\hat{b}(\sigma)|\sigma)h(\frac{\sigma - \hat{\sigma}^\circ}{\Delta})\mathcal{T}(\sigma)dF(\sigma) \geq 0$ by (12). The case $\sigma_0 > \hat{\sigma}^\circ$ is analogous and therefore omitted. Consequently,

$$\frac{\partial W(\hat{\sigma}^\circ)}{\partial \Delta} = \int_{\hat{\sigma}^\circ - \Delta}^{\hat{\sigma}^\circ} g(\hat{b}(\sigma)|\sigma) \frac{\partial \hat{b}(\sigma)}{\partial \Delta} \mathcal{T}(\sigma) dF(\sigma) + \int_{\hat{\sigma}^\circ}^{\hat{\sigma}^\circ + \Delta} g(\hat{b}(\sigma)|\sigma) \frac{\partial \hat{b}(\sigma)}{\partial \Delta} \mathcal{T}(\sigma) dF(\sigma) > 0$$

for all $\Delta \in [0, \bar{\Delta}]$. In addition, welfare (3) is continuous in $\Delta \in [0, \infty)$. Hence, some legal uncertainty increases welfare. \square

Proof of Proposition 3: Lemma 2 states the first-order condition (4) for the optimal policy σ if there is legal certainty, $\Delta = 0$. Rewrite this condition as

$$G(\pi|\hat{\sigma}^C)\mathbb{E}(x|\hat{\sigma}^C) = \int_0^\pi \alpha b dG(b|\hat{\sigma}^C) - (1 - G(\pi|\hat{\sigma}^C))((1 - \alpha)\pi - \kappa). \quad (13)$$

Notice that

$$\begin{aligned} 1 - G(\pi|\hat{\sigma}) &= (1 - G(\pi H(1)|\hat{\sigma})) \underbrace{H(1)}_{=1} - (1 - G(\pi H(-1)|\hat{\sigma})) \underbrace{H(-1)}_{=0} = \\ &= \int_{-1}^1 \frac{\partial(1 - G(\pi H(\delta)|\hat{\sigma}))H(\delta)}{\partial \delta} d\delta = \\ &= \int_{-1}^1 h(\delta) \left(1 - G(\pi H(\delta)|\hat{\sigma}) - g(\pi H(\delta)|\hat{\sigma})\pi H(\delta)\right) d\delta = \\ &= \int_{-1}^1 g(\pi H(\delta)|\hat{\sigma})h(\delta) \left(\frac{1 - G(\pi H(\delta)|\hat{\sigma})}{g(\pi H(\delta)|\hat{\sigma})} - \pi H(\delta)\right) d\delta = \\ &= \int_{\hat{\sigma} - \Delta'}^{\hat{\sigma} + \Delta'} g(\pi H(\frac{\sigma - \hat{\sigma}}{\Delta'})|\hat{\sigma})h(\frac{\sigma - \hat{\sigma}}{\Delta'}) \left(\frac{1 - G(\pi H(\frac{\sigma - \hat{\sigma}}{\Delta'})|\hat{\sigma})}{g(\pi H(\frac{\sigma - \hat{\sigma}}{\Delta'})|\hat{\sigma})} - \pi H(\frac{\sigma - \hat{\sigma}}{\Delta'})\right) \frac{1}{\Delta'} d\sigma \end{aligned}$$

for any $\Delta' > 0$. In addition,

$$\begin{aligned} \int_0^\pi b dG(b|\hat{\sigma}) &= \int_0^{\pi H(1)} b dG(b|\hat{\sigma}) - \int_0^{\pi H(-1)} b dG(b|\hat{\sigma}) = \int_{-1}^1 \frac{\partial \int_0^{\pi H(\delta)} b dG(b|\hat{\sigma})}{\partial \delta} d\delta = \\ &= \int_{-1}^1 g(\pi H(\delta)|\hat{\sigma})\pi H(\delta)\pi h(\delta) d\delta = \int_{\hat{\sigma} - \Delta'}^{\hat{\sigma} + \Delta'} g(\pi H(\frac{\sigma - \hat{\sigma}}{\Delta'})|\hat{\sigma})\pi h(\frac{\sigma - \hat{\sigma}}{\Delta'})\pi H(\frac{\sigma - \hat{\sigma}}{\Delta'}) \frac{1}{\Delta'} d\sigma \end{aligned}$$

