Optimal Policy for Software Vulnerability Disclosure[1]

Ashish Arora, Rahul Telang, Hao Xu

H. John Heinz III School of Public Policy and Management
Carnegie Mellon University, Pittsburgh PA 15213

Email: {ashish; rtelang; xhao}@andrew.cmu.edu

Abstract

Software vulnerabilities represent a serious threat to cyber security, most cyber-attacks exploit known vulnerabilities. Unfortunately, there is no agreed-upon policy for their disclosure. Disclosure policy (which sets a protected period given to a vendor to release the patch for the vulnerability) indirectly affects the speed and quality of the patch that a vendor develops. Thus CERT/CC and similar bodies acting in the public interest can use disclosure to influence the behavior of vendors and reduce social cost. This paper develops a framework to analyze the optimal timing of disclosure. We formulate a model involving a social planner who sets the disclosure policy and a vendor who decides on the patch release. We show that the vendor typically release the patch less expeditiously than is socially optimal. The social planner optimally shrinks the protected period to push the vendor to deliver the patch more quickly and sometimes the patch release time coincides with disclosure. We extend the model to allow the proportion of users implementing patches to depend upon the quality (chosen by the vendor) of the patch. We show that a longer protected period does not always results in a better patch quality. Another extension allows for some fraction of users to use "work-arounds". We show that the possibility of work-arounds can provide the social planner more leverage and hence the social planner shrinks the protected period. Interestingly, possibility of work-arounds can sometimes increase the social cost due to the negative externalities imposed by the users who can use the work-arounds on the users who can not.

Keyword: Economics of Cyber-Security, Software Vulnerability, Disclosure Policy, Instant Disclosure, Patching, Patch Quality.

forthcoming: Management Science

[1] The authors thank the participants at the Third workshop on Economics and Information Security (WEIS 2004), Minneapolis, the Ninth INFORMS Conference on Information Systems and Technology (CIST) 2004, Denver, the ZEW Conference in Mannheim (2005) and seminar participants at Stanford University, for their valuable feedback. We also thank the DE, the AE, and two anonymous reviewers for many valuable suggestions, and Ed Barr for suggesting many improvements in the writing. This research was partially supported through a grant from Cylab, Carnegie Mellon University. Rahul Telang acknowledges the generous support of National Science Foundation through the CAREER award CNS - 0546009.

"First, the Nation needs a better-defined approach to the disclosure of vulnerabilities. The issue is complex because exposing vulnerabilities both helps speed the development of solutions and also creates opportunities for would be attackers."

The National Strategy to Secure Cyberspace, (2003: p 33)

1 Introduction

Information security breaches pose a significant and increasing threat to national security and economic well-being. In the Symantec Internet Security Threat Report (2003), companies reported experiencing about 30 attacks per week. These attacks often exploit software defects or vulnerabilities. The number of reported vulnerabilities has increased dramatically over time. The same Symantec report documented 2,524 vulnerabilities discovered in 2002, affecting over 2000 distinct products, an 81.5 percent increase over 2001. The CERT/CC (Computer Emergency Response Team/Coordination Center) received and cataloged 8,064 vulnerabilities in 2006 alone and reported more than 82,000 incidents involving various cyber-attacks. Although precise estimates are not available, losses from cyber-attacks can be substantial. The CSI-FBI survey estimates show that the loss per company was more than $500,000 in 2004, and more than $200,000 in 2005 (CSI-FBI 2005). Software vendors, including Microsoft, have announced their intention to reduce vulnerabilities in their products. Despite this, it is likely that vulnerabilities will continue to be discovered in the foreseeable future.

1.1 Vulnerability Disclosure Policies

There is considerable debate about how software vulnerabilities should be disclosed. In one view, discoverers should report vulnerabilities to vendors and wait until the vendor develops a patch. However, since a vendor is unlikely to fully internalize all user-losses when a vulnerability is exploited, some believe that often patches are excessively delayed. This belief fueled the creation of full-disclosure mailing lists in late 90's, such as "Bugtraq" where vulnerability information is disclosed immediately and discussed openly. Proponents of instant disclosure claim it increases public awareness, presses vendors to issue patches quickly, and improves the quality of software over time. However, many believe that the disclosure of vulnerabilities, especially without a patch, is dangerous for it leaves users defenseless against attackers. Richard Clarke, President

Bush's former special advisor for cyberspace security, said: "It is irresponsible and ... extremely damaging to release information before the patch is out." (Clarke, 2002).

While Bugtraq tends to favor full and quick disclosure, organizations like CERT follow a more cautious approach. After learning of a vulnerability, CERT contacts the vendor(s) and provides a time window to patch the vulnerability; the de facto policy is to give vendors 45 days. After that, the vulnerability is publicly disclosed. Other organizations have proposed their own policies. For example, OIS, which represents a consortium of 11 software vendors, suggests a 30 day window.[2] In addition, firms such as iDefense and 3Com/Tippingpoint buy vulnerability information from users on behalf of their clients, an arrangement which may be socially harmful absent appropriate vulnerability disclosure guidelines (Kannan and Telang, 2005).

1.2 Research Questions

A lack of consensus on the appropriate disclosure policy necessitates a conceptual framework to analyze and guide disclosure policy.[3] Thus, the goal of this paper is, (i) to develop a model of the optimal policy for vulnerability disclosure, which provides actionable recommendations to the policy maker, and (ii) to analyze how the optimal policy is conditioned by various factors. In particular, we examine how long a vendor should be allowed to keep a vulnerability secret (henceforth the "protected period") to optimally balance the need to protect users while providing vendors with incentives to develop a patch expeditiously. Our model can be used to analyze alternatives and to suggest improvements in policies of entities, such as CERT, acting on behalf of society at large.

The optimal disclosure policy depends on the behavior of vendors, potential attackers, and users. In this paper, a vendor is assumed to minimize costs, and hence its choice of when to deliver the patch minimizes the sum of the cost of developing a patch and the portion of the user losses it internalizes. Developing a patch early is costly. However, developing a patch late is costly for the users because attackers can find the vulnerability on their own, and this probability is increasing in time. Thus the vendor trades-off these costs. However, since the

[2] Organization for Internet Safety members include @stake, BindView, Caldera International (The SCO Group), Foundstone, Guardent, ISS, Microsoft, NAI, Oracle, SGI, and Symantec. For details, see http://www.oisafty.org

[3] See also the "Full Disclosure Debate Bibliography", http://www.wildernesscoast.org/bib/disclosure-by-date.html (Accessed, August 24, 2005). See also Preston and Lofton (2002).

vendor does not internalize all customer losses, it has insufficient incentives to produce the patch expeditiously. The social planner can potentially influence the vendor decision by credibly threatening to disclose the vulnerability information after a protected period (thereby making it available to attackers as well). Thus, the social planner chooses the protected period by trading-off customer losses due to disclosure against the benefits of an earlier patch from the vendor (which also reduces customer loss). One key result is that the vendor is more responsive to disclosure if it internalizes more customer loss. However, even when the vendor internalizes a small fraction of the customer loss, an optimal disclosure policy can generate significant social benefits. More interestingly, the social planner can achieve the first best outcome even if the vendor does not internalize customer loss fully.

While our set up is general, we make simplifying assumptions for tractability. However, we show in Section 4.3 that our results are robust to various extensions. In extensions of the basic model, we allow the users' patching rate to vary with the quality of the patch and we allow the vendor to choose the quality of the patch. We show that giving more time to the vendor does not always result in a higher quality patch. In another extension, we allow some users (*smart users*) to defend themselves by applying work-arounds instead of waiting for a patch, if informed of the vulnerability. We show that the social planner can be more aggressive and disclose the vulnerability early, even if the vendor internalizes very little of the customer loss. However, for intermediate values of the proportion of smart users, the presence of smart users increases the social loss.

In section 2, we review the relevant literature. We present the basic set up and assumptions in section 3 and the vendor's decision and the choice of the socially optimal protected period in section 4. Section 4.4 analyzes the case where the speed with which users apply the patch is a function of the quality of the patch. Section 5 extends the model to allow the users to implement work-arounds, and section 6 summarizes and concludes.

2 Prior Literature

This paper contributes to the emerging literature on the economic and policy aspects of cyber security, hitherto a near exclusive domain of computer scientists and technologists (Gordon

and Loeb, 2002). Typical empirical work in this domain has been devoted to trend analysis of vulnerabilities. Arbaugh, Browne, McHugh and Fithen (2001) show that the number of the cumulative incidences (C) follow specific trend over time (M). In particular, they estimate $C = \alpha + \beta\sqrt{M}$. Arbaugh, Fithen and McHugh (2000) propose a life-cycle model of a vulnerability and show that the number of attacks exhibit a bell shaped curve over time since the discovery of the vulnerability. Arora, Nandkumar and Telang (2006) find that disclosing a vulnerability leads to more attacks, and the number of attacks are higher if the patch is not available at the time of disclosure. Arora, Krishnan, Telang and Yang (2005a) find that early disclosure prompts vendors to release patches more quickly.

Some recent papers analyze economic issues related to vulnerability disclosure. Kannan and Telang (2005) show that a market for software vulnerability would lower social welfare if buyers choose to disclose the vulnerabilities they buy. August and Tunca (2005) analyze how unpatched users exert externalities on patched users, and show that the presence of the externality affects the vendors' incentives to improve network security. We also explore such externalities by allowing a fraction of users to implement work-arounds and impose the externalities on the users that lack this ability. Arora, Caulkins and Telang (2005) develop an analytical model where the possibility of patching a software product after it has been released creates incentives for the vendors to rush to the market with buggier products, especially in larger markets. Cavusoglu et al. (2005) present a model of risk sharing between the vendor and software users where the risk arises due to vulnerabilities. These papers do not deal with the issue of disclosure directly.

Choi, Fershtman and Gandal (2005) present a model in which the vendor chooses to disclose vulnerability information along with a patch when a vulnerability is discovered. They show that the vendor may not disclose the vulnerability information even when it is socially optimal to do so. They do not model the threat of disclosure. Png, Tang and Wang (2006) model a game between users and attackers. They show that externalities cause users to underinvest in security and suggest policy measures to remedy the problem. Nizovtsev and Thursby (2007) model the incentives of benign users to disclose software vulnerabilities through an open public forum, whereas in our model, a benign user only contacts CERT which then chooses the disclosure window. Cavusoglu et al. (2004) also analyze the question of vulnerability disclosure. However, their operationalization of social cost differs from ours. Thus, unlike in our model, they find

that vendors may release the patch before the socially optimal time.

3 Basic Set-up and Assumptions

There are four participants in our model – a social planner, a vendor, representative users (customers of the vendor's products) and attackers. Customers' and attackers' behavior is exogenously fixed and we focus on the decisions of the social planner and the vendor. We model a situation (see Figure 1) where a vulnerability is discovered by a benign discoverer (different from the vendor or attackers) and is reported to a social planner (like CERT) at time '0'.[4] The social planner immediately informs the vendor and sets a protected period, T, after which it commits to publicly disclose this information. The vendor makes a one-time decision on when to release a patch.[5] For simplicity, patch release time, τ, is assumed to be deterministic.

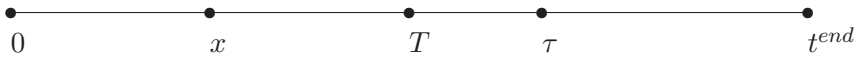


Figure 1: Timeline of Vulnerability

Users incur losses when attackers exploit the vulnerability in their systems. Attackers exploit the vulnerability when they become aware of it and if the customers have not patched. The disclosure policy is binary: either all information is disclosed or none is. To ensure that losses are bounded, we treat the product life-cyle (or version life-cyle), t^{end} as large but finite. Instant disclosure means $T = 0$ while a secrecy policy implies $T > t^{end}$, which essentially means that the information is never disclosed (any action after t^{end} is economically irrelevant in our model). We assume, for now, that customers apply the patch as soon as it is released.

Attackers may discover the vulnerability at time x, where x is a random variable with a $p.d.f.$, $f(x)$. Attackers exploit the vulnerability at time x or at time T, whichever is earlier. Estimates suggest that about 60 percent of the documented vulnerabilities can be exploited almost instantly, either because exploit tools are widely available or because no exploit tool is

[4]Since our goal is to study the socially optimal protected period, we examine the case where the vulnerability is reported to the social planner. If the vendor were to find the vulnerability, it would act as if the protected period were infinite. If the attacker were to find the vulnerability, it would be as if the protected period were 0.

[5]This assumption makes sense if the vendor has to commit resources for patching for a period of time. However, implicitly it requires that the vendor releases a patch as soon as the patch is ready.

needed (Symantec, 2003). Allowing for a deterministic period of exploit-tool development is straight-forward. A key assumption in our model, relaxed in section 5, is that users remain unprotected until a patch is released.

Our model has two stages. In the first stage, the social planner chooses the optimal protected period T^* and in the second stage, the vendor chooses a patch development time, τ^*, in response to T. The social planner sets T taking into account the vendor's response. Accordingly, we first solve the second stage and then solve the first stage. Before we proceed, we list the notation in Table 1 below and outline key assumptions.

τ	patch release time set by the vendor
T	protected period set by the social planner
$V(\tau, T)$	vendor expected cost function
$S(\tau(T), T)$	social planner cost function
q	quality of the patch
$C(\tau, q)$	vendor patch development cost when the quality of the patch is q
$\zeta(z)$	instantaneous loss for unpatched customers at time z after the patch release
$l(y)$	cumulative pre-patch customer loss when exposed to attacks for duration y
$L(\tau(T), T)$	expected pre-patch customer loss
$\tilde{L}(\tau(T), t^{end})$	expected post-patch customer loss
$\mathcal{L}(\tau(T), T, t^{end})$	total (pre and post-patch) expected customer loss
λ	proportion of customer loss internalized by the vendor
$p(z, q)$	proportion of users who have applied the patch of quality q by the elapsed time z since patch release.
$F(x)$	probability of attacker finding the vulnerability by time x
τ^s	socially optimal patch release time (if the social planner could release patch)
τ^∞	patch release time by the vendor under secrecy policy
T^k	kink point in vendor's reaction function w.r.t to T
α	proportion of smart users who can implement work-around if informed by the social planner
w	cost of work-around
$V^w(\tau, T)$	vendor cost function when smart users implement work-around
$S^w(\tau(T), T)$	social planner cost function when smart users implement work-around
τ^w	vendor patching time when smart users implement work-around.
T^w	cut-off point such that for $T \leq T^w$, vendor induces work-around and vice-versa
$\hat{\alpha}$	minimum proportion of smart users such that work-around is socially optimal

Table 1: Notation

3.1 Assumptions

Let $l(y)$ be the cumulative customer loss when users are exposed to an unpatched vulnerability for duration y. The longer users are exposed without a patch, the greater the losses suffered.[6] Even though a given user, once attacked, might not suffer additional losses even as she remains exposed, as the period of exposure increases, the number of users who are likely to be attacked will increase. This is because the number of attackers aware of the vulnerability and in possession of the attack scripts increases. As Arbaugh et al. (2000) note "intrusions increase once the community discovers a vulnerability and the rate of intrusions accelerates as news of the vulnerability spreads to a wider audience". This suggests the following assumption.

