

Tax evasion with a conscience

Martin Dufwenberg^{1,2,3}  | Katarina Nordblom^{4,5} 

¹Department of Economics, University of Arizona, Tucson, USA

²University of Gothenburg, Gothenburg, Sweden

³CESifo, Munich, Germany

⁴Department of Economics and CeCAR, University of Gothenburg, Gothenburg, Sweden

⁵UCFS, Uppsala University, Uppsala, Sweden

Correspondence

Martin Dufwenberg, Department of Economics, University of Arizona, Tucson, USA.

Email: martind@eller.arizona.edu

Funding information

Swedish Research Council,
Grant/Award Numbers: 2016-01485-3,
348-2014-3770

Abstract

How do moral concerns affect fraud and detection, for example, tax compliance and the need for audits? We propose answers by exploring a psychological $2 \times 1 \times 2 \times 1$ inspection game which incorporates belief-dependent guilt, unawareness, and third-party audience effects. Novel conclusions are drawn regarding whose behavior is affected by moral concerns (it's the inspector's more than the inspectee's) as well as the policy issue whether to allow the use of a principle of public access whereby tax returns are made public information.

1 | INTRODUCTION

A clear conscience may be a good pillow (as the proverb goes). Yet, this may not be enough to deter criminal behavior. Shoplifting, fare dodging, use of narcotics, reneged arms control treaties, embezzlement, corruption, moral hazard in insurance, and tax evasion happens. Law enforcement (or contracting partners) react by engaging in costly audits, inspections, raids, and crackdowns. The interaction is studied in the literature on inspection games.¹ Mostly, no notion of integrity, virtue, or morale is invoked though. It is natural to wonder whether conclusions

¹See, for example, Avenhaus et al. (2002) for an introduction and survey.

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2021 The Authors. *Journal of Public Economic Theory* published by Wiley Periodicals LLC

change if such matters are taken into account. We explore that issue, for a stylized and special setting.

We focus on the impact of guilt. We assume that anticipated pangs of this emotion temper individuals' inclinations to break the law, if they thereby hurt (or are perceived to hurt) others. This is consistent with psychological evidence and experimental work by behavioral economists, which Battigalli and Dufwenberg (2007) (B&D) cite to motivate a formal approach to how players' guilt shapes outcomes in games. They build on the framework of so-called psychological game theory (Battigalli & Dufwenberg, 2009, 2021; Geanakoplos et al., 1989), in which players' utilities depend on beliefs (about beliefs) about choices and not only on which end node is reached (as in standard games). We apply B&D's theory to explore a 4-player $2 \times 1 \times 2 \times 1$ inspection game. There are two active players: an inspector and an inspectee, and two passive players: the inspectee's alter ego, who is unaware of a legal rule, and an observer who is hurt if the law is violated. While these additional players cannot actively choose, they are crucial to the psychological analysis via their impact on beliefs.

From now on, we mainly frame our analysis in terms of tax evasion. Many scholars in the extensive related literature (starting with Allingham & Sandmo, 1972) argued that "moral sentiments" may be crucial for understanding compliance.² Two caveats should be noted. First, as a referee pointed out, the world of tax evasion is complex and we consider but a limited special setting, involving a taxpayer's binary choice to report (or not) some extra income. We disregard many additional realistic aspects (e.g., income heterogeneity, risk preferences, and partial evasion) which the tax evasion literature often considered. Second, the insights we derive may apply more broadly than to tax evasion. The readers should keep in mind the examples we mentioned above (and perhaps in particular shoplifting and free-riding on a tram).

Following B&D, we explore two forms of moral anguish. Under "simple guilt" an evading taxpayer cares about the extent to which he *actually* hurts fellow citizens relative to what they expect. Under "guilt-from-blame" he rather worries about others' impressions regarding his *intentions* to cause such harm. Each of these sentiments may make intuitive sense and it is not clear whether one is more empirically relevant than the other. Our analysis contrasts these forms of bad conscience with each other, and with a classical world with no remorse.

Whose behavior is most affected by the presence of taxpayer guilt? It may seem intuitive that it would be the taxpayer, but in much of our analyses the amount of evasion does not depend on the degree of guilt aversion. Instead, the inspection rate is affected. It would be a mistake to portray this as a failure of guilt sentiments to reign in tax evasion. Less inspection is needed when an equilibrium is played, and this may involve savings of public funds.

We also derive a novel conclusion regarding tax policy, namely whether to make tax returns public information. The insight hinges crucially on an aspect that is idiosyncratic to psychological game theory. Namely, conclusions depend on the information structure *across end nodes*. This could never happen in traditional games. When comparing private versus public tax returns, these schemes provide different information to tax payers' neighbors. Neighbors make no choices, yet, under guilt-from-blame, a taxpayer cares about the neighbor's inferences

²See, for example, Andreoni et al. (1998), Slemrod (2019), and Alm (2019) for excellent surveys of the tax evasion literature. As regards tax morale, relevant work includes theory by Erard and Feinstein (1994), Bosco and Mittone (1997), and Eisenhauer (2008) (unlike us, these authors do not use psychological game theory), experimental studies by Coricelli et al. (2010), and Dulbeck et al. (2016), and the broad discussion by Luttmer & Singhal (2014) and Alm (2019).



regarding whether or not evasion occurs. Whether tax returns are private or public turns out to be critical.