for any $\Delta' > 0$. Suppose $f'(x) = 0$ in the relevant range. Then inserting the last two equations into the first-order condition (13) yields

$$\begin{aligned} G(\pi|\hat{\sigma}^C)\mathbb{E}(x|\hat{\sigma}^C) \frac{\Delta'}{\pi} &= \quad (14) \\ &= \int_{\hat{\sigma}^C - \Delta'}^{\hat{\sigma}^C + \Delta'} g(\pi H(\frac{\sigma - \hat{\sigma}^C}{\Delta'})|\hat{\sigma}^C)h(\frac{\sigma - \hat{\sigma}^C}{\Delta'}) \left(H(\frac{\sigma - \hat{\sigma}^C}{\Delta'}) (\pi - \kappa) - \frac{1 - G(\pi H(\frac{\sigma - \hat{\sigma}^C}{\Delta'})|\hat{\sigma}^C)}{g(\pi H(\frac{\sigma - \hat{\sigma}^C}{\Delta'})|\hat{\sigma}^C)} \frac{(1 - \alpha)\pi - \kappa}{\pi} \right) d\sigma. \end{aligned}$$

for any $\Delta' > 0$. Comparing the first-order conditions with legal uncertainty (5) and without legal uncertainty (14), we see that the right-hand side of both conditions is the same up to a multiplicative constant if $\Delta' = \Delta$ and f is constant in the relevant range. Additionally, the right-hand side of (14) is constant in $\hat{\sigma}^C$, because the right-hand side of (13) is constant in $\hat{\sigma}^C$. Suppose that, in addition,

$$0 \geq \mathbb{E}_{J(\delta)}(\delta) = \int_{-1}^1 \delta dJ(\delta) \quad (15)$$

with the image measure $J(\delta) = G(\pi H(\delta)|\hat{\sigma})/G(\pi|\hat{\sigma})$. Notice that

$$\begin{aligned} \int_{-1}^1 \delta dJ(\delta) &= \int_{-1}^1 \delta \frac{g(\pi H(\delta)|\hat{\sigma}^C)}{G(\pi|\hat{\sigma}^C)} \pi h(\delta) d\delta = \int_{\hat{\sigma}^C - \Delta}^{\hat{\sigma}^C + \Delta} \frac{\sigma - \hat{\sigma}^C}{\Delta} \frac{g(\hat{b}(\sigma)|\hat{\sigma}^C)}{G(\pi|\hat{\sigma}^C)} \frac{\pi}{\Delta} h\left(\frac{\sigma - \hat{\sigma}^C}{\Delta}\right) d\sigma = \\ &= \int_{\hat{\sigma}^C - \Delta}^{\hat{\sigma}^C + \Delta} \frac{\sigma}{\Delta} \frac{g(\hat{b}(\sigma)|\hat{\sigma}^C)}{G(\pi|\hat{\sigma}^C)} \frac{\pi}{\Delta} h\left(\frac{\sigma - \hat{\sigma}^C}{\Delta}\right) d\sigma - \frac{\hat{\sigma}^C}{\Delta}. \end{aligned}$$

Hence, condition (15) implies

$$\mathbb{E}(x|\hat{\sigma}^C) = \hat{\sigma}^C \geq \int_{\hat{\sigma}^C - \Delta}^{\hat{\sigma}^C + \Delta} \sigma \frac{g(\hat{b}(\sigma|\hat{\sigma}^C))}{G(\pi|\hat{\sigma}^C)} \frac{\pi}{\Delta} h\left(\frac{\sigma - \hat{\sigma}^C}{\Delta}\right) d\sigma. \quad (16)$$

$$\begin{aligned} \text{Therefore } \int_{\hat{\sigma}^C - \Delta}^{\hat{\sigma}^C + \Delta} g(\hat{b}(\sigma)|\hat{\sigma}^C) h\left(\frac{\sigma - \hat{\sigma}^C}{\Delta}\right) \mathbb{E}(x|\sigma) d\sigma &= \\ = \int_{\hat{\sigma}^C - \Delta}^{\hat{\sigma}^C + \Delta} \sigma \frac{g(\hat{b}(\sigma)|\hat{\sigma}^C)}{G(\pi|\hat{\sigma}^C)} G(\pi|\hat{\sigma}^C) \frac{\Delta}{\pi} \frac{\pi}{\Delta} h\left(\frac{\sigma - \hat{\sigma}^C}{\Delta}\right) d\sigma &\leq G(\pi|\hat{\sigma}^C) \frac{\Delta}{\pi} \mathbb{E}(x|\hat{\sigma}^C). \end{aligned}$$