A1: $l(y)$ *is increasing and strictly convex in* $y, t^{end} \geq y \geq 0$ *and* $l(0) = 0$.

Customer losses depends upon the gap between when the vulnerability is discovered by attackers (or disclosed to them) and when the patch is released. If the patch is released before disclosure, customers suffer a loss only if an attacker rediscovers the vulnerability prior to the patch. From Figure 1, x is when an attacker finds the vulnerability and τ is when the patch is released. Customers may be attacked between time x and τ, and do not suffer losses beyond t^{end}. Hence, cumulative customer loss is $l(\tau - x)$ (or $l(t^{end} - x)$ if $\tau > t^{end}$). If the patch is released after T and the customers apply the patches instantly, there are two possibilities: First, attackers could find the vulnerability on their own and exploit it for $\tau - x$ periods. Alternatively, at time T, attackers learn about the vulnerability when it is disclosed and exploit it until the patch is released at τ, for $\tau - T$ periods. The probability that attackers discover the vulnerability on their own by time x is given by the *cdf* $F(x)$. Thus, the expected pre-patch customer loss $L(\tau, T)$, which is a function of the protected period T and the patch release time τ, can be written as,

$$L(\tau, T) = \begin{cases} \int_0^{\tau} l(\tau - x) dF(x), \ when \ \tau < T \\ \int_0^{T} l(\tau - x) dF(x) + (1 - F(T)) l(\tau - T), \ when \ \tau \geq T \end{cases} \quad (1)$$

The first part of (1) is the customer loss when a patch is released before the protected period T but an attacker discovers the vulnerability at $x < \tau$, exposing customers to attacks for the duration $\tau - x$. The second part is when the patch is released after disclosure, and attackers

[6] Customer loss also depends on vulnerability and customer specific factors, which we ignore here.

either find it before T and attack for the duration $\tau - x$ or learn of the vulnerability at T when it is publicly disclosed and attack for the duration $\tau - T$.

The customer loss that the vendor internalizes is in the form of either a loss in reputation or a loss in future sales or as customer support costs to which it is contractually obligated. We represent the proportion of customer loss internalized by the vendor as λ and call it the internalization factor. Although vendors in the United States do not face any liability for defects in software products, λ may be interpreted as contractual liability. We assume that $\lambda < 1$.

We use *subscripts* to denote partial derivatives of functions with respect to their arguments, except that we use $l^{'}(\cdot)$ to denote the derivative of $l(\cdot)$. We use *superscripts* to denote particular values of variables.

Even after the vendor releases the patch, customers may not apply the patch instantly, and may continue to incur losses. Since a patch may disclose additional details about the vulnerability, we allow the post-patch losses incurred (by un-patched customers) to differ from those before the patch release. If z is the time elapsed since the patch was released, let $p(z, q)$ denote the proportion of customers that have applied that patch by time z, where q ($q \geq 0$) is the quality of the patch. Since higher quality patches are easier for the customers to download and apply, we assume that

A2: $p_q(z, q) > 0$

The expected cumulative post-patch loss is $\tilde{L}(t^{end} - \tau, q) = \int_0^{t^{end}-\tau} \zeta(z)(1 - p(z, q))dz$, where $\zeta(z)$ is the *instantaneous* post-patch loss for unpatched customers at time z. The total (pre and post-patch) expected customer loss is given by $\mathcal{L}(\tau, T, q)$, and can be written as -

$$\mathcal{L}(\tau, T, q) = \begin{cases} \int_0^{\tau} l(\tau - x)dF(x) \ + \int_0^{t^{end}-\tau} \zeta(z)(1 - p(z, q))dz, & \text{when} \quad \tau \leq T \\ \underbrace{\int_0^{T} l(\tau - x)dF(x) + (1 - F(T))l(\tau - T)}_{L(\tau,T)} + \underbrace{\int_0^{t^{end}-\tau} \zeta(z)(1 - p(z, q))dz}_{\tilde{L}(\tau,q)}; & \text{when} \quad \tau > T \end{cases} \tag{2}$$

Note that $\tilde{L}(\cdot)$ captures the post-patch losses and does not depend on T. Similarly, $L(\tau, T)$ does not depend upon q. We will show in the next section that $L(\cdot)$ is convex. However, since we

require the total expected customer loss to be convex, a sufficient though not necessary condition is that $\tilde{L}(\cdot)$ be convex. We assume -

A3: $\tilde{L}(\tau, q)$ *is strictly convex in* (τ, q)

Concretely, *A3* requires that $\frac{\zeta'(t^{end}-\tau)}{\zeta(t^{end}-\tau)} > \frac{p'(t^{end}-\tau,q)}{1-p(t^{end}-\tau,q)}$, i.e., the growth rate of instantaneous losses for unpatched users is greater than the rate at which unpatched users decline. Convexity of $\tilde{L}(\tau, q)$ also requires that $p_{qq}(z, q) < 0$, or that the share of patched users increases at a diminishing rate as quality increases.[7]

A4: (i) $C_\tau(\tau, q) < 0$, $C_{\tau\tau}(\tau, q) > 0$, $C(0, q) = \infty$, *and* $C_\tau(0, q) = -\infty$. *(ii)* $C(\tau, q)$ *is strictly convex in* (τ, q), $C_q(\tau, q) > 0$, $C_{\tau q}(\tau, q) < 0$

We assume that patch development cost $C(\tau, q)$ is decreasing and convex in τ. The more resources the vendor allocates to develop a patch, shorter is the time taken to patch. However, the benefits of delay are decreasing. We need $C(0, q) = \infty$, and $C_\tau(0, q) = -\infty$ for technical convenience. The second part of *A4* deals with the impact of quality on patch development cost. It is likely that accelerating patch development and increasing the quality of the patch draw upon scarce resources. Hence, we assume that the shorter is the time allocated for patch development, the costlier is the patch quality and the marginal cost with respect to quality is also higher.

Finally, we need the following condition for tractability

$$A5\text{: } \min_{\tau} \left\{ C(\tau) + \lambda \int_0^{\tau} l(\tau - x) dF(x) \right\} + \lambda \int_0^{t^{end}-\tau} \zeta(z)(1 - p(z, q)) dz < \lambda \int_0^{t^{end}} l(t^{end} - x) dF(x)$$

The inequality ensures that, left to itself, the vendor would voluntarily develop a patch before the end of the life-cycle, and moreover, this is socially desirable as well. The left hand side of the inequality is the vendor loss if the vendor releases a patch. The right-hand side of the inequality is the vendor loss when the vendor does not develop a patch. Essentially, we require t^{end} to be long, or λ to be large. Note that if post-patch losses are large, *A5* may not hold.

The assumption is empirically sensible in that the vast majority of the vulnerabilities handled by CERT are patched, which suggests that the relative to the typical time taken to patch, the product life-cycles are long. However, t^{end} would be small, if, for instance, a vulnerability is

[7] There is one final restriction on $p(z, q)$ and $\zeta(z)$ of the requirement that the relevant matrix of second order derivatives be positive definite. There is no meaningful economic interpretation of this restriction.

discovered shortly before a new version of the product is scheduled to be released. In this case, optimal disclosure policy is trivial – one should simply wait for the vulnerability to be addressed in the new version.[8] If the assumption were violated, then every interior solution for the vendor's problem on the optimal patch release time would have to be compared to the payoff from not releasing the patch. We discuss this further in section 4.4, along with other extensions.

3.2 Vendor's Objective Function

Given a commitment by the social planner to a protected period T, and customer patching rate $p(z, q)$, the vendor chooses a patch development time τ and the patch quality q to minimize its costs. The social planner's commitment is assumed to be credible because either this is a repeated game (which we do not explicitly model) or the social planner is concerned about its reputation, or both. The vendor's expected cost function is

$$V(\tau, q; T) = C(\tau, q) + \lambda \mathcal{L}(\tau, T, q)$$

This cost function has two terms. The first term is the cost of patch development, $C(\tau, q)$ and the second is the portion of expected user loss internalized by the vendor, $\lambda \mathcal{L}(\tau, T, q)$.

4 Model and Analysis

We are now ready to analyze, (i) the vendor's decision to release the patch for a given protected period, and (ii) the socially optimal protected period. We begin by analyzing the case where customer apply the patches instantly ($p(0, q) = 1$) so that post-patch losses are zero ($\tilde{L}(\cdot) = 0$) and hence patch quality is exogenously set at some level q. We relax these assumption later in section 4.4.

First, we outline the vendor's decision and then we analyze the optimal disclosure policy and how various factors condition the policy.

[8] If the vendor does not fix it in the new release, then it is as if the clock were restarted - i.e., t^{end} is large.

4.1 Vendor's Decision

The expected user loss is as given in equation (1) and the vendor's objective function is [9]

$$V(\tau;T) = C(\tau) + \lambda L(\tau,T) \tag{3}$$

From equation (1), $L(\tau,T)$ is continuous everywhere, and is differentiable everywhere except perhaps at $\tau = T$. Lemma 1 shows that it is also convex in τ. Since $C(\tau)$ is also convex in τ, that vendor cost function $V(\tau,T)$ is strictly convex in τ as well. Lemma 1 shows that there always exists a unique optimal τ^* for a given T. (All proofs are in the online appendix)

Lemma 1. *The expected customer loss function $L(\tau,T)$ is strictly convex in patch development time τ. For any given T, there exists a unique optimal patch development time τ^*.*

However, an optimal τ^* may be at a kink point because the vendor cost function is not differentiable in τ at $\tau = T$ (unless $l'(0) = 0$). In the following, we analyze the possibility of such a kink point and its properties.[10] Figure 2 shows the two components of the vendor's objective function and highlights such a kink point.

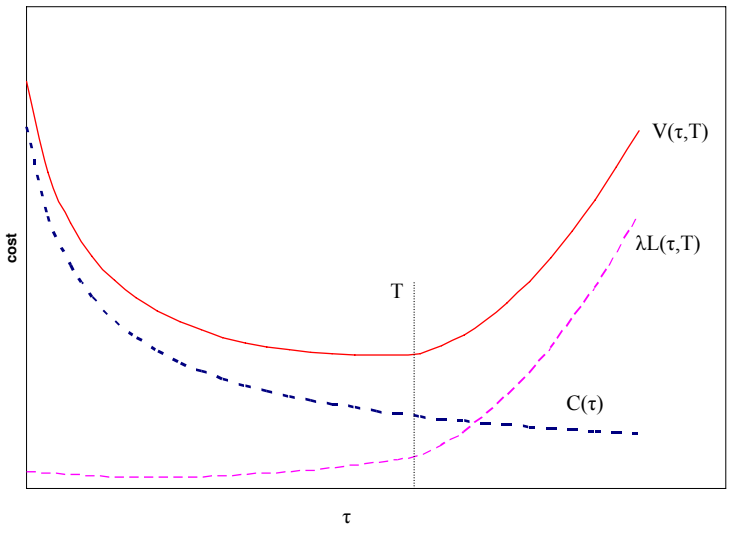


Figure 2: Kink Point in Vendor Cost function

At any such kink point, the left hand side derivative of V w.r.t τ is $V_\tau^-(T)\Big|_{\tau=T} = \Big(C_\tau(\tau) +$

[9] For notational clarity, we suppress q from the cost function $C(\cdot)$ and the loss function $L(\cdot)$ when q is not a decision variable

[10] We are grateful to a reviewer for alerting us to the existence and importance of the kink point.

$\lambda \int_0^T l^{'}(\tau - x)dF(x)\Big)$. The right hand side derivative is given by

$$V_\tau^+(T)\Big|_{\tau=T} = C_\tau(T) + \lambda \int_0^T l^{'}(T-x)dF(x) + \lambda(1-F(T))l^{'}(0).$$

Since $l'(0) > 0$, the right hand side derivative is bigger than the left hand side derivative. The economic interpretation is that a kink will exist if attackers can exploit even a very small gap between the disclosure and the patch. Empirical results in Arora, Nandkumar and Telang (2006) show that typically the disclosure is followed by a spike in attacks, which subside following the release of the patch, but only with some delay. This suggests that $l^{'}(0) > 0$ is plausible and perhaps even likely.

At a "kink" equilibrium $\tau^* = T$, the right hand side derivative must be non-negative and the left hand side derivative must be negative. Define T^k such that $V_\tau^+(T)\Big|_{\tau=T^k} = 0$. Note that the second derivative of the right hand side w.r.t. T is equal to $C_{\tau\tau}(T) + \lambda \int_0^T l^{''}(T-x)dF(x) > 0$. Thus, the right hand side derivative is increasing in T. Since, $V_\tau^+(T)\Big|_{\tau=T^k} = 0$ for $T < T^k$, $V_\tau^+(T)\Big|_{\tau=T} < 0$. Hence, for all $T < T^k, \tau^* \neq T$.

Define the socially optimal patching time, τ^s, i.e., the patch release time that minimizes the unconstrained social cost (i.e., $\lambda = 1$) as:

$$\tau^s = \arg\min_\tau \Big\{ C(\tau) + \int_0^\tau l(\tau - x)dF(x) \Big\}. \tag{4}$$

Let τ^∞ denote the optimal patch development time given secrecy policy (i.e., $\tau^\infty = \arg\min_\tau \ C(\tau) + \lambda \int_0^\tau l(\tau - x)dF(x)$).

Lemma 2. *T^k, τ^s and τ^∞ exist.*

The difference between T^k and τ^∞ (the time when the vendor will release the patch if left alone) is due to the impact of disclosure. If even a vanishingly small delay in releasing the patch (after disclosure) leads to a loss (so that $l^{'}(0) > 0$), then $T^k < \tau^\infty$. Any increase in $l^{'}(0)$ will decrease T^k and increase the gap between T^k and τ^∞. A vulnerability for which no exploit code is needed, which can be remotely exploited, or where attackers have large numbers of "zombie" computers under their control, are likely to be characterized by higher values of $l^{'}(0)$ and lower values of T^k. Correspondingly, T^k plays an important role in our analysis because we show

below that even if $\lambda < 1$, the social planner may be able to achieve the first best outcome by setting $T \geq T^k$. Thus the factors that affect T^k have important implications for the disclosure policy. To further characterize T^k, we show below that T^k decreases when the internalization factor (λ) increases or the probability of attackers finding the vulnerability increases (a first order stochastic dominant shift in $F(x)$).[11]

Lemma 3. *(i) T^k is decreasing in λ. (ii) If $G(x)$ is a c.d.f. defined over the same domain as $F(x)$ such that $G(x) \geq F(x)$ then T^k corresponding to $G(x)$ is smaller than the T^k corresponding to $F(x)$.*

Note that T^k is the smallest protected period such that the vendor releases the patch within the protected period. A higher λ and an increased probability of the vulnerability being discovered by attackers imply higher marginal cost to the vendor of delaying the patch for a given T.

We are now ready to define the optimal vendor behavior for a given T. Vendor's best response function is the implicit function of $V_\tau(\tau, T) = 0$.

Theorem 1. *For $T \in [0, T^k)$, the vendor patches after disclosure i.e., $T < \tau^* < T^k$ and the slope of $\tau^*(T)$ is strictly less than one. For $T \in [T^k, \tau^\infty]$, the vendor patches at T i.e., $\tau^* = T$, and hence slope of $\tau^*(T)$ is equal to one. For $T \in [\tau^\infty, t^{end}]$, the vendor patches at τ^∞, and hence the slope of $\tau^*(T)$ is equal to zero.*

Theorem 1 shows, as many full disclosure proponents believe, that reducing T results in the vendor releasing the patch more quickly (i.e., $\partial\tau^*/\partial T > 0$, but only if $T < \tau^\infty$). Further, for any $T < T^k$, τ^* is an interior point and the vendor patches after the protected period elapses; for $\tau^\infty \geq T \geq T^k$, $\tau^* = T$ so that T^k possibly marks a kink in the best response function $\tau(T)$. Figure 3 shows τ^* as a function of T: τ^* increases in T until τ^∞ and is flat after that. Moreover, after T^k is reached, $\tau^* = T$. Also notice that because $\tau > 0$ when $T = 0$ and $\partial\tau^*/\partial T < 1$, the gap between τ^* and T shrinks.

[11] Note that a "kink" can arise in other ways as well. Punitive disclosure, such as where a vendor releasing a patch after the disclosure is publicly "named and shamed", will not affect T^k. It will, however, increase the possibility of the "kink" solution by making the vendor objective function discontinuous at the disclosure point. Such punitive disclosure punishes the vendor without imposing losses on customers, thereby increasing the potency of the disclosure policy. (See also Section 4.2.1 below.)

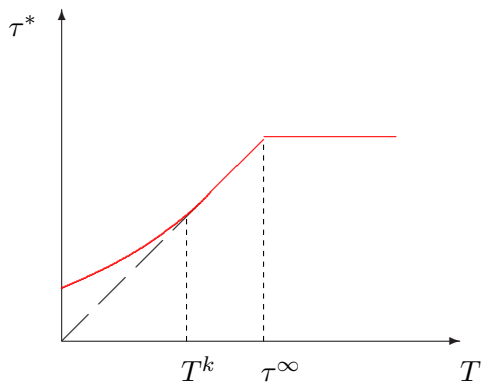


Figure 3: Patch development time τ as function of Protected Period T

The lower is λ, the slower is the vendor to patch. However, this is only true when $T < T^k$, after which the vendor chooses to patch at T regardless of λ. (Recall from Lemma 3 that T^k itself is decreasing in λ.) This is formalized in Corollary 1.

Corollary 1. *A higher internalization factor implies an earlier patch, $\partial\tau^*/\partial\lambda < 0$ for any $T < T^k$. When $T \geq T^k$, $\partial\tau^*/\partial\lambda = 0$.*

4.2 The Social Planner's Decision: Optimal Disclosure Policy

The social planner chooses optimal T^* to minimize total social cost, $S(T)$, taking into account the vendor's best response function $\tau(T)$. The social cost is given by

$$S(T) = C(\tau(T)) + L(\tau(T), T) \tag{5}$$

Social cost differs from vendor cost in that the former includes the entire expected user loss, whereas the latter includes only a fraction λ of the expected user loss. When the vendor internalizes only a portion of customer loss, i.e. $\lambda \in (0, 1)$, the vendor's incentives and the social planner's incentives are not aligned.

We can now derive the socially optimal policy. We assume that $S(T)$ admits only a single minimum (see the online appendix for the sufficient condition for a unique T.) A variety of functional forms, including the exponential and quadratic loss functions, yield a single minima.

Recall that τ^s is the socially optimal time to deliver the patch, i.e., the time a vendor would release the patch on its own if it internalized the entire customer loss. Since the vendor does

not internalize the entire customer loss, absent a threat of disclosure, the vendor delivers the patch after τ^s. Given a protected period T, the vendor will deliver the patch after T if $T < T^k$, and exactly at time T for $T \geq T^k$. Hence, as long as $\tau^s \geq T^k$, the social planner can choose $T = \tau^s$ and the vendor would patch exactly at τ^s. However for $\tau^s < T^k$, the socially optimal T, denoted by T^*, lies between τ^s and T^k as shown in the next proposition and in Figure 4.

Theorem 2. *When $\tau^s < T^k$, the socially optimal protected period T^* is bounded within (τ^s, T^k) i.e. $\tau^s < T^* \leq \tau(T^*) \leq T^k$. When $\tau^s \geq T^k$, $T^* = \tau(T^*) = \tau^s$.*

Clearly, whenever $\tau^s \geq T^k$, the social planner can achieve the socially best outcome even though $\lambda < 1$. Thus, disclosure policy can be an effective and potent tool even though the social planner can affect vendor behavior only indirectly.

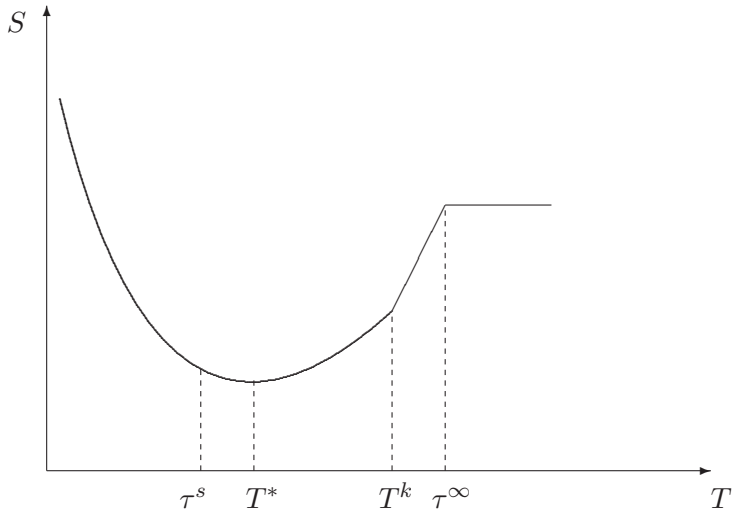


Figure 4: Social Cost as Function of T

4.2.1 Factors Affecting the Optimal Disclosure Policy

An increase in λ will cause the vendor to release the patch earlier because the vendor internalizes a larger fraction of customer losses. A higher λ also implies that the vendor is more sensitive to disclosure. Hence, the social planner will optimally reduce T^*. The optimal protected period, T^*, decreases with λ until λ reaches λ_0 and is constant thereafter. The intuition is that T^k is decreasing in λ (from Lemma 3), so that for high enough λ, $\tau^s \geq T^k$ holds. If $\tau^s \geq T^k$, the social planner can choose $T = \tau^s$ and the vendor would patch exactly at the socially optimal time τ^s. This is formalized in Theorem 3.

Theorem 3. *There exists a* $\lambda^0 \in (0,1)$ *such that for* $\lambda \geq \lambda^0$, $\tau^s \geq T^k$, *and* $\tau^* = T^* = \tau^s$ *and* T^* *is independent of* λ. *For* $\lambda < \lambda^0$, $\tau^s < T^k$ *and the socially optimal protected period,* T^*, *is decreasing in* λ.

When $\tau > T$, there is a period when customers are exposed. The gap between T and τ falls with λ , and τ becomes more responsive to T. In short, the social planner has greater leverage with the vendor when λ is higher. This result is important; disclosure policy relies upon the sensitivity of the vendor to customer losses. When the vendor is more sensitive to customer losses (higher λ), the social planner has more leverage in forcing vendors to release the patch on time. In this respect, our result is counter-intuitive: One expects greater alignment between the firm's objective function and social welfare to weaken the need for regulation, but here the reverse is true, because a greater alignment between the two also increases the efficacy of regulation. However, our numerical analysis (see Section 4.5) suggests that even for low λ, suitably chosen disclosure can generate significant social benefits.

There are two ways to a higher λ. One is when customers are able to punish the vendor by switching to a competing product. We conjecture that competition increases λ. Second, larger users are more likely to contract with vendors about the patching support. Thus, vendors whose market base consists of large users will have higher λ.

4.3 Robustness of the model

Though not analyzed here, the social planner plausibly has a spectrum of disclosure possibilities. For instance, the social planner could issue a general warning that a particular product is insecure, which could hurt the vendor without imposing large losses on users. In terms of our model, one would add a term to the vendor's cost, so the modified vendor cost is $V(\cdot) = C(\tau) + \lambda L(\tau, T) + \psi(\tau, T)$ and the social cost is $S(\cdot) = C(\tau) + L(\tau, T)$, where $\psi(\tau, T)$ captures the damage suffered by the vendor not due to the customer loss. We expect that $\psi(\tau, T) = 0$ for $\tau \leq T$ and is positive and increasing in $\tau - T$. Including such a term will cause a discontinuity in the cost function at $\tau = T$, and cause cost function to be concave around $\tau = T$, thereby increasing the likelihood of the vendor patching at T. This modification can also be used to analyze the case where the vendor suffers losses that are not included in the social cost function

(e.g., where the vendor loses some customers to other vendors by releasing the patch late).

It is also plausible that the post disclosure loss function differs from $l(\cdot)$. If we let $m(\tau - T)$ represent the post disclosure losses, then as long as $m(0) = 0$, $m(\cdot)$ is convex, and $V_\tau^+(T)\Big|_{\tau=T}$ is monotonically increasing in T, our basic results should continue to hold.[12] Disclosure by the social planner may also increase λ by forcing the vendor to acknowledge responsibility. This is as if the vendor's (but not social) post-patch loss function were $m(\cdot) = \sigma l(\cdot)$, where $\sigma > 1$.

We have assumed through *A5* that despite a finite product life-cycle, "not patching" is not an optimal choice. The online appendix shows that if assumption *A5* is violated, then there exists a threshold, $T^{NP}(t^{end})$, such that $T > T^{NP}$ implies that the vendor does not patch. It is intuitive that T^{NP} increases with t^{end}. Further, since the vendor does not internalize the entire customer loss, there is a range of values of t^{end} such that the social planner's choice of the protected period is constrained by the need to get the vendor to patch. The unconstrained choice of the protected period would result in the vendor not releasing the patch. When the constraint binds, the protected period is shorter than when the constraint does not bind. Finally, there is a threshold value of t^{end} (possibly zero) below which the social planner chooses *secrecy* and the vendor does not develop a patch.

4.4 Customers do not Patch Instantly

Thus far we have assumed that all customers patch as soon as the patch is available (or $p(0, q) = 1$). The .NET passport vulnerability is a good example. A fix on the server side stops the invasion and customers need no patch (Infoworld, 2003). However, many vulnerabilities require that customers download and apply patches. Not all customers apply patches immediately upon release (Rescorla 2003). Six months after the DDOS attacks that paralyzed several high-profile Internet sites, more than 100,000 machines were still unpatched and vulnerable (InternetNews.com, 2000).

Users may not patch their systems immediately because (i) it takes time to find out about the patch, (ii) applying the patch may take time and may be difficult for unskilled users, (iii) poor quality patches may themselves create new problems. For example, the initial Microsoft

[12] If $m(0) > 0$ and $m''(\cdot) < 0$, the cost functions may be non-convex, implying that patching soon after disclosure is sub-optimal, and thus we should see many cases of patching coinciding with disclosure.

patch for a vulnerability CVE-2001-0016 disabled many updates of service pack 2 of Windows NT, making the patched system even more vulnerable to attacks (Beatie et. al. 2002). The following statement clearly points to how the patch quality can affect the patch uptake:

"About 95 percent of exploits occur after bulletins and patches are put out....(T)he reason the exploit is effective is because the patch uptake is too low. The reason the patch uptake is too low is it's too hard to patch, and the quality of the patch is not consistent enough that people can feel safe patching right away.", Microsoft chief security strategist Scott Charney.[13]

The speed with which customers apply patches depends on two factors: the time elapsed since the patch is released (z) and the quality of the patch (q). Quality can be interpreted to include those features as well that make it easier for users to download and install the patches.

4.4.1 Quality of Patch is Exogenous

We first consider the case when quality is exogenously fixed at some level q. The expected loss function is of the form given in equation (2). Thus the vendor's objective function is

$$\begin{aligned} V(\tau, T) &= C(\tau) + \lambda \mathcal{L}(\tau, T; q) \\ &= C(\tau) + \lambda L(\tau, T) + \lambda \tilde{L}(\tau; q) \end{aligned}$$

One can show that Theorems 1-3 continue to hold.[14]

Since post-patch losses fall as the patch is delayed, τ^*, τ^s, T^k, and τ^∞ are all larger than when patches are installed instantaneously. The vendor therefore delays the patch. If $\frac{d\tau}{dT}$ is (weakly) smaller for a given T in the presence of post-patch losses, then the social planner also optimally allows more time (higher T^*). Indeed, in the extreme case, where users never patch, not developing a patch is both privately and socially optimal.

Theorem 4. *When users do not patch instantly, (i) T^k is higher, (ii) the vendor slows patch development, and (iii) if $\frac{d\tau}{dT}$ in the presence of post-patch loss is no higher than $\frac{d\tau}{dT}$ in the absence of post-patch loss, the social planner allows more time before disclosure.*

The theorem formalizes the intuition that the vendor and the social planner are both less

[13] http://entmag.com/news/article.asp?EditorialsID=5833; (accessed September 1, 2006)

[14] The proofs are analogous to those in the previous section and are omitted here. They are available from the authors upon request.

aggressive in the presence of post-patch losses. Moreover, we are less likely to see cases where the patch coincides with the disclosure, since T^k moves to the right.

The rate of user implementation of patches can be low. In a case study of the OpenSSL Remote Buffer Overflow vulnerability (exploited by the notorious slapper worm), Rescorla (2003) reports that 60 percent of the investigated servers did not patch even after two weeks of the release of the patch. Our results imply that in such cases, vendors should be given more time to develop patches.

4.4.2 Quality of Patch is Endogenous

A common argument for giving vendors more time is that it facilitates the development of higher quality patches. High quality patches are those that customers trust, are easier to download and apply, and less likely to cause problems later. Simply put, a higher quality patch increases the uptake of patches. When the vendor can choose both q and τ, the vendor's cost function is:

$$\begin{aligned} V(\tau, q) &= C(\tau, q) + \lambda \mathcal{L}(\tau, T, q) \\ &= C(\tau, q) + \lambda L(\tau, T) + \lambda \tilde{L}(\tau, q) \end{aligned}$$

Consider the impact of increasing the protected period, T. An increase in T will lead to an increase in τ. The impact of higher τ on q depends on the sign of $V_{\tau q}(\cdot)$. It is easy to see from equation (2) that $L_{\tau q}(\cdot) = 0$ and $\tilde{L}_{\tau q}(\cdot) > 0$ so that $V_{\tau q}(\cdot) = C_{\tau q}(\cdot) + \lambda \tilde{L}_{\tau q}(\cdot)$ is of indeterminate sign. If the development cost effect, $C_{\tau q}(\cdot)$, dominates, then $V_{\tau q}(\cdot) < 0$, and the vendor will increase q. However, if the post-patch loss effect, $\lambda \tilde{L}_{\tau q}(\cdot)$, dominates, then $V_{\tau q}(\cdot) > 0$, and the vendor will decrease q.