Condition (16) ensures the inequality. Hence, at $\hat{\sigma}^C$ the left-hand side of the first-order condition (5) with legal uncertainty is smaller than the right-hand side. Consequently, legal uncertainty increases the threshold of legality if condition (15) is satisfied.

Analogously, legal uncertainty decreases the threshold of legality if $\mathbb{E}_{J(\delta)}(\delta) \geq 0$. \square

Proof of Corollary 1: The cut-off for active firms equals

$$\hat{b}(\sigma) = \begin{cases} 0 & \text{if } \sigma < \hat{\sigma} - \Delta \\ \frac{\lambda H(\frac{\sigma - \hat{\sigma}}{\Delta})}{1 - \gamma H(\frac{\sigma - \hat{\sigma}}{\Delta})} & \text{if } \hat{\sigma} - \Delta \leq \sigma < \hat{\sigma} + \Delta \text{ and } \gamma H(\frac{\sigma - \hat{\sigma}}{\Delta}) < 1 \\ \frac{\lambda}{1 - \gamma} & \text{if } \sigma \geq \hat{\sigma} + \Delta \text{ and } \gamma < 1 \\ \infty & \text{if } \sigma \geq \hat{\sigma} - \Delta \text{ and } \gamma H(\frac{\sigma - \hat{\sigma}}{\Delta}) \geq 1. \end{cases}$$

By Lemma 3, the derivative of welfare with respect to legal uncertainty Δ is

$$\int_{\hat{\sigma}^* - \Delta}^{\hat{\sigma}^* + \Delta} g(\hat{b}(\sigma)|\sigma) \frac{\partial \hat{b}(\sigma)}{\partial \Delta} \left(\mathbb{E}(x|\sigma) - \hat{b}(\sigma) \frac{\pi - \kappa}{\pi} \right) dF(\sigma).$$

Denote the term in brackets by $\tilde{\mathcal{T}}(\sigma) = \mathbb{E}(x|\sigma) - \hat{b}(\sigma) \frac{\pi - \kappa}{\pi}$. Analogously to the proof of Theorem 1, the function $\tilde{\mathcal{T}}(\sigma)$ decreases in $\sigma \in (\hat{\sigma}^* - \Delta, \hat{\sigma}^* + \Delta)$ for sufficiently small $\Delta > 0$. The first-order condition for the optimal policy $\hat{\sigma}^*$ is

$$\int_{\hat{\sigma}^* - \Delta}^{\hat{\sigma}^* + \Delta} g(\hat{b}(\sigma)|\sigma) \frac{\partial \hat{b}(\sigma)}{\partial \hat{\sigma}} \tilde{\mathcal{T}}(\sigma) dF(\sigma) = 0.$$

In the relevant range, $\partial \hat{b}(\sigma)/\partial \hat{\sigma} \leq 0$ with strict inequality for some circumstances σ . As $\tilde{\mathcal{T}}(\sigma)$ decreases in σ , this first-order condition implies that there are circumstances $\sigma_0 \in (\hat{\sigma}^* - \Delta, \hat{\sigma}^* + \Delta)$ such that

$$\tilde{\mathcal{T}}(\sigma) \begin{cases} > 0 & \text{for all } \sigma < \sigma_0 \\ < 0 & \text{for all } \sigma > \sigma_0. \end{cases}$$

This first-order condition also implies that

$$\int_{\hat{\sigma}^\circ - \Delta}^{\hat{\sigma}^\circ} g(\hat{b}(\sigma)|\sigma) \frac{\partial \hat{b}(\sigma)}{\partial \hat{\sigma}^\circ} \tilde{\mathcal{T}}(\sigma) dF(\sigma) = - \int_{\hat{\sigma}^\circ}^{\hat{\sigma}^\circ + \Delta} g(\hat{b}(\sigma)|\sigma) \frac{\partial \hat{b}(\sigma)}{\partial \hat{\sigma}^\circ} \tilde{\mathcal{T}}(\sigma) dF(\sigma) \leq 0. \quad (17)$$

The second term is non-positive, because $-\partial \hat{b}(\sigma)/\partial \hat{\sigma}^\circ > 0$ and $\tilde{\mathcal{T}}(\sigma)$ changes sign exactly once. Assume for the moment $\sigma_0 \leq \hat{\sigma}^\circ$. Then $\tilde{\mathcal{T}}(\sigma) < 0$ for all $\sigma \in [\hat{\sigma}^\circ, \hat{\sigma}^\circ + \Delta)$. Remember that Lemma 1 shows that the derivative $\partial \hat{b}(\sigma)/\partial \Delta$ is negative for circumstances $\sigma \in (\hat{\sigma}^\circ, \hat{\sigma}^\circ + \Delta)$. Hence,

$$\int_{\hat{\sigma}^\circ}^{\hat{\sigma}^\circ + \Delta} g(\hat{b}(\sigma)|\sigma) \frac{\partial \hat{b}(\sigma)}{\partial \Delta} \tilde{\mathcal{T}}(\sigma) dF(\sigma) \geq 0,$$

because $\frac{\partial \hat{b}(\sigma)}{\partial \Delta} \tilde{\mathcal{T}}(\sigma) > 0$ for all circumstances σ in the relevant range. In addition,

$$\begin{aligned} \int_{\hat{\sigma}^\circ - \Delta}^{\hat{\sigma}^\circ} g(\hat{b}(\sigma)|\sigma) \frac{\partial \hat{b}(\sigma)}{\partial \Delta} \tilde{\mathcal{T}}(\sigma) dF(\sigma) &= \int_{\hat{\sigma}^\circ - \Delta}^{\sigma_0} g(\hat{b}(\sigma)|\sigma) \frac{\hat{\sigma}^\circ - \sigma}{\Delta^2} \frac{\lambda h(\frac{\sigma - \hat{\sigma}^\circ}{\Delta})}{(1 - H(\frac{\sigma - \hat{\sigma}^\circ}{\Delta}))^2} \underbrace{\tilde{\mathcal{T}}(\sigma)}_{>0} dF(\sigma) + \\ &+ \int_{\sigma_0}^{\hat{\sigma}^\circ} g(\hat{b}(\sigma)|\sigma) \frac{\hat{\sigma}^\circ - \sigma}{\Delta^2} \frac{\lambda h(\frac{\sigma - \hat{\sigma}^\circ}{\Delta})}{(1 - H(\frac{\sigma - \hat{\sigma}^\circ}{\Delta}))^2} \underbrace{\tilde{\mathcal{T}}(\sigma)}_{<0} dF(\sigma) > \\ &> \frac{\hat{\sigma}^\circ - \sigma_0}{\Delta^2} \left(\int_{\hat{\sigma}^\circ - \Delta}^{\sigma_0} g(\hat{b}(\sigma)|\sigma) \frac{\lambda h(\frac{\sigma - \hat{\sigma}^\circ}{\Delta})}{(1 - H(\frac{\sigma - \hat{\sigma}^\circ}{\Delta}))^2} \tilde{\mathcal{T}}(\sigma) dF(\sigma) + \int_{\sigma_0}^{\hat{\sigma}^\circ} g(\hat{b}(\sigma)|\sigma) \frac{\lambda h(\frac{\sigma - \hat{\sigma}^\circ}{\Delta})}{(1 - H(\frac{\sigma - \hat{\sigma}^\circ}{\Delta}))^2} \tilde{\mathcal{T}}(\sigma) dF(\sigma) \right) = \\ &= \frac{\hat{\sigma}^\circ - \sigma_0}{\Delta} \int_{\hat{\sigma}^\circ - \Delta}^{\hat{\sigma}^\circ} g(\hat{b}(\sigma)|\sigma) (-1) \frac{\partial \hat{b}(\sigma)}{\partial \hat{\sigma}^\circ} \tilde{\mathcal{T}}(\sigma) dF(\sigma) \geq 0, \end{aligned}$$

because $\int_{\hat{\sigma}^\circ - \Delta}^{\hat{\sigma}^\circ} g(\hat{b}(\sigma)|\sigma) \frac{\partial \hat{b}(\sigma)}{\partial \hat{\sigma}^\circ} \tilde{\mathcal{T}}(\sigma) dF(\sigma) \leq 0$ by (17). The case $\sigma_0 > \hat{\sigma}^\circ$ is analogous and therefore omitted. Consequently, $\partial W(\hat{\sigma}^\circ)/\partial \Delta > 0$ for sufficiency small Δ . Hence, some legal uncertainty increases welfare. \square