Theorem 5. *When patch quality is endogenous, the optimal time for patch development τ^* is increasing in disclosure time T, i.e., $\frac{d\tau^*}{dT} > 0$. If $V_{\tau q}(\cdot) \leq 0$, patch quality is increasing in T, ($\frac{\partial q}{\partial T} \geq 0$) and if $V_{\tau q}(\cdot) > 0$, patch quality is decreasing in T, ($\frac{\partial q}{\partial T} \leq 0$) .*

Theorem 5 appears counter-intuitive. It is commonly believed that providing more time to vendors will improve the quality of the patch. However this view is only partially true. An increase in τ (due to an increase in T) has two opposing effects on the marginal benefit of quality. On the one hand, the marginal cost of quality falls since $C_q(\cdot)$ falls with τ, but on the

other hand, increasing τ also reduces the marginal benefit of quality because a delayed patch also reduces the marginal benefit of quality in mitigating post-patch loss. When the post-patch loss effect dominates, increasing the protected period reduces the patch quality.

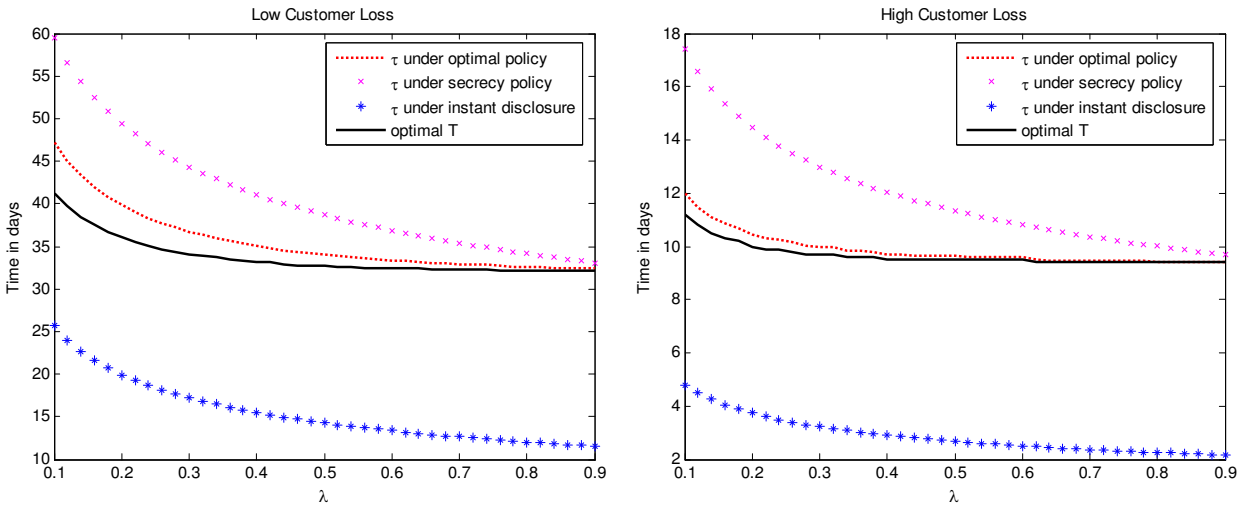
What happens to the optimal protected period when quality is set endogenously? In general, the protected period can be higher or lower with endogenous quality. However, if the post-patch loss effect dominates the patch development cost effect, then the optimal protected period is lower when quality is endogenous. When the patching cost effect dominates, the impact on T is ambiguous: reducing T hastens the patch but lowers its quality (see online appendix for proof).

4.5 Numerical Analysis

To gain more insight into the optimal disclosure policy, we performed numerical simulation, with the following functional forms: $C(\tau) = 50000/\tau^{0.75}$, $l(y) = 25y^2$, $F(x)$ uniform over $[0, z]$ and we let z vary from 200 to 300 days. We let λ vary from 0.1 to 0.9. Since the effectiveness of the policy depends on the relative magnitudes of the customer loss and the patching cost, we repeat the simulation with customer loss $l(\cdot) = 2500y^2$, keeping the rest same.

Given the paucity of information on the patching costs and the customer loss functions, we chose parameter values that match observed patch release behavior. In a related empirical study (Arora, Krishnan, Telang, and Yang 2005), we estimate that instant disclosure leads to a patch in 31 days and secrecy leads to patch in about 61 days, on average. With the chosen forms, for $\lambda = 0.1$, under instant disclosure, the patch release time τ varies from about 25 days (low customer loss) to 5 days (high customer loss), and under secrecy, τ correspondingly varies from 58 days to 18 days. We calculate the optimal patch release times under instant disclosure, under secrecy, and under the optimal protected period. To avoid clutter, we plot these values for only $z = 300$; lower values of z make both the vendor and the social planner more aggressive.

In figure 5, λ is on the x-axis and time (in days) is on the y-axis. Note that the patch takes the longest time under secrecy and the shortest under instant disclosure. As expected, the patch release time falls with λ. Also the difference between the patch release time and the protected period (the middle two lines) shrinks with λ. Beyond $\lambda = 0.5$, the difference between T and τ is small. In short, when the vendor is responsive enough, disclosure is very effective in forcing the vendor to release the patch in time. But even for small λ, the optimum disclosure policy is



(a) fig 5(a)

(b) fig 5(b)

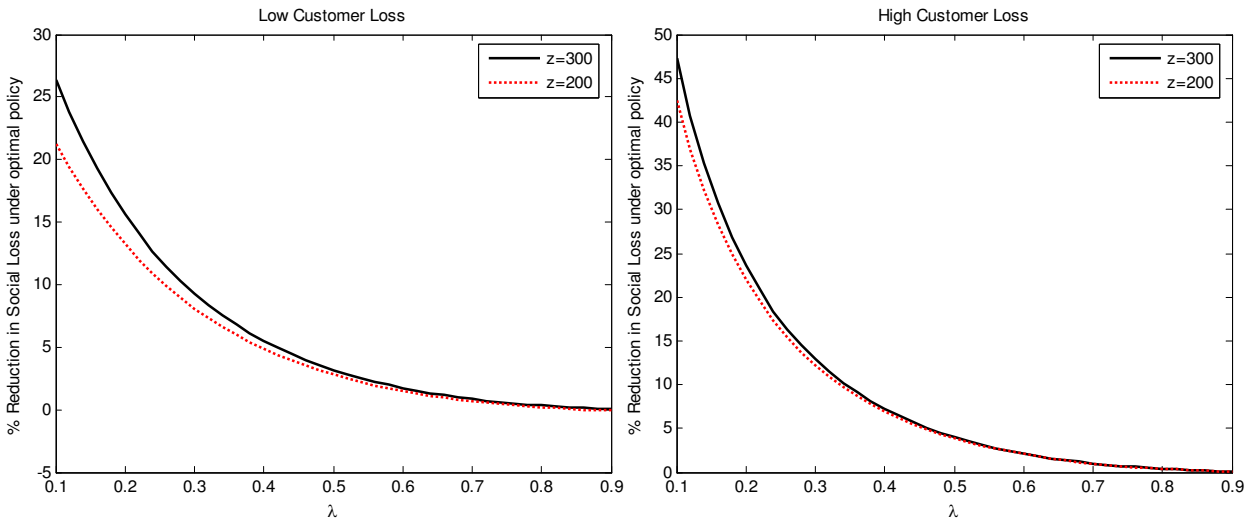
Figure 5: τ and T by λ

effective in reducing the vendor's patch release time (from 60 to 48 days in figure 5(a) and from 17 to 12 days in figure 5(b)). When the customer loss is high (figure 5(b)), τ is lower as is the protected period, T, and the difference between the two is smaller.

Turning to the social losses, we find that the instant disclosure performs poorly compared with either the secrecy or the optimal policy, with social losses 300-400 percent greater under the instant disclosure than under the optimal policy. Accordingly, in figures 6(a) and 6(b) we only plot the percentage reduction in the social loss under the optimal policy compared to the secrecy policy.

Note that the difference between z =200 and z =300 is small. For low values of λ, the optimal policy generates benefits (social loss reductions) of about 25-30 percent in figure 6(a) and about 40-45 percent for figure 6(b). To put this in perspective, if we take \$1 billion as a conservative lower bound for the losses arising from information security breaches due to inappropriate disclosure, the implied savings from the optimal disclosure policy range from \$250 million to \$450 million. More importantly, by offering a credible alternative to the instant disclosure, CERT also potentially reduces some cases of instant disclosure, which are significantly more costly.

In both figures, as λ increases, secrecy (no disclosure) becomes almost as effective as optimal policy even though the gap between τ and T remains large. More interestingly, we showed earlier



(a) fig 6(a)

(b) fig 6(b)

Figure 6: percentage reduction in social loss under the optimal policy

that with higher λ, the protected period is short. However, even when λ is small and the social planner can not be very aggressive, the social benefits of optimal policy are high. Thus policy has *bite* in terms of reducing the social losses when λ is small even though the vendor is not as responsive.

Note that the optimal policy matters more when the customer losses (relative to the patching costs) are higher (figure 6(b) vs. figure 6(a)). Higher customer loss makes the vendor more responsive but higher customer loss also raises the social loss. Since the vendor internalizes only a fraction of the customer loss, the socially optimal policy relative to the secrecy policy is more effective when the customer loss is higher. One interpretation of this result is in terms of market size. An increase in the market size is equivalent to an increase in the customer loss relative to the patch development cost, since the latter is roughly invariant to changes in market size. Arora, Caulkins and Telang (2005) show that vendors patch more expeditiously in larger markets. The numerical simulations reported here confirm that finding and also show that if the vendor internalizes only a fraction of customer loss, the socially optimal policy relative to the secrecy is more effective in larger markets.

We also experimented with skewed distributions for $F(\cdot)$. When $F(\cdot)$ is skewed to the right (i.e. the hackers have a higher chance of finding vulnerability later) then gains from the optimal

policy are even higher. Left skewed distributions, on the other hand, produced lower gains.

Finally, for λ larger than 50 percent, the vendor's response under the secrecy is very similar to that under the optimal disclosure policy. As long as the vendor internalizes at least 50 percent of the user loss, the vendor's decision to release a patch on its own creates the additional social costs of the order of only 5 percent. We also explored kink solutions and they lead to similar predictions.

Empirical results agree with our model. Arora, Krishnan, Telang, and Yang (2006a) find that vendors typically patch after disclosure and early disclosure leads to an early patch. In Figure 7, disclosure is on the x-axis and the patch release time is on the y-axis. As can be seen, early disclosure leads to a quicker patch. Moreover, except for when T is small, the gap between T and τ shrinks as T increases.

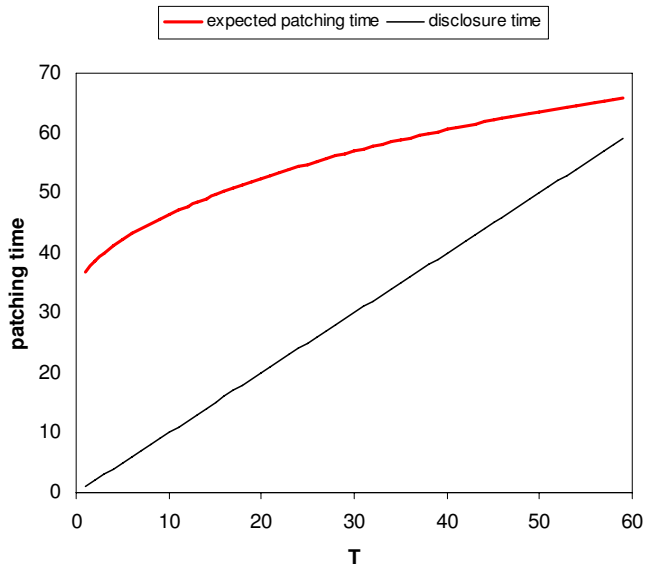


Figure 7: Patching time Vs Disclosure Time, *Data source*: Arora et al. (2006a)

Figure 7 uses actual disclosure. Since actual disclosure times may differ from the promised protected period, Arora, Forman, Nandkumar and Telang (2006b) exploit the variation in the number of vendors with a common vulnerability to estimate how the threat of disclosure affects the patch release time. The results are consistent with the theory in that a higher likelihood of early disclosure leads to an earlier patch. The authors also find in their data that in almost 15 percent of the cases, the patch release time coincides with disclosure. This may plausibly reflect

un-modeled communication between CERT and the vendor. However, it is also predicted by our model when $T^k < \tau^s$. In terms of our model, it appears that $\lambda > 0.5$ corresponds to $\tau^s > T^k$.

5 Customers can implement work-arounds

So far, we assumed that customers have no choice but to wait for the patch. However, sometimes customers can implement work-arounds instead. Examples of work-arounds include shutting off a port, changing the default settings, disabling a service, revising and recompiling a portion of code.

Suppose α fraction of users (henceforth *smart users*) can implement a work-around at a one time cost w, *if informed of the vulnerability by the social planner.*[15]A work-around will be implemented only if the expected loss of waiting for a patch, $l(\tau - T)$ is greater than w. Thus the expected loss for users implementing work-arounds is $\int_0^T l(T-x)dF(x) + w$. The remaining $(1-\alpha)$ percent of users must wait for the patch and their expected loss is as before. For brevity, we ignore kink solutions and post-patch losses, and patch quality. We also assume that all customers are fully informed about τ and T. Since there is no uncertainty in the model and customers are not being strategic, the vendor can credibly announce τ.

The possibility of work-arounds lowers the expected post disclosure loss for *smart* users, and in turn, the vendor has an incentive to delay the patch, which hurts users who can not apply work-arounds. However, the choice of the patch release time also affects customers' choice of work-around. In particular, if the vendor chooses τ such that $l(\tau - T) < w$ then smart users would not apply work-arounds and instead wait for the patch. In this case, the vendor loss function is

$$V(\tau, T) = C(\tau) + \lambda\Big(\int_0^T l(\tau - x)dF(x) + (1 - F(T))l(\tau - T)\Big)$$

$$\text{s.t.} \quad l(\tau - T) < w$$

If, however, the vendor chooses a τ such that $l(\tau - T) > w$, then smart users can apply a

[15]If *smart* users could apply work-around when the exploitation by the hackers start and not wait for the social planner to inform, then the potency of disclosure reduces. The social planner has no incentive to use disclosure as a tool to encourage work-around.

work-around and the vendor loss function, denoted by $V^w(\tau, T)$ is

$$V^w(\tau,T) = C(\tau) + \lambda\alpha\Big(\int_0^T l(T-x)dF(x) + w\Big) + \lambda(1-\alpha)\Big(\int_0^T l(\tau-x)dF(x) + (1-F(T))l(\tau-T)\Big)$$
$$\text{s.t.} \quad l(\tau - T) \geq w \tag{6}$$

Let $\tau^w = \arg\min_\tau\{V^w(\tau;T)\}$ and similarly $\tau = \arg\min_\tau\{V(\tau;T)\}$ denote the vendor's optimal choices of τ, conditional on allowing and not allowing the work-arounds respectively. If neither constraint binds, the patch release time in the work-around case is equivalent to the patch release time with no work-around but the internalization factor equal to $\lambda \cdot (1-\alpha)$. Thus for a given T, $\tau^w = \tau(\lambda \cdot (1-\alpha), T)$. While the patch release time in the no work-around case is $\tau(\lambda, T)$.