Proof of Proposition 4: Welfare equals

$$\int (1 - G(\pi(\sigma)|\sigma)) \left(\mathbb{E}(b|\sigma, b \geq \pi(\sigma)) - \mathbb{E}(x|\sigma) - \kappa \mathbf{1}_{\pi(\sigma) > 0} \right) dF(\sigma)$$

with $\mathbf{1}_{\pi(\sigma) > 0} = 1$ for all σ with $\pi(\sigma) > 0$ and $\mathbf{1}_{\pi(\sigma) > 0} = 0$ otherwise. Pointwise optimization of welfare with respect to penalties $\pi(\sigma)$ yields

$$-g(\pi(\sigma)|\sigma) \left(\mathbb{E}(b|\sigma, b \geq \pi(\sigma)) - \mathbb{E}(x|\sigma) - \kappa \right) + (1 - G(\pi(\sigma)|\sigma)) \left(\frac{\partial \mathbb{E}(b|\sigma, b \geq \pi(\sigma))}{\partial \pi(\sigma)} \right) = 0$$

if positive penalties are imposed and $\pi(\sigma) > 0$. Lemma 3 (p.17) together with the chain rule results in

$$\frac{\partial \mathbb{E}(b|\sigma, b \geq \pi(\sigma))}{\partial \pi(\sigma)} = g(\pi(\sigma)|\sigma) \frac{\mathbb{E}(b|\sigma, b \geq \pi(\sigma)) - \pi(\sigma)}{1 - G(\pi(\sigma)|\sigma)}.$$

Inserting this derivative into the last equation yields

$$g(\pi(\sigma)|\sigma) (\mathbb{E}(x|\sigma) - \pi(\sigma) + \kappa) = 0 \quad \text{and for } g(\pi(\sigma)|\sigma) > 0: \quad \pi^*(\sigma) = \mathbb{E}(x|\sigma) + \kappa$$

if positive penalties are optimal. Yet, the penalties have to satisfy additional constraints. Consider a given σ . If penalties are zero and $\pi(\sigma) = 0$, the enforcement authority could

save the enforcement costs κ . Then welfare equals $\mathbb{E}(b - x|\sigma)$. If penalties are positive and $\pi(\sigma) > 0$, welfare equals

$$(1 - G(\pi(\sigma))) \left(\mathbb{E}(b|b \geq \pi(\sigma)) - \mathbb{E}(x|\sigma) - \kappa \right).$$

Comparing both terms shows that positive penalties are optimal if

$$(1 - G(\pi(\sigma)|\sigma))\kappa \leq (1 - G(\pi(\sigma)|\sigma))\mathbb{E}(b|\sigma, b \geq \pi(\sigma)) - \mathbb{E}(b|\sigma) + G(\pi(\sigma)|\sigma)\mathbb{E}(x|\sigma).$$

Rearranging this condition yields

$$(1 - G(\pi(\sigma)|\sigma))\kappa \leq G(\pi(\sigma)|\sigma)(\mathbb{E}(x|\sigma) - \mathbb{E}(b|\sigma, b \leq \pi(\sigma))).$$

Intuitively, imposing penalties $\pi(\sigma)$ deters a fraction $G(\pi(\sigma)|\sigma)$ of firms with average benefits of $\mathbb{E}(b|\sigma, b \leq \pi(\sigma))$, while costs are $(1 - G(\pi(\sigma)|\sigma))\kappa$. Plugging in optimal penalties $\pi^*(\sigma)$ defines the threshold $\hat{\sigma}^\circ$ by

$$\kappa = G(\mathbb{E}(x|\hat{\sigma}^\circ) + \kappa|\hat{\sigma}^\circ)(\kappa + \mathbb{E}(x|\hat{\sigma}^\circ) - \mathbb{E}(b|\hat{\sigma}^\circ, b \leq \mathbb{E}(x|\hat{\sigma}^\circ) + \kappa)) \quad (18)$$

Consequently, the penalties given in the proposition are optimal. \square

Proof of Proposition 5: Implement penalties of

$$\pi(\sigma^S) = \begin{cases} 0 & \text{for } \sigma^S < \hat{\sigma}^\circ - \Delta \\ \mathbb{E}(x|\sigma^S) + \kappa + \Delta \int_{-1}^1 \delta dH(\delta) & \text{for } \sigma^S \geq \hat{\sigma}^\circ - \Delta \end{cases} \quad (19)$$

with $\hat{\sigma}^\circ$ as defined in Eq. (18). It is easy to check that for all circumstances $\sigma \notin (\hat{\sigma}^\circ - 2\Delta, \hat{\sigma}^\circ)$ expected penalties and probabilities of conviction are the same as with legal certainty and $\Delta = 0$: A firm with circumstances $\sigma \leq \hat{\sigma}^\circ - 2\Delta$ expects to pay no penalties and a firm with circumstances $\sigma \geq \hat{\sigma}^\circ$ expects to pay penalties equal to $\mathbb{E}(\pi(\sigma^S)|\sigma) =$

$$\int_{-1}^1 \pi(\sigma - \Delta\delta) dH(\delta) = \int_{-1}^1 \mathbb{E}(x|\sigma - \Delta\delta) dH(\delta) + \kappa + \Delta \int_{-1}^1 \delta dH(\delta) = \sigma + \kappa = \mathbb{E}(x|\sigma) + \kappa$$

For all circumstances $\sigma \in (\hat{\sigma}^\circ - 2\Delta, \hat{\sigma}^\circ)$, expected penalties and probabilities of conviction increase compared to legal certainty and $\Delta = 0$. In particular, probabilities of conviction are $H((\sigma - \hat{\sigma}^\circ + \Delta)/\Delta)$ now. Then $\mathbb{E}(x|\hat{\sigma}^\circ) > \mathbb{E}(b|\hat{\sigma}^\circ, b \leq \mathbb{E}(x|\hat{\sigma}^\circ) + \kappa)$ by definition of $\hat{\sigma}^\circ$ in Eq. (18). Hence, the additional deterrence increases welfare for Δ sufficiently small. This effect is even stronger, because deterrence is not uniform, but those firms with low benefits are deterred. Yet, additional enforcement costs κ reduce welfare. Additional costs are

$$\kappa \int_{\hat{\sigma}^\circ - 2\Delta}^{\hat{\sigma}^\circ} \left(1 - G \left(\int_{-1}^{\frac{\sigma - \hat{\sigma}^\circ + \Delta}{\Delta}} \pi(\sigma - \Delta\delta) dH(\delta) \middle| \sigma \right) \right) H \left(\frac{\sigma - \hat{\sigma}^\circ + \Delta}{\Delta} \right) dF(\sigma)$$

with the first term in brackets capturing the fraction of active firms and the second term capturing probabilities of conviction. These costs can be reduced by fragmenting the set $(\hat{\sigma}^\circ - \Delta, \hat{\sigma}^\circ)$ into a partition. Consider an even $n \in \mathbb{N}$ and modify penalties in Eq. (19) to

$2\pi(\sigma)$ for all $\sigma \in (\hat{\sigma}^\circ - \Delta(2i + 1)/n, \hat{\sigma}^\circ - \Delta 2i/n)$ and all $i \in \{1, 2, \dots, n/2 - 1\}$ and to 0 otherwise. For n sufficiently large, deterrence remains almost unchanged, but probabilities of conviction are reduced. Therefore the benefits of legal uncertainty remain almost constant, while the additional costs are reduced. Continuing this procedure ensures that small amounts of legal uncertainty increase welfare. Since penalties $\pi(\sigma^S)$ in Eq. (19) are not optimal in general for $\Delta > 0$, welfare increases by setting penalties optimally. Consequently, legal uncertainty has even bigger positive welfare effects. \square

B Additional Material for the Numerical Example

As a benchmark, consider constrained first-best as depicted in Figure 2. All firms with low circumstances and $\sigma \leq \alpha b$ are active. Hence, welfare equals

$$\mathbb{E}(G(\sigma/\alpha|\sigma)(\alpha\mathbb{E}(b|\sigma, b \geq \sigma/\alpha) - \mathbb{E}(x|\sigma))) \approx 12$$