Define T^w as the protected period where the vendor is indifferent between inducing work-arounds and not, so that $V^w(\tau^w(T^w), T^w) = V(\tau(T^w), T^w)$. The following theorem establishes the conditions that T^w exists.[16]

Theorem 6. *(i) For all w such that $l(\tau(\lambda \cdot (1-\alpha)), 0) > w$ and $C(\tau^\infty) < \lambda w$, there exists a $0 < T^w < \tau^\infty$ such that the vendor induces work-arounds by smart users if $T \in [0, T^w)$ and does not induce work-arounds if $T > T^w$.*

(ii) $\frac{\partial T^w}{\partial \alpha} > 0$, and $\frac{\partial T^w}{\partial \lambda} < 0$.

By choosing a T smaller than T^w, the social planner can induce the vendor to choose a τ such that the work-around is feasible: releasing the patch very quickly after disclosure (the only way to head off a work-around) is costly. A T greater than T^w will, on the other hand, induce the vendor to release the patch soon after disclosure and avoid a work-around. An increase in α, the percent of smart users, will decrease $V^w(\cdot)$, and therefore increase T^w. A more subtle reasoning implies that increasing λ will decrease T^w. When the vendor is indifferent between inducing work-arounds or not, the patch development cost is smaller in the work-around case, and hence, the user-loss is greater. An increase in λ increases the weight of the user-loss component in the vendor cost function, so that the point of indifference is at a smaller T. The second part of the Theorem 6 formalizes this intuition.

The second part of Theorem 6 also points to the interaction between λ and α in conditioning

[16] For technical convenience, we assume that T^w is unique. The assumption is satisfied by a variety of functional forms for $l(\cdot)$, including the quadratic function.

T^w and, indirectly, also the protected period T. Thus, low λ and high α would lead to more work-arounds for a fixed T. However, they also affect optimal T.
Let the social cost, with and without work-around, respectively be:

$$S^w(\alpha) = C(\tau^w(T)) + \alpha\left(\int_0^T l(T-x)dF(x) + w\right) + (1-\alpha)\left(\int_0^T l(\tau^w(T)-x)dF(x) + (1-F(T))l(\tau^w(T)-T)\right)$$
s.t. $T < T^w(\alpha)$.

$$S(\alpha) = C(\tau(T)) + \int_0^T l(\tau(T)-x)dF(x) + (1-F(T))l(\tau(T)-T)$$
s.t. $T > T^w(\alpha)$.

Recall that T^w is a T such that the vendor is indifferent between work-around and no work-around ($V^w = V$). Since $\lambda < 1$, it is immediate that at T^w, $S^w(\cdot) > S(\cdot)$. Thus if the social planner finds it optimal to induce work-around, it will always set $T^* < T^w$.

If $S^w(\alpha) - S(\alpha)$ is decreasing in α, then we can find a $\hat{\alpha}$ such that $S^w(\hat{\alpha}) = S(\hat{\alpha})$.[17] Any $\alpha < \hat{\alpha}$ implies that the social planner will set a long enough protected period which does not induce a work-around. Conversely, when the percentage of *smart* users is greater than $\hat{\alpha}$, the social planner chooses a short protected period and discloses early, and the vendor delays the patch.

Theorem 6 provides insights into the interplay between λ and $\hat{\alpha}$. In particular, we conjecture that the social planner is more likely to prefer work-arounds and, hence, early disclosure when proportion of *smart* users is high and when λ is low. We perform numerical simulations using the same functional forms as before to analyze this ($C = 50000/\tau^{0.75} \quad L = 25.(\tau - T)^2$). We fix $z = 100$ and $w = 20$, $\lambda = 0.4$.

In Figure 8, as λ falls, the minimum percentage of smart users required for early disclosure, $\hat{\alpha}$, falls and, consequently, the disclosure policy is more aggressive. This relationship is conditioned by the cost of the work-around, w. As w increases, the protected period is larger. Note when λ is small, $\hat{\alpha}$ is also small. Intuitively, for a small λ, the patch release time is longer and the social planner will optimally disclose early enough to induce a work-around, even with a small fraction of smart users.

[17] As long as $S^w(\alpha) - S(\alpha)$ is decreasing in α and $w < C(\tau^s) + \int_0^{\tau^s} l(\tau^s - x)dF(x)$, a unique $\hat{\alpha}$ exists. Existence follows upon noting that $S^w(0) - S(0) > 0$ and $S^w(1) - S(1) < 0$ and $S^w(\alpha) - S(\alpha)$ is continuous in α. Further, $S(\alpha)$ weakly increases with α as well (the constraint $T > T^w$ becomes tighter). Thus, as long as $S^w(\alpha) - S(\alpha)$ is decreasing in α, $\hat{\alpha}$ is unique.

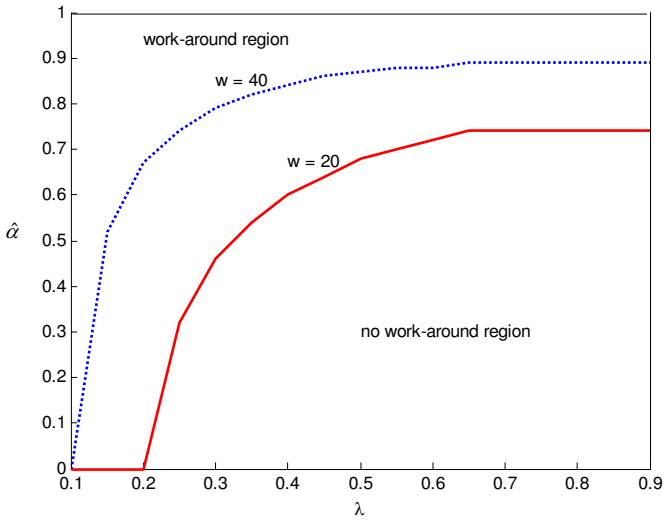


Figure 8: Required proportion of smart users to induce work-around

To summarize, when at least some users can implement work-arounds at a reasonable cost, disclosure acquires additional potency because it empowers these users. Clearly, the higher the fraction of users able to implement work-arounds, the more important it is to empower them via disclosure, and, thus shorter the protected period. Interestingly, when the vendor internalizes only a small fraction of user-losses, the empowerment aspect of disclosure becomes more significant: the social planner will choose quick disclosure even with a smaller fraction of users able to implement work-arounds.

One might expect that the "option value" implicit in *smart* users would always be socially beneficial. However, this intuition is incomplete because the vendor does not fully internalize user losses. Smart users convey a negative externality upon other users because smart users' presence leads the vendor to prefer the work-around option even when it is not socially desirable. As a result, the social planner may be forced to choose a longer protected period to avoid a work-around. This negative externality may even offset the beneficial effect leading to a net increase in social cost.[18] Indeed, for an intermediate range of α, where the vendor prefers to induce work-arounds but the social costs are lower without work-arounds, the presence of *smart* users raises total social cost.

Theorem 7. *If $S^w(\alpha) - S(\alpha)$ is decreasing in α then there exists a region between $[\hat{\hat{\alpha}}, \hat{\alpha}]$ such*

[18] We are grateful to an anonymous reviewer for pointing us in this direction.

that the presence of smart *users results in a higher social cost.*

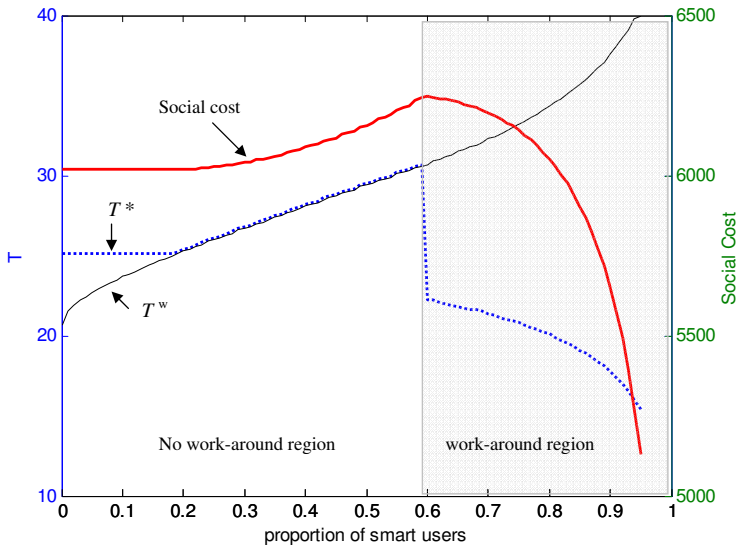


Figure 9: Social cost, T^* and T^w as a function of smart users

In Figure 9, we plot T and T^w on the left y-axis, and the social cost on the right y-axis. At $\alpha = 0$, the social cost is approximately 6000. For intermediate values of α, the social cost is higher than 6000. In Figure 9, $T^w = T^*(\alpha = 0)$ at $\alpha = 0.2$. Beyond that, the social planner strictly prefers no work-arounds but the only way to prevent work-arounds is by setting $T^* = T^w$. Once $\hat{\alpha}$ is reached, the social planner strictly prefers work-arounds (notice the jump in optimal T^*) and discloses early. Further increases in the fraction of smart users reduce social cost, and at $\alpha > 0.78$, the social cost drops below 6000.

The effect of market size follows the same intuition as in the previous section. A larger market increases the marginal cost of customer loss relative to the patch development cost, and hence decreases τ and T, and also reduces the gap between τ and T (see Figure 6). Thus, a larger market reduces T^w, makes a work-around less likely, and increases $\hat{\alpha}$.

6 Conclusions

How and when vulnerabilities should be disclosed is an important policy issue. A sensible disclosure policy must balance the need to protect users against attackers and the need to prod vendors to develop patches expeditiously. We develop a model that outlines how the policy

maker can optimally influence vendor behavior to minimize social cost.

We find that as long as the vendor does not internalize the entire user loss, the vendor will release the patch later than is socially optimal, unless threatened with disclosure. In some cases, the policy maker can force the vendor to release the patch at the socially optimal time while in other cases the optimal protected period is such that the vendor releases the patch after disclosure, though still earlier than the vendor would otherwise have. The more responsive the vendor is to user losses, the more aggressive the social planner can be by setting a shorter protected period. In general, both an instant disclosure and a secrecy policy are sub-optimal, though numerical simulations suggest that instant disclosure is particularly inefficient.

These results are robust to a partial implementation of the patch by users and to endogenous variations in the quality of the patch. When users take time to apply patches, the protected period should be longer. When the vendor also chooses patch quality (and higher quality patches are applied faster), contrary to conventional wisdom, a longer protected period may even reduce the quality of the patch and increase social cost.

When users can defend themselves via work-arounds, disclosure empowers users and increases the potency of disclosure policy, leading to a shorter protected period. However, this also creates a negative externality for users incapable of defending themselves. The vendor opts for work-arounds too readily, leading the social planner to extend the protected period in some cases. As a result, the social cost may actually rise as the proportion of users capable of implementing work-arounds increases. This suggests that unless the defensive measures are within the reach of a large enough number of users, encouraging their use may be counter-productive.

Our results are subject to a variety of qualifications. First, we leave for future research the case where a vendor can respond to disclosure by accelerating the rate of patch development, preferring to focus on the insights from a simpler static model. Thus, our model is best thought of relating to the policy rather than a patch release decision support system. We conjecture that allowing for such measures will lower the cost of disclosure, implying that the optimal protected period would be shorter. Second, in our model, costs and benefits are known with certainty. Thus it ignores the possibility that if the patch development cost is more expensive than expected and the vendor is making a good faith effort to develop a patch, the protected period may be optimally extended in a dynamic model with communication between the social planner and the

vendor. Third, we ignore the complications created when more than one vendor is affected by a single vulnerability. This case is empirically relevant and is the subject of a companion piece (Arora, Forman, Nandkumar and Telang 2006b), whose results suggest that the basic insights developed here are robust. Finally, we ignore the possibility that early disclosure would force vendors to provide secure software in the first place.

Despite these qualifications, our simple model is consistent with the observed evidence. Specifically, it correctly predicts that typically vendors release a patch after the disclosure, and that the gap between when the patch is released and the disclosure falls with the protected period. Moreover, though we work with continuous and convex loss functions, a kink in the vendor's response to the protected period arises naturally, yielding the prediction that the patch release time for some patches would coincide with disclosure.

Different assumptions may well lead to different conclusions about the optimal disclosure policy, but our model can be tailored to reflect those differences without major changes to the basic structure. In this sense, our model highlights the key areas where additional empirical evidence is required, by bringing out the key implications of the assumptions we have made. The contribution of this paper, therefore, lies not only in the specific results obtained but also in the framework developed that allows for various generalizations and highlights the possibilities and limits of social disclosure policy.

References

Arbaugh W.A., Fithen W. L., and McHugh J. (2000), "Windows of Vulnerability: A Case Study Analysis", *IEEE Computer.*

Arbaugh W.A., Browne H., McHugh J., and Fithen W.L. (2001), "A Trend Analysis of Exploitations". *IEEE Symposium on Security and Privacy.* Oakland, California, USA.

Arora A., Caulkins J.P., and Telang R. (2005), "Research Note - Sell First, Fix Later: Impact of Patching on Software Quality," *Management Science*, 52(3),465-471.

Arora A., Krishnan R., Telang R., and Yang Y. (2006a), "How Quickly do they Patch? An Empirical Analysis of Vendor Response to Disclosure Policies", *ICIS*, Milwaukee 2006.

Arora A., Nandkumar A., and Telang, R. (2006), "Does Information Security Attack Frequency Increase With Vulnerability Disclosure? - An Empirical Analysis", *Information Systems Frontier* 8, 350-362.

Arora A., Forman C., Nandkumar A., and Telang R. (2006b) "Competitive and Strategic Effects in the Timing of Patch Release" , *WEIS 2006*, Cambridge, England.

August T., and Tunca T. (2005), "Network Software Security and User Incentives," *Management Science*, 52 (11), 1703-1720.

Beattie S., Arnold S., Cowan C., Wagle, P., and Wright, C. (2002), "Timing the Application of Security Patches for Optimal Uptime", *Proc of LISA: 16th Systems Administration Conference*

Cavusoglu H., Cavusoglu H., and Raghunathan S. (2004), "How Should We Disclose Software Vulnerabilities?", 14^{th} *WITS*, Washington D.C.

Cavusoglu H., Cavusoglu H., and Zhang J. (2005), "Security Patch Management: Share the Burden or Share the Damage?,"Working paper, Tulane University.

Choi J., Fershtman C., and Gandal N. (2005), "Internet Security, Vulnerability Disclosure and Software Provision", *WEIS 2005*, Boston, MA June 2-3.

Clark R., (2002) http://www.blackhat.com/html/bh-usa-02/bh-usa-02-speakers.html#Richard Clarke, (*accessed July 19, 2006*).

CSI-FBI (2005) "CSI/FBI Computer Crime and Security Survey", 2006)

Gordon L.A., and Loeb M.P. (2002). "The Economics of Information Security Investment". *ACM Transactions on Information and System Security*, 5.

Infoworld, 2003, www.infoworld.com/article/03/06/30/HNpass_1.html; www.internetnews.com/dev-news/article.php/2203651 respectively (accessed, 08/24/05)

Kannan K., and Telang R. (2005), "Market for Software Vulnerabilities? Think Again", *Management Science,* 51(5), 726-740.

NetworkMagazine.com (2000), "The Pros and Cons of Posting Vulnerabilities". http://www.networkmagazine.com/article/NMG20001003S0001 (Accessed, 08/24/2005)

National Strategy to Secure Cyberspace (2003), http://www.whitehouse.gov/pcipb (Accessed, 08/24/2005)

Nizovtsev D., and Thursby M., (2005), "To disclose or not? An analysis of software user behavior", *Information Economics and Policy,* 19(1), 43-64.

Png Ivan, Tang C.Q., and Wang S.Y., (2006), "Information Security: User Precautions and Hacker Targeting", *WEIS 2006*, Cambridge, England.

Preston E., and Lofton J. (2002). "Computer Security Publications: Information Economics, Shifting Liability and the First Amendment", *Whittier Law Review*, 24, 71-142.

Rescorla E. (2003), "Security Holes... Who Cares?", *Proceedings of the 12th USENIX Security Conference*, August.

Symantec Inc (2003), “Symantec Internet Security Threat Report”. http://www.symantec.com

This page is intentionally blank. Proper e-companion title page, with INFORMS branding and exact metadata of the main paper, will be produced by the INFORMS office when the issue is being assembled.

Online Appendix

When *A5* is violated and not-patching may be privately optimal

A5 could be violated if post patch losses are high, i.e., if users do not apply patches. In this case, disclosure policy is pointless. So, in what follows, we discuss the case where post patch losses are zero, i.e., if $p(z,q)=1$. If *A5* is still violated, there are two possibilities. First, the vendor will not patch even under instant disclosure i.e., $\min_{\tau \leq t^{end}} \Big\{ C(\tau) + \lambda l(\tau) \Big\} \geq \lambda l(t^{end})$. In this case, there is no role for disclosure policy. If, on the other hand, $\min_{\tau \leq t^{end}} \Big\{ C(\tau) + \lambda l(\tau) \Big\} < \lambda l(t^{end})$ the role of disclosure policy exists.

The discussion here closely parallels the discussion of work-arounds. Define T^{NP} such that the vendor is indifferent between patching or not patching. Formally, $\min_{\tau} \Big\{ C(\tau) + \lambda \Big(\int_0^{T^{NP}} l(\tau - x) dF(x) + (1 - F(T^{NP})) l(\tau - T^{NP}) \Big) \Big\} = \lambda \Big(\int_0^{T^{NP}} l(t^{end} - x) dF(x) + (1 - F(T^{NP})) l(t^{end} - T^{NP}) \Big)$ Note that as long as $\min_{\tau} \Big\{ C(\tau) + \lambda l(\tau) \Big\} < \lambda l(t^{end})$ i.e., as long as it is privately optimal to patch under instant-disclosure, T^{NP} exists because the violation of *A5* implies that the vendor's cost with patching is strictly greater than without patching at $T = t^{end}$. Since the cost with and without patching is continuous in T, T^{NP} exists. Further, the derivative of the cost with patching with respect to T is $-(1 - F(T)) l'(\tau(T) - T)$ and that of the cost without patching is $-(1 - F(T)) l'(t^{end} - T)$. The second derivative is smaller (larger in absolute value) because $t^{end} \leq \tau(T)$. Thus, the cost without patching falls faster than the cost with patching as the protected period increases. This implies that T^{NP} is unique. It is immediate that $T \leq T^{NP}$ implies that the vendor will patch and $T > T^{NP}$, the vendor will not patch. Finally, note that $\lambda \int_0^{T^{NP}} l(t^{end} - x) dF(x) + (1 - F(T^{NP})) l(t^{end} - T^{NP})$ is increasing in t^{end}, so that T^{NP} is increasing in t^{end}.

It is also obvious that the social cost with patching is strictly less than the social cost without patching at $T = T^{NP}$.

$$S(T^{NP}) = \min_{\tau} \Big\{ C(\tau) + \int_0^{T^{NP}} l(\tau - x) dF(x) + (1 - F(T^{NP})) l(\tau - T^{NP}) \Big\} <$$

$$\int_0^{T^{NP}} l(t^{end} - x) dF(x) + (1 - F(T^{NP})) l(t^{end} - T^{NP}) = S^{NP}(T^{NP}; t^{end})$$

If the social planner wishes for the vendor not to patch, the optimal protected period is t^{end}. The social cost is $\int_0^{t^{end}} l(t^{end} - x) dF(x)$. If the social planner wishes for the vendor to patch, it chooses $T \leq T^{NP}$ to minimize $S(T)$.

Define $\hat{\hat{t}}$ as the lifecycle length such that the (unconstrained) optimal protected period is just short enough to induce patching i.e., $\arg\min_{T} \Big\{ C(\tau(T)) + \int_0^{T} l(\tau(T) - x) dF(x) + (1 - F(T)) l(\tau(T) - T) \Big\} = T^{NP}(\hat{\hat{t}})$. Define $\hat{t}$ as such a t^{end} that the social planner is indifferent between having the vendor patch or not i.e., $\min_{T} \Big\{ C(\tau(T)) + \int_0^{T} l(\tau(T) - x) dF(x) + (1 - F(T)) l(\tau(T) - T) \quad \text{s.t.} \quad T \leq$

$T^{NP}(\hat{t})\Big\} = \int_0^{\hat{t}} l(\hat{t}-x)dF(x)$. Then, for $t^{end} < \hat{t}$, the social planner simply maintains the vulnerability secret and the vendor does not patch. For $\hat{\hat{t}} \geq t^{end} \geq \hat{t}$, the social planner prefers to have the vendor develop a patch but is constrained by the need to induce patching and sets a protected period $T = T^{NP}(t^{end})$. For $t^{end} > \hat{\hat{t}}$ the social planner's choice is unconstrained and corresponds to the case analyzed in the text.

Proof of Lemma 1

Step (i) and (ii) show that $L(\tau, T)$ is convex in τ for $\tau < T$ and for $\tau > T$. Utilizing these results, we show in (iii) that it is convex everywhere.

(i) We first show that for any $\tau > T$, $L_{\tau\tau} > 0$.

From equation (1),

$$L_\tau = \lambda\Big(\int_0^T l^{'}(\tau - x)dF(x) + (1 - F(T))l^{'}(\tau - T)\Big)$$
$$L_{\tau\tau} = \lambda\Big(\int_0^T l^{''}(\tau - x)dF(x) + (1 - F(T)l^{''}(\tau - T)\Big)$$

Since $l(.)$ is strictly convex in τ, $L_{\tau\tau} > 0$, i.e. $L(\tau)$ is strictly convex in τ.

(ii) We show that for any $\tau < T$, $L(\tau)$ is strictly convex in τ. From equation (1), $L_\tau = \lambda \int_0^\tau l^{'}(\tau - x)dF(x)$. Hence, $L_{\tau\tau} = \lambda \int_0^\tau l^{''}(\tau - x)dF(x) + l^{'}(0) > 0$

(iii) We need to prove that, for any $\tau_1 < \tau_2$ and $0 < k < 1$,

$$kL(\tau_1) + (1-k)L(\tau_2) > L(k\tau_1 + (1-k)\tau_2) \tag{O1}$$

From (i) and (ii), when $\tau_1 < \tau_2 \leq T$ or $T \leq \tau_1 < \tau_2$, (O1) holds. We only need to prove the inequality holds when $\tau_1 \leq T \leq \tau_2$.

Define $\hat{k}$ such that $\hat{k}\tau_1 + (1-\hat{k})\tau_2 = T$. When $k \leq \hat{k}$, $k\tau_1 + (1-k)\tau_2 = \tau \leq T$. Let $M(k) = L(\tau) - kL(\tau_1) - (1-k)L(\tau_2)$. Convexity is equivalent to $M(k) < 0$.

We first show that $M(\hat{k}) < 0$. Note that

$$\begin{aligned}
\int_0^{\tau_2} l(\tau_2 - x)dF(x) &= \int_0^T l(\tau_2 - x)dF(x) + \int_T^{\tau_2} l(\tau_2 - x)dF(x) \\
&\leq \int_0^T l(\tau_2 - x)dF(x) + \int_T^{\tau_2} l(\tau_2 - T)dF(x) \\
&\leq \int_0^T l(\tau_2 - x)dF(x) + \int_T^{\infty} l(\tau_2 - T)dF(x) \\
&\leq \int_0^T l(\tau_2 - x)dF(x) + (1 - F(T))l(\tau_2 - T) = L(\tau_2)
\end{aligned}$$

Therefore, $\hat{k}.L(\tau_1) + (1-\hat{k})L(\tau_2) \geq \hat{k}\int_0^{\tau_1} l(\tau_1 - x)dF(x) + (1-\hat{k})\int_0^{\tau_2} l(\tau_2 - x)dF(x)$

The convexity of $\int_0^{\tau} l(\tau - x)dF$ in τ implies that $k\int_0^{\tau_1} l(\tau_1 - x)dF(x) + (1-k)\int_0^{\tau_2} l(\tau_2 - x)dF(x) > \int_0^{\tau} l(\tau - x)dF(x)$

At $k = \hat{k}$, $\hat{k}\int_0^{\tau_1} l(\tau_1 - x)dF(x) + (1-\hat{k})\int_0^{\tau_2} l(\tau_2 - x)dF(x) > \int_0^{T} l(T - x)dF(x) = L(T)$. Hence $\hat{k}L(\tau_1) + (1-\hat{k})L(\tau_2) > L(T)$

Recall $M(\hat{k}) = L(T) - \hat{k}.L(\tau_1) - (1-\hat{k})L(\tau_2)$. Thus $M(\hat{k}) < 0$

Note that $\frac{dM}{dk} = L_\tau \frac{\partial \tau}{\partial k} - L(\tau_1) + L(\tau_2) = (\tau_1 - \tau_2)L_\tau(\tau) - L(\tau_1) + L(\tau_2)$

And $\frac{d^2M}{dk^2} = (\tau_1 - \tau_2)^2 L_{\tau\tau}(\tau)$

For all $k \neq \hat{k}$, $L_{\tau\tau}(\tau) > 0$. Therefore, $\frac{d^2M}{dk^2} > 0$ when $k \neq \hat{k}$. Also, M(0) = M(1)= 0 and M($\hat{k}$) < 0. Furthermore, $M(k) > 0$ is not possible because it would require $\frac{d^2M}{dk^2} < 0$ over some interval. Hence we must have that $M(k) < 0, \forall 0 < k < 1$. **QED**

Existence and uniqueness of τ^*

From equation (3), $V(\tau, T) = C(\tau) + \lambda L(\tau)$ is a continuous function defined over a compact domain $[0, t^{end}]$, and must attain a minimum. By *A4*, $C_\tau(0) = -\infty$ so that developing the patch instantaneously cannot be optimum, and by *A5*, the vendor voluntarily develops the patch before t^{end}, i.e., $arg\min_{\tau}\{C(\tau) + \int_0^T l(\tau - x)dF(x)\} < t^{end}$. Hence the minimum point must be in the interior of the domain so that $0 < \tau^* < t^{end}$. By the strict convexity of V in τ, τ^* must be unique. **QED**

Proof of Lemma 2

(i) A unique T^k exists.

Recall that $V_\tau^+(T)\Big|_{\tau=T} = \left(C_\tau(\tau) + \lambda\left(\int_0^T l'(\tau - x)dF(x) + (1 - F(T))l'(0)\right)\right)\Big|_{\tau=T}$. To save on notation, define $\phi(T) = V_\tau^+(T)\Big|_{\tau=T}$. Then T^k is unique if $\phi(T) = 0$ has a unique solution. By *A4*, $\phi(0) < 0$. By *A5*, $arg\min_{\tau}\{C(\tau) + \int_0^T l(\tau - x)dF(x)\} < t^{end}$ and convexity of V implies that $\phi(t^{end}) = C_\tau(t^{end}) + \lambda\int_0^{t^{end}} l'(t^{end} - x)dF(x) > 0$. By the Intermediate Value theorem, there exists at least one T^k such that $\phi(T_K) = 0$. We have already shown in the text that $\phi(T)$ is increasing in T, so that for any $T < T^k$, $\phi(T) < 0$, and for any $T > T^k$, $\phi(T) > 0$. Hence, T^k must be unique.

(ii) There exists a unique τ^s.

The proof that $C_\tau(\tau) + \int_0^\tau l'(\tau - x)dF(x) = 0$ has a unique solution is analogous to the proof of uniqueness of τ^* shown in Lemma 1 and hence not repeated here.

(iii) There exists a unique τ^∞.

A continuous function on a compact domain attains a minimum. $C(\tau) + \lambda \int_0^\tau l(\tau - x)dF(x)$ is continous in τ in the range $[0, t^{end}]$. Therefore τ^∞ exists. By *A5* $\tau^\infty < t^{end}$. By *A4*, $\tau^\infty > 0$. Hence τ^∞ is in the interior. $C(\tau)+\lambda \int_0^\tau l(\tau - x)dF(x)$ is strictly convex in τ in the range $[0, t^{end}]$. Therefore τ^∞ is unique. Moreover, $\tau^\infty > T^k$ because $C_\tau(T^k)+ \int_0^{T^k} l^{'}(T^k - x)dF(x) < \phi(T^k) = 0$. **QED**

Proof of Lemma 3

(i) Note that $\frac{\partial T^k}{\partial \lambda} = -\frac{\phi_\lambda}{\phi_T}$. $\phi_T > 0$ and $\phi_\lambda = \int_0^T l^{'}(T - x)dF(x) + (1 - F(T))l^{'}(0) > 0$. Thus $\frac{\partial T^k}{\partial \lambda} < 0$.

(ii) Define $\phi(T) = V_\tau^+(T)\Big|_{\tau=T}$. We can rewrite $\phi(T)$ after integrating by parts as $\phi(T) = C_\tau(T) + \lambda.\Big(\int_0^T l^{''}(T - x)F(x)dx + l^{'}(0)\Big)$. $\phi(T^k) = 0$ implicitly defines T^k. Replacing $F(x)$ by $G(x)$ would increase increase the second term, $\lambda.\Big(\int_0^T l^{''}(T - x)F(x)dx\Big)$, and leave the other two unchanged. Since $\phi(T)$ is increasing in T, it must be that T^k corresponding to $G(x)$ is smaller than the T^k corresponding to $F(x)$. **QED**

Proof of Theorem 1

Theorem 1 has four major results. We prove them one by one.

(i). When $T \in [0, T^k)$, $T < \tau^* < T^k$

Recall that $\phi(T)$ is monotonically increasing in T, and $\phi(T^k) = 0$. Hence, for $T \in [0,\ T^k)$, $\phi(T) < \phi(T^k) = 0$. If the vendor's cost were minimized at $\tau = T$, $V_\tau^-(T) < 0$ and $V_\tau^+(T) \equiv \phi(T) \geq 0$, which leads to a contradiction. Hence, τ^* cannot be at T . Further, $\tau^* < T$ also leads to a contradiction because it implies that T does not constrain the vendor's patching time i.e., $\tau^* = \tau^\infty$. In other words, $\tau^* < T$ implies $T > \tau^\infty$. Note that $\tau^\infty > T^k$ (by Lemma 2) which implies $T > T^k$, which contradicts the condition that $T \in [0,\ T^k)$.

(ii). For $T \in [0,\ T^k)$, $0 < \frac{\partial \tau^*}{\partial T} < 1$.