(62 and 142, resp., for the alternative welfare weights). Remember that I denote the distribution of private benefits b by $G(b|\sigma)$. As a second benchmark, consider welfare without legal uncertainty that equals

$$F(\hat{\sigma}^C)(\alpha\mathbb{E}(b) + \mathbb{E}(x|\sigma \leq \hat{\sigma}^C)) = F(300\alpha)(300\alpha + \mathbb{E}(x|\sigma \leq 300\alpha)) \approx 8$$

(40 and 108, resp.). Remember that I denote the distribution of circumstances σ by $F(\sigma)$. As a final benchmark, consider extreme policies. If all actions are prohibited, policy is $\hat{\sigma} = -\infty$ and welfare equals 0. If all actions are allowed, policy is $\hat{\sigma} = \infty$ and welfare equals $\alpha\mathbb{E}(b) - \mathbb{E}(x) = 300\alpha - 200 = -100$ (0 and 100, resp.). Therefore there are significant welfare gains by legal uncertainty and optimal policies. Remember that the optimal policy is $\hat{\sigma}^C = \alpha\mathbb{E}(b) = 300\alpha$ without legal uncertainty. With legal uncertainty, Proposition 1 determines optimal policies by

$$\int_{\hat{\sigma}^\circ - \Delta}^{\hat{\sigma}^\circ + \Delta} \mathbb{E}(x|\sigma) - \alpha\hat{b}(\sigma) dF(\sigma) = 0.$$

Table 2 shows optimal policies for some values.

	$\Delta = 0$, Legal Certainty	$\Delta = 1$	$\Delta = 50$	$\Delta = 100$	$\Delta = 300$	$\Delta = 1000$
$\alpha = 1/3$	100	100	107	100	-23	-655
$\alpha = 2/3$	200	200	200	200	200	200
$\alpha = 1$	300	299	271	264	300	533

Table 2: Optimal policies $\hat{\sigma}^\circ$ depending on Δ and α

For $\alpha = 2/3$, the average deterrence $\alpha\hat{b}(\hat{\sigma}^\circ) = 200$ just equals the mean μ of the circumstances distribution F . Around the mean the normal distribution F is symmetric. In addition, the image measure $J(\delta)$ from Proposition 3 is undistorted, i.e., $\mathbb{E}_{J(\delta)}(\delta) = 0$, because private benefits and error terms are uniformly distributed. Therefore, the optimal policy is 200 and constant in Δ . In general, optimal policies depend on legal uncertainty Δ .

References

- Arnott, R. and Stiglitz, J. E. (1988). Randomization with Asymmetric Information. *RAND Journal of Economics*, 19(3):344–362.
- Ben-Shahar, O. (1997). Playing without a rulebook: Optimal enforcement when individuals learn the penalty only by committing the crime. *International Review of Law and Economics*, 17(3):409–421.
- Besanko, D. and Spulber, D. F. (1989). Antitrust Enforcement Under Asymmetric Information. *Economic Journal*, 99(396):408–425.
- Besanko, D. and Spulber, D. F. (1993). Contested Mergers and Equilibrium Antitrust Policy. *Journal of Law, Economics and Organization*, 9(1):1–29.
- Calfee, J. E. and Craswell, R. (1984). Some Effects of Uncertainty on Compliance with Legal Standards. *Virginia Law Review*, 70(5):965–1003.
- Choné, P. and Linnemer, L. (2008). Assessing horizontal mergers under uncertain efficiency gains. *International Journal of Industrial Organization*, 26(4):913–929.
- Christiansen, A. and Kerber, W. (2006). Competition Policy with Optimally Differentiated Rules Instead of 'Per Se Rules Vs. Rule of Reason'. *Journal of Competition Law and Economics*, 2(2):215–244.
- Craswell, R. and Calfee, J. E. (1986). Deterrence and Uncertain Legal Standards. *Journal of Law, Economics and Organization*, 2(2):279–303.
- D'Amato, A. (1983). Legal Uncertainty. *California Law Review*, 71(1):1–55.
- Demougin, D. and Fluet, C. (2008). Rules of proof, courts, and incentives. *RAND Journal of Economics*, 39(1):20–40.
- Ehrlich, I. and Posner, R. A. (1974). An economic analysis of legal rulemaking. *Journal of Legal Studies*, 3(1):257–286.
- European Commission (2010). Guidelines on Vertical Restraints. *Official Journal of the European Union*, C 130(01):1–46.
- Friedman, E. and Wickelgren, A. L. (2014). A New Angle on Rules versus Standards. *American Law and Economics Review*, 16(2):499–549.
- Gauthier, S. and Laroque, G. (2014). On the value of randomization. *Journal of Economic Theory*, 151:493–507.
- Gergen, M. P. (2011). Uncertainty and Tax Enforcement: A Case for Modest Fault-Based Penalties. *Tax Law Review*, 64(4):453–487.
- Kaplow, L. (1990). Optimal Deterrence, Uninformed Individuals, and Acquiring Information about Whether Acts Are Subject to Sanctions. *Journal of Law, Economics and Organization*, 6(1):93–128.
- Kaplow, L. (1992). Rules Versus Standards: An Economic Analysis. *Duke Law Journal*, 42:557–629.
- Kaplow, L. (1994). The Value of Accuracy in Adjudication: An Economic Analysis. *Journal of Legal Studies*, 23(1):307–401.
- Kaplow, L. (1995). A Model of the Optimal Complexity of Legal Rules. *Journal of Law, Economics and Organization*, 11(1):150–163.
- Kaplow, L. (2011). Optimal Proof Burdens, Deterrence, and the Chilling of Desirable Behavior. *American Economic Review*, 101(3):277–280.
- Kaplow, L. and Shapiro, C. (2007). Antitrust. *Harvard Olin Discussion Paper Series*, 575.