We have already shown that τ^* is an interior point in this range, so that $V_\tau(\tau^*) = 0$. Differentiating both sides of the equation w.r.t T, we get $V_{\tau\tau}\frac{\partial \tau^*}{\partial T} + V_{\tau T} = 0$. Thus, $\frac{\partial \tau^*}{\partial T} = -\frac{V_{\tau T}}{V_{\tau\tau}}$.

Substituting we get, $\frac{\partial \tau^*}{\partial T} = \frac{\lambda\Big(1-F(T)\Big)\Big(l^{"}(\tau - T)\Big)}{C_{\tau\tau}+\lambda \int_0^T l^{"}(\tau - T)dF(x)+\lambda\Big(1-F(T)\Big)l^{"}(\tau - T)}$. Since the numerator and denominator are positive and the denominator is greater than the numerator, $0 < \frac{\partial \tau^*}{\partial T} < 1$.

(iii). For $T \in [T^k, \tau^\infty)$, $\tau^* = T$ and $\frac{\partial \tau^*}{\partial T} = 1$

As in part (i) above, $\tau^* < T$ is not possible for $T < \tau^\infty$. For $\tau^* > T$,

$V_\tau = \left(C_\tau(\tau) + \lambda \int_0^T l'(\tau - x) dF(x) + \lambda(1 - F(T)) l'(\tau - T) \right) > \phi(T) > \phi(T^k) = 0$. Therefore $\tau^* > T$ is not possible either. Since a minimum always exists, it must be that $\tau^* = T$. Since this holds for all $T \in [T^k, \tau^\infty)$, it follows that $\frac{\partial \tau^*}{\partial T} = 1$.

(iv) For $T \in [\tau^\infty, t^{end}), \tau^* = \tau^\infty$.

The result is obvious upon noting that $\tau^* \leq \tau^\infty$. **QED**

Proof of Corollary 1

From Theorem 1, when $T \geq T^k$, the vendor always patches at T or at τ^∞, whichever is smaller. Hence, a change in λ does not change τ^*. When $T < T^k$, the vendor's cost is minimized at an interior point. Totally differentiating $V_\tau = 0$ (the first order condition for an interior optimum) and re-arranging terms we get $\frac{\partial \tau^*}{\partial \lambda} = -\frac{V_{\tau\lambda}}{V_{\tau\tau}}$. Note that the numerator $V_{\tau\lambda} = L_\tau > 0$ and denominator is always positive. Thus, it follows that $\frac{\partial \tau^*}{\partial \lambda} < 0$. **QED**

Proof for the Existence and Uniqueness of T^*

(i) There exists an optimal disclosure time T^*.

Increases in T beyond τ^∞ leave social and private costs unchanged because the vendor will always patch at τ^∞. Thus we can restrict our attention to the range [0,τ^∞]. $S(T)$ is continuous on the compact range and therefore must attain a minimum in this range. **QED**

(ii) If $\frac{C_{\tau\tau}(\tau)}{\lambda} \frac{l'(\tau - T)}{l''(\tau - T)} + \frac{l'(\tau - T)}{l''(\tau - T)} \int_0^T l''(\tau - x) dF(x) + \lambda(1 - F(T)) l'(\tau - T) < (1 - \lambda) \int_0^T l'(\tau - x) dF(x)$ holds then T^* is unique.

Suppose T^1 marks the first local minimum for social cost. We show that for any $T > T^1$, $\frac{dS(T)}{dT} > 0$ implying that $S(T)$ does not have any stationary points after T^1. Hence T^1 must be the only point where $S(T)$ reaches a minimum. Recall from the definition of S that

$$\frac{dS}{dT} = S_\tau . \frac{d\tau}{dT} + S_T = (1 - \lambda) \left[\int_0^T l'(\tau - x) dF(x) + (1 - F(T)) l'(\tau - T) \right] \frac{d\tau}{dT} + (F(T) - 1) l'(\tau - T)$$

and $\frac{d\tau}{dT} = \frac{\lambda(1 - F(T)) l"(\tau - T)}{C_{\tau\tau}(\tau) + \lambda \int_0^T l"(\tau - x) dF(x) + \lambda(1 - F(T)) l"(\tau - T)}$. Thus, $\frac{dS}{dT} > 0$ *iff*

$$\frac{\lambda(1 - F(T)) l"(\tau - T)}{C_{\tau\tau}(\tau) + \lambda \int_0^T l"(\tau - x) dF(x) + \lambda(1 - F(T)) l"(\tau - T)} > \frac{(1 - F(T)) l'(\tau - T)}{(1 - \lambda) \left[\int_0^T l'(\tau - x) dF(x) + (1 - F(T)) l'(\tau - T) \right]}$$

. Simplifying it leads to the required result. Note that $S(T)$ is not differentiable at T^k. However, if the first minimum is at T^k the condition applied to the right derivative of T^k is sufficient to rule out an interior minimum beyond T^k.

Proof of Theorem 2

(i) When $\tau^s \geq T^k, T^* = \tau^s$

In this case it is obvious that setting $T > \tau^s$ cannot be optimal because whenever $T \geq T^k$ the vendor will patch at T yielding a social cost equal to $C(T) + \int_0^T l(T-x)dF(x)$. By definition, this function is minimized at $T = \tau^s$. It only remains to check that any $T < T^k$ yields strictly higher social cost. The social cost in this case is $C(\tau) + \int_0^T l(\tau - x)dF(x) + (1 - F(T))l(\tau - T)$. Note that $\int_0^T l(\tau - x)dF(x) + (1 - F(T))l(\tau - T) \geq \int_0^T l(\tau - x)dF(x) + \int_T^{t^{end}} l(\tau - T)dF(x) > \int_0^T l(\tau - x)dF(x) + \int_T^{t^{end}} l(\tau - x)dF(x) > \int_0^\tau l(\tau - x)dF(x)$. It follows that the social cost in this case, $C(\tau) + \int_0^T l(\tau - x)dF(x) + (1 - F(T))l(\tau - T) > C(\tau) + \int_0^\tau l(\tau - x)dF(x) > C(\tau^s) + \int_0^{\tau^s} l(\tau^s - x)dF(x)$. Therefore, $T < T^k$ is not optimal. Hence, $T^* = \tau^s$ is the only possibility.

ii) When $\tau^s < T^k$, $\tau^s < T^* \leq \tau(T^*) \leq T^k$.

When $T \in [0, T^k)$, Theorem 1 implies that $T < \tau(T)$. In the next theorem, we show that $\frac{dT^*}{d\lambda} < 0$. Since τ^s is the optimal T^* when $\lambda = 1$, it must be that optimal $T^* > \tau^s$ when $\lambda < 1$.

Next we show that $T^* \leq T^k$. The first derivative of S evaluated at $T > T^k$ is positive, as shown below.

$\frac{dS}{dT} = S_\tau . \frac{\partial \tau^*}{\partial T} + S_T = S_\tau + S_T$ because $\frac{\partial \tau^*}{\partial T} = 1$ when $T > T^k$. $S_\tau + S_T = C_\tau(T) + \int_0^T l'(T-x)dF(x) > C_\tau(\tau^s) + \int_0^{\tau^s} l'(\tau^s - x)dF(x) = 0$, because $T > \tau^s$ and $C_\tau(T) + \int_0^T l'(T-x)dF(x)$ is increasing in T. Therefore, it must be that $S_\tau + S_T > 0$.

Hence $\tau^s < T^* \leq T^k$. As already proved, $T < T^k$ implies $\tau^s < T^* \leq \tau(T^*) \leq T^k$. **QED**

Proof of Theorem 3

i) There exists a λ_0 so that $\lambda \geq \lambda_0$ implies $\tau^s \geq T^k$ and $\lambda < \lambda_0$ implies $\tau^s < T^k$.

We first show that T^k is monotonically decreasing in λ . By definition, $\phi(T^k) = V_\tau^+ \Big|_{\tau = T^k} = 0$. Differentiating both sides w.r.t. λ, $\frac{\partial \phi}{\partial T}\frac{dT^k}{d\lambda} + \frac{\partial \phi}{\partial \lambda} = 0$. From the proof of Lemma 3, $\frac{\partial \phi}{\partial \lambda} > 0$. Also, $\frac{\partial \phi}{\partial T} > 0$, hence $\frac{dT^k}{d\lambda} < 0$. From (4), τ^s stays constant when λ changes.

When $\lambda = 1$, $\phi(T^k) = C_\tau(T^k) + \int_0^{T^k} l'(T^k - x)dF(x) + (1 - F(T^k))l'(0) = 0$. Recall that $C_\tau(T) + \int_0^T l'(T-x)dF(x) = 0$ solves for τ^s. Comparing the conditions, it is clear that $T^k < \tau^s$. When $\lambda = 0$, $T^k = \infty > \tau^s$. Thus, by the Intermediate Value Theorem, there exists a $0 \leq \lambda_0 \leq 1$ such that $T^k(\lambda_0) = \tau^s$. Hence, when $\lambda \geq \lambda_0$, $\tau^s \geq T^k(\lambda)$ and $\lambda < \lambda_0$ implies $\tau^s < T^k(\lambda)$.

ii) For $\lambda < \lambda_0$, $\tau^s < T^k$. In this range, a higher internalization ratio, λ, implies a lower socially optimal protected period, T^*.

We want to prove that $\frac{dT^*}{d\lambda} < 0$. When $\tau^s < T^k$, T^* is either an interior point or $T^* = T^k$. When $T^* = T^k$ then $\frac{dT^*}{d\lambda} < 0$. To see this, first note that $\frac{dT^k}{d\lambda} < 0$ from (i) above. If T^* does not decrease with λ then $T^* > T^k$ for some λ which is not possible.

If T^* is in the interior then satisfies first-order condition, i.e. $\frac{dS}{dT} = 0$. For ease of notation, let $Q(T) = \frac{dS}{dT} = S_\tau.\frac{\partial \tau^*}{\partial T} + S_T$. So $Q(T^*) = 0$. By the implicit function theorem, $\frac{dT^*}{d\lambda} = -\frac{\frac{dQ}{d\lambda}}{\frac{dQ}{dT}}$. The denominator is simply $\frac{d^2S}{dT^2}$ which is positive for any interior minimum. Thus, we need to show that $\frac{dQ}{d\lambda} > 0$.

$$\begin{aligned}\frac{dQ}{d\lambda} = \frac{d^2S}{dTd\lambda} &= S_{\tau\tau}.\frac{\partial \tau}{\partial \lambda}.\frac{\partial \tau^*}{\partial T} + S_\tau.\frac{\partial^2 \tau}{\partial T \partial \lambda} + S_{\tau T}.\frac{\partial \tau}{\partial \lambda} \\ &= \frac{\partial \tau}{\partial \lambda}\left(S_{\tau\tau}.\frac{\partial \tau^*}{\partial T} + S_{\tau T}\right) + S_\tau.\frac{\partial^2 \tau}{\partial T \partial \lambda}\end{aligned}$$

Recall that $S_{\tau\lambda} = 0$ and $V_\tau = 0$ at any interior minimum. Thus, we must have $C_\tau = -\lambda.L_\tau$. Therefore $S_\tau = (1-\lambda)L_\tau > 0$. From Corollary 1, we also know that $\frac{\partial \tau^*}{\partial \lambda} < 0$. In summary, it is sufficient for the numerator to be positive if term in parenthesis above is negative and $\frac{\partial^2 \tau}{\partial T \partial \lambda} > 0$. We show them respectively.

1. We first show that $S_{\tau\tau}.\frac{\partial \tau^*}{\partial T} + S_{\tau T} < 0$.
   Recall that $\frac{\partial \tau^*}{\partial T} = -\frac{V_{\tau T}}{V_{\tau\tau}} = -\frac{\lambda.S_{\tau T}}{V_{\tau\tau}}$. Substituting this above

   $$\begin{aligned}S_{\tau\tau}.\frac{\partial \tau^*}{\partial T} + S_{\tau T} &= -S_{\tau\tau}\frac{\lambda.S_{\tau T}}{V_{\tau\tau}} + S_{\tau T} \\ &= S_{\ tauT}\left(1 - \lambda\frac{S_{\tau\tau}}{V_{\tau\tau}}\right)\end{aligned}$$

   Note than $S_{\tau T} = -(1-F(T))l^{"}(\tau - T) < 0$ and $\lambda.\frac{S_{\tau\tau}}{V_{\tau\tau}} = \lambda.\frac{C_{\tau\tau}(\tau)+L^{"}(\tau)}{C_{\tau\tau}(\tau)+\lambda L^{"}(\tau)} < 1$ Thus, $S_{\tau\tau}.\frac{\partial \tau^*}{\partial T} + S_{\tau T} < 0$.

2. We show that $\frac{\partial^2 \tau}{\partial T \partial \lambda} > 0$.
   $\frac{\partial \tau^*}{\partial T} = -\frac{V_{\tau T}}{V_{\tau\tau}}$. Note that the numerator $V_{\tau T} = \lambda.S_{\tau T}$, hence its partial derivative w.r.t. λ is $S_{\tau T}$. The partial derivative of $V_{\tau\tau}$ w.r.t. λ is $L_{\tau\tau}$. Thus, $\frac{\partial^2 \tau}{\partial T \partial \lambda} = -\frac{S_{\tau T}.V_{\tau\tau} - \lambda.S_{\tau T}.L_{\tau\tau}}{V_{\tau\tau}^2} = \frac{S_{\tau T}.(\lambda L_{\tau\tau} - V_{\tau\tau})}{V_{\tau\tau}^2}$. Moreover, $V_{\tau\tau} = C_{\tau\tau} + \lambda.L_{\tau\tau}$ where $C_{\tau\tau} > 0$. Since $S_{\tau T} < 0$, it must be that $\frac{\partial^2 \tau}{\partial T \partial \lambda} > 0$

Thus, $\frac{dG}{d\lambda} > 0$ and hence $\frac{dT^*}{d\lambda} < 0$. **QED**

Proof of Theorem 4

We want to show that when customers do not patch, then (i) T^k moves to the right. (ii) Both vendor and social planner become less aggressive. Recall that $0 < p(z) < 1$.

(i) For readability, let $\tilde{T}^k$ be the solution of $\phi(T) = V_\tau^+(T)|_{\tau=T}$ in the presence of post-patch

losses. By definition of $\tilde{T}^k$, we have

$$\phi(\tilde{T}^k) = C_\tau(\tau) + \lambda\Big(\int_0^T l^{'}(\tau - x)dF(x) + (1 - F(T))l^{'}(0) - \zeta(t^{end} - \tau)(1 - p(t^{end} - \tau))\Big)\Big|_{\tau=\tilde{T}^k} = 0$$
$$= C_\tau(\tilde{T}^k) + \lambda\Big(\int_0^{\tilde{T}^k} l^{'}(\tilde{T}^k - x)dF(x) + (1 - F(\tilde{T}^k))l^{'}(0) - \zeta(t^{end} - \tilde{T}^k)(1 - p(t^{end} - \tilde{T}^k))\Big) = 0 \tag{O2}$$

In the absence of post-patch loss, T^k was defined as

$$V_\tau^+\Big|_{\tau=T^k} = C_\tau(T^k) + \lambda\Big(\int_0^{T^k} l^{'}(T^k - x)dF(x) + (1 - F(T^k))l^{'}(0)\Big) = 0$$

Since (O2) is increasing T (by *A3* and the convexity of C and l in τ), it must be that $\tilde{T}^k > T^k$.

(ii) Suppose τ^* and T^* were the optimal patching time and protected period in the absence of post-patch loss. Fix T^* and consider the optimal $\tilde{\tau}^*$ in the presence of post-patch loss. The vendor solves:

$$V_{\tilde{\tau}} = C_{\tilde{\tau}} + \lambda\Big(\int_0^{T^*} l^{'}(\tilde{\tau} - x)dF(x) + (1 - F(T^*))l^{'}(\tilde{\tau} - T^*) - \zeta(t^{end} - \tilde{\tau})(1 - p(t^{end} - \tilde{\tau}))\Big) = 0$$

It is immediate that optimal $\tilde{\tau}^* > \tau^*$ for the same T^*. We also know that social planner solves:

$$\frac{dS}{dT}\Big|_{\tau=\tau^*, T=T^*} = S_\tau . \frac{\partial \tau}{\partial T} + S_T = 0 \tag{O3}$$

Now, consider what happens to the optimal protected period in the presence of post-patch loss. Consider

$$\frac{dS}{dT}\Big|_{\tilde{\tau}=\tilde{\tau}^*, T=T^*} = S_{\tilde{\tau}} . \frac{\partial \tilde{\tau}}{\partial T} + S_T \tag{O4}$$

$S_{\tilde{\tau}} = C_{\tilde{\tau}}(\tilde{\tau})(1 - \frac{1}{\lambda})$, $S_\tau = C_\tau(\tau)(1 - \frac{1}{\lambda})$. Since, $C_\tau < C_{\tilde{\tau}}$, and $(1 - \frac{1}{\lambda}) < 1$, it must be that $S_{\tilde{\tau}} < S_\tau$. Similarly $S_T|_{\tau=\tilde{\tau}} = -(1 - F(T))l^{'}(\tilde{\tau} - T) < S_T|_{\tau=\tau}$. Finally, we need that $\frac{\partial \tilde{\tau}}{\partial T} < \frac{\partial \tau}{\partial T}$ which we assume. [19]

In summary, all positive terms in (O4) are smaller than in (O3) and the negative term is larger. Therefore, $\frac{dS}{dT}\Big|_{\tau=\tau^*, T=T^*} < 0$. Since S is convex and has an interior optimum, the social planner should set optimal $\tilde{T}^* > T^*$. **QED**

[19] It is hard to show $\frac{\partial \tilde{\tau}}{\partial T} < \frac{\partial \tau}{\partial T}$ without additional functional form restrictions. However, from (O3), $\frac{\partial \tau}{\partial T} = -\frac{S_T}{S_\tau} = \frac{(1-F(T))l^{'}(\tau - T)}{C_\tau(1-\frac{1}{\lambda})}$. $\frac{\partial \tilde{\tau}}{\partial T} = \frac{\lambda(1-F(T))l^{"}(\tilde{\tau}-T)}{V_{\tilde{\tau}\tilde{\tau}}}$. Thus all we need is $\frac{V_{\tilde{\tau}\tilde{\tau}}}{l^{"}(\tilde{\tau}-T)} > \frac{C_\tau(\lambda-1)}{l^{'}(\tau-T)}$ for $\frac{\partial \tilde{\tau}}{\partial T} < \frac{\partial \tau}{\partial T}$. Intuitively, it is easy to see that this holds. Particularly, consider the neighborhood of $[T^k, \tilde{T}^k]$. We know that $\frac{\partial \tau}{\partial T} = 1$ and $\frac{\partial \tilde{\tau}}{\partial T} < 1$ in this region.

Proof of Theorem 5:

(i) When $T < T^k$, vendor cost is minimized at an interior point.

We need to show that $\frac{d\tau^*}{dT} > 0$; and $\frac{dq^*}{dT} \geq 0 \Leftrightarrow V_{q\tau} \leq 0$. Let $H(\tau, q)$ be the determinant of the Hessian formed from V. Since V is convex, $H(\tau, q) > 0$. By Cramer's Rule, and upon noting that $V_{qT} = 0$ we get

$$\frac{d\tau^*}{dT} = -\frac{V_{\tau T}.V_{qq}}{H(\tau, q)}$$

Note that $V_{\tau T} = -(1 - F(T))l''(\tau - T) < 0$. Since, $V_{qq} > 0$, we have $\frac{d\tau^*}{dT} > 0$.

Similarly, we have that

$$\frac{dq^*}{dT} = \frac{V_{\tau T}.V_{\tau q}}{H(\tau, q)}$$

Thus if $V_{\tau q} \leq 0$ then $\frac{dq^*}{dT}$ is positive and negative otherwise.

(ii) When $T \geq T^k$, the vendor cost is minimized at the kink, i.e. $\tau^* = T$

$V_q|_{\tau=\tau^* \& q=q^*} = 0$. Taking the total derivative w.r.t. q and noticing that $V_{qT} = 0$, $\frac{dq^*}{dT} = -\frac{V_{\tau q}}{\tilde{V}_{qq}}$, it follows that $\frac{dq^*}{dT} \geq 0 \Leftrightarrow V_{\tau q} \leq 0$. **QED**

Proof of the Statement: The optimal protected period is lower in case of endogenous quality when post-patch loss dominates patch development cost

To prove this, we compare optimal T in when quality is exogenous and when it is endogenous. For the comparison to be worthwhile, we fix the quality in exogenous case to be the same as in endogenous case.

Let $S(T)$ be the social cost. Then in the case when quality is endogenous, the social cost is

$$S(T) = C(\tau^*(T), q^*(T), T) + L(\tau^*(T), T) + \tilde{L}(\tau^*(T), q^*(T))$$

The socially optimal protected period, T^*, is implicitly given by the first order condition

$$\frac{dS}{dT} = \left[C_\tau + L_\tau + \tilde{L}_\tau\right] \frac{d\tau}{dT} + \left(C_q + \tilde{L}_q\right) \frac{dq}{dT} + L_T = 0 \tag{O5}$$

Consider $\frac{dS}{dT}$ evaluated at $T = T^*$ when q be exogenously fixed at q^*. Since q is exogenous, the derivative of S with respect to T is given by

$$\frac{dS}{dT}\Big|_{T=T^*,\ q \text{ is exogenous}} = \left[C_\tau + L_\tau + \tilde{L}_\tau\right] \frac{d\tau}{dT} + L_T \tag{O6}$$

If $\frac{dS}{dT}\Big|_{T=T^*}$ given in (O6) is negative then optimal T when quality is exogenous must be larger than T^* (because S has as an interior minima). Note that the patch release time τ is same in both (O5) and (O6) but $\frac{d\tau}{dT}$ is different in the two equations. However, one can show that $\frac{d\tau}{dT}$ is larger when quality is endogenous ($\frac{d\tau}{dT}$ in (O5) is larger than $\frac{d\tau}{dT}$ in (O6)).[20]

From the vendor's first order condition for q we know that $C_q + \tilde{L}_q < 0$. When the post-patch loss effect dominates, $\frac{dq}{dT} < 0$ (from Theorem 5). Recall that $S_\tau = C_\tau + L_\tau + \tilde{L}_\tau > 0$, so that the derivative of social cost with respect to the protected period T, evaluated at T^*, $\frac{dS}{dT}\Big|_{T=T^*,\ q \text{ is exogenous}} = [C_\tau + L_\tau + \tilde{L}_\tau]\frac{d\tau}{dT} + L_T$ is negative when quality is held fixed at q^*. If T is unique, then it follows that the optimal protected period will be greater in the exogenous quality case if the post-patch loss effect dominates. Intuitively, if the post-patch loss effect dominates, then quality will increase as the protected period shrinks. Here, it makes sense to shrink the protected period because it encourages higher quality and quicker patching. However, the reverse is not true because when the patch development cost effect dominates, reducing T does bring forth a quicker patch but of lower quality. **QED**

Proof of Theorem 6

(i) Recall that ignoring the constraints, for a given T, $\tau^w = \tau(\lambda(1-\alpha), T) > \tau = \tau(\lambda, T)$. Fix a w such that $l(\tau^w) > w$. This ensures that at $T = 0$, the constraint will not bind if work-around is induced but may or may not bind in case of no work-around (i.e. $l(\tau) <> w$). As T increases, the gap between τ (and τ^w) and T decreases. Eventually, at $\hat{T}$, $l(\tau - \hat{T}) = w$ and for any $T > \hat{T}$, the constraint will not bind when a work-around is not induced. As T increases further, the gap between τ^w and T will shrink enough so that at $T = \hat{\hat{T}}$, $l(\tau^w - T) = w$. For $T > \hat{\hat{T}}$ the constraint will bind when vendor wants to induce work-around. Figure 10 explains the intuition.

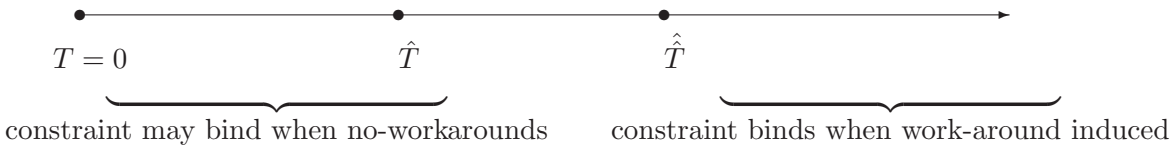


Figure 10: Work-around regions with T

For $\alpha = 0$, $\hat{\hat{T}} = \hat{T}$. Further, for any T, the patch release time required to induce users to work-around must be weakly greater than the patch release time required to induce them not to work-around, it also follows that $\hat{\hat{T}} \geq \hat{T}$. Recall that the unconstrained optimal patch release time in the work-around case is simply the τ chosen for an internalization factor equal to $\lambda(1-\alpha)$. As α increases, the patch release time chosen under work-around increases and thus $\hat{\hat{T}}$ increases as well. By definition, $\hat{T}$ does not change with α.

We now prove that T^w exists. We first show that $V^w \leq V$ for all $T < \hat{\hat{T}}$.

[20] When quality is exogenous, $\frac{d\tau}{dT} = \frac{V_{\tau T}}{V_{\tau\tau}}$. When quality is exogenous, $\frac{d\tau}{dT} = \frac{V_{\tau T} \cdot V_{qq}}{V_{\tau\tau} \cdot V_{qq} - (V_{\tau q})^2}$. A little algebraic manipulation yields the desired result.

Note that at $\alpha = 0$, $V^w \leq V$ for all $T \leq \hat{T}$. When $\alpha = 0$, V^w and V are identical except that V is the (weakly) constrained value for $T \leq \hat{T}$. If $l(\tau(\lambda), T) > w$, the constraint binds for $T < \hat{T}$ and $V^w < V$. $T \in [\hat{T}, \hat{\hat{T}}]$ (when both V and V^w are unconstrained) $V^w = V$. Since V does not depend on α, if we show that V^w decreases in α then it must be that $V^w < V$ for all $T \in [0, \hat{\hat{T}}]$. By the envelope theorem

$$\frac{dV^w}{d\alpha} = V^w_\alpha = \lambda.\Big(\int_0^T l(T-x)dF(x) + w - \int_0^T l(\tau^w - x)dF(x) - (1-F(T))(\tau^w - T) \Big)$$

Since $w < l(\tau^w - T)$ when $T < \hat{\hat{T}}$,

$$\frac{dV^w}{d\alpha} < \lambda.\Big(\int_0^T l(T-x)dF(x) + l(\tau^w - T) - \int_0^T l(\tau^w - x)dF(x) - (1-F(T))(\tau^w - T) \Big)$$

Since l in convex and $l(0) = 0$, it must be that

$$\begin{aligned}\int_0^T l(\tau^w - x)dF(x) = \int_0^T l(\tau^w - T + T - x)dF(x) &> \int_0^T l(\tau^w - T)dF(x) + \int_0^T l(T-x)dF(x) \\ &= l(\tau^w - T)F(T) + \int_0^T l(T-x)dF(x)\end{aligned}$$

Substituting in the expression of $\frac{dV^w}{d\alpha}$ above and rearranging the terms, we have $\frac{dV^w}{d\alpha} < 0$. Therefore $V^w < V$ for all $T \in [0, \hat{\hat{T}}]$ for $\alpha > 0$. The assumption that $C(\tau^\infty) < \lambda w$ ensures that $V^w > V$ for $T = \tau^\infty$. To see this note that V^w is decreasing in α. Further, at $\alpha = 1$, the vendor will not patch so that, evaluated at $T = \tau^\infty$, $V^w = \lambda\Big(\int_0^{\tau^\infty} l(\tau^\infty - x)dF(x) + w \Big)$, and $V = C(\tau^\infty) + \lambda\Big(\int_0^{\tau^\infty} l(\tau^\infty - x)dF(x) \Big)$, so that $V - V^w = C(\tau^\infty) - \lambda w$. By continuity, there is some T^w such that $V^w = V$. It also follows that $T^w \geq \hat{\hat{T}}$. If, as assumed in the text, T^w is unique, this implies that at T^w, $V^w_T - V_T > 0$ because V^w cuts V from below.

(ii) $\frac{\partial T^w}{\partial \alpha} = -\frac{V^w_\alpha - V_\alpha}{V^w_T - V_T}$. The divisor has been shown to be positive in the neighborhood of T^w. Further, $V_\alpha = 0$ and $V^w_\alpha < 0$ as shown earlier. This implies that $\frac{\partial T^w}{\partial \alpha} > 0$.

Similarly, $\frac{\partial T^w}{\partial \lambda} = -\frac{V^w_\lambda - V_\lambda}{V^w_T - V_T}$, it is sufficient to show that $V_\lambda - V^w_\lambda < 0$, when evaluated at $T = T^w$.

$$\begin{aligned}\Big(V_\lambda - V^w_\lambda \Big)\Big|_{T=T^w} &= \Big(V(\tau, T^w) - C(\tau) \Big) - \Big(V^w(\tau^w, T^w) - C(\tau^w) \Big) \\ &= C(\tau^w) - C(\tau) < 0\end{aligned}$$

QED

Proof of Theorem 7

Proof: Define $T^*(\alpha = 0)$ as the optimal protected period in the absence of smart users. $T^w(\alpha = 0) < T^*(\alpha = 0)$ and $T^w(\alpha)$ is strictly increasing in α from Theorem 6. Therefore, there exists a $\hat{\hat{\alpha}}$ such that $T^w(\hat{\hat{\alpha}}) = T^*(\alpha = 0)$. However, at any such T^w, $S^w > S$. Since, by definition

of $\hat{\alpha}$, $S^w(\hat{\alpha}) - S(\hat{\alpha}) = 0$, there exists a region $[\hat{\hat{\alpha}}, \hat{\alpha}]$ where social planner strictly prefers no work-arounds. However, in this region $T^w > T^*(\alpha = 0)$. Therefore, the social planner has to set optimal $T^* = T^w(\alpha)$ to prevent work-arounds. However, an increase in T^* so that $T^* > T^*(\alpha = 0)$, but such that there are no work-arounds can only increase social cost. **QED**