- Kaplow, L. and Shavell, S. (1992). Private versus Socially Optimal Provision of Ex Ante Legal Advice. *Journal of Law, Economics and Organization*, 8(2):306–320.
- Katsoulacos, Y. and Ulph, D. (2009). On Optimal Legal Standards for Competition Policy: A General Welfare-Based Analysis. *Journal of Industrial Economics*, 57(3):410–437.
- Katsoulacos, Y. and Ulph, D. (2016a). Legal Uncertainty, Competition Law Enforcement Procedures and Optimal Penalties. *European Journal of Law and Economics*, 41(2):255–282.
- Katsoulacos, Y. and Ulph, D. (2016b). Regulatory decision errors, Legal Uncertainty and welfare: A general treatment. *International Journal of Industrial Organization*, forthcoming.
- Lando, H. (2006). Does wrongful conviction lower deterrence? *The Journal of Legal Studies*, 35(2):327–337.
- Lang, M. and Wambach, A. (2013). The Fog of Fraud – Mitigating Fraud by Strategic Ambiguity. *Games and Economic Behavior*, 81:255–275.
- Malik, A. S. (1990). Avoidance, Screening and Optimum Enforcement. *RAND Journal of Economics*, 21(3):341–353.
- Neven, D. J. and Roller, L.-H. (2005). Consumer surplus vs. welfare standard in a political economy model of merger control. *International Journal of Industrial Organization*, 23(9-10):829–848.
- Osofsky, L. (2011). The Case Against Strategic Tax Law Uncertainty. *Tax Law Review*, 64(4):489–538.
- Png, I. (1986). Optimal subsidies and damages in the presence of judicial error. *International Review of Law and Economics*, 6(1):101–105.
- Polinsky, A. M. and Shavell, S. (1989). Legal Error, Litigation, and the Incentive to Obey the Law. *Journal of Law, Economics and Organization*, 5(1):99–108.
- Polinsky, A. M. and Shavell, S. (2000). The Economic Theory of Public Enforcement of Law. *Journal of Economic Literature*, 38(1):45–76.
- Schinkel, M. P. (2010). Market Oversight Games. *Amsterdam Center for Law & Economics Working Paper*, 2010-11.
- Schinkel, M. P. and Tuinstra, J. (2006). Imperfect competition law enforcement. *International Journal of Industrial Organization*, 24(6):1267–1297. The Economics of Cartels, Cartel Policy, and Collusion.
- Strausz, R. (2006). Deterministic versus stochastic mechanisms in principal-agent models. *Journal of Economic Theory*, 128(1):306–314.
- Strausz, R. (2011). Regulatory risk under optimal incentive regulation. *Economic Journal*, 121(553):740–762.