We believe we are the first to incorporate different notions of taxpayer guilt using psychological game theory. On a broader level, our work relates to an emerging literature involving behavioral analysis of tax compliance. Hashimzade et al. (2013) review models that consider, e.g., regret, ambiguity aversion, and prospect theory. We also relate to the literature on the interplay between taxpayers and authorities and optimal policy. Rather than taking policy as given, like as Reinganum and Wilde (1985) and Graetz et al. (1986), we model a game between taxpayer and tax authority. Gahvari and Micheletto (2020), Meunier and Schumacher (2020), and Jung et al. (2021) are recent examples that study the consequences of taxpayer behavior on optimal policy.

Section 2 describes the inspection game around which our subsequent analysis centers, and derives equilibria assuming standard preferences. Section 3 introduces the two forms of guilt aversion, and studies how equilibrium behavior changes. Section 4 shows that our results from Section 3 can be given a population reinterpretation. This “robustness check” offers an alternative way to think about the results which may be realistic. Section 5 examines the effects of a principle of public access to tax returns. Section 6 offers concluding remarks.

2 | PRELIMINARIES: AN INSPECTION GAME

2.1 | The game form

While our set-up may be of relevance for a variety of inspection settings (as we discussed in the introduction), we frame it as involving a taxpayer who receives a certain income and makes the dichotomous choice of whether to declare it and pay the corresponding tax or whether to evade. This may seem restrictive and is certainly not applicable to all kinds of tax evasion. In many countries most income is third-party reported for employed individuals, so the room for tax evasion is limited.³ However, what we have in mind is some extra income, which is not automatically reported and where the taxpayer has a high degree of discretion. It is becoming increasingly frequent that people have side hustles in addition to their regular jobs.⁴ Such extra income is typically not automatically reported to the authorities, but is to be self-reported by the individual. Hence, there is room for the individual to choose whether to declare the income or not.

Our approach also endogenizes the tax authority's inspection behavior. Many previous models of tax evasion take that as given. As forcefully argued by Graetz et al. (1986), such an approach (a tradition that goes back to, Allingham & Sandmo, 1972) may be too restrictive. Authorities anticipate taxpayer behavior as much as taxpayer behavior depends on inspection rules. Arguably, this is best modeled such that a taxpayer and the authority make their moves *simultaneously*, neither party being able to condition its choice on that of the other. The game we explore is chosen to be as simple as possible subject to the constraint that it be rich enough to capture those strategic interdependencies, and some other key concerns: Player 1 is a

³See the discussion in Kleven et al. (2011).

⁴Think of gains from lotteries or poker games, an investor who has earnings from abroad, or someone renting out their apartment through Airbnb.

taxpayer who chooses to declare (D) or to evade (E) a certain income. Player 2, the tax authority, simultaneously chooses to inspect (I) or not (N).⁵

The game also includes player 3, who is a neighbor, a silent observer. The neighbor is a dummy player with no choice in the particular game, but who is crucial to the analysis once we incorporate guilt. We think of player 3 as a law-abiding citizen without the particular extra income that player 1 has.⁶ Player 3 forms beliefs about the behavior of player 1, but does not directly interact with him.

The law says that the income should be declared, but we allow that it requires awareness to know that. That some taxpayers are unaware of tax rules and uncertain about what taxes they really should pay has recently been pointed out as being important to overall compliance (see, e.g., Alm, 2019; Slemrod, 2019). Especially in relation to the so-called gig economy, many people who earn some extra income find tax rules overly complicated and do not know if or how they should declare these side-hustle incomes.⁷

To capture this, we assume that only a share $\delta \in (0, 1]$ of the taxpayer population are aware that this specific income has to be declared, while the rest are unaware and therefore (unconsciously) evade with certainty.⁸ Players 2 and 3 cannot tell whether the taxpayer is aware that the extra income has to be declared or not. They assign probabilities δ and $1 - \delta$ to either case.

Figure 1 presents the game form, with the players' material payoffs. Player 1 is the aware taxpayer while 1' (who formally speaking is a separate player with the same material payoffs as player 1) is unaware.

If 1 declares, his (material) payoff is normalized to 0, irrespective of whether he is inspected. If 1 evades, his payoff depends on whether 2 inspects: If inspected, the payoff is $-f$, which is the perceived cost of punishment. If not inspected, the payoff is instead the saved tax payment t . Payoffs for the authority when not inspecting is t if 1 declares and 0 if he evades, that is, when 1 is evading tax revenue isforegone. Inspection comes at a cost, c , irrespective of whether or not 1 is evading. Hence, inspecting an honest player 1 would yield a payoff $t - c$. Inspecting an evader, however, brings revenue $x \in [0, f]$, so that net payoff is $t - c + x$. A fine, were there no administrative transaction costs, would involve $x = f$. A jail sentence would bring no revenue, so $x = 0$. A case where $0 < x < f$ would reflect some combination, maybe including also administrative costs of conviction.⁹ If the taxpayer's experienced cost of the punishment, f , were the same, the cost to society would thus depend on the nature of punishment, that is, the value of x . To avoid solutions where it is trivially in 2's interest to always choose N , we assume that $t - c + x > 0$.

⁵Note that our approach gives more discretion to the tax authority than, for example, the model by Jung et al. (2021), where the authority receives a certain budget and inspects accordingly.

⁶Hence, we disregard more complex social interactions where individuals mutually affect each-others' evasion behavior.

⁷See, for example, Bruckner (2016) and Thayer (2020). As we shall later see, with some forms of guilt in the picture, this is crucial. If tax rules are complicated or unknown to some people, there is a possibility to "hide" behind this fact. If you "didn't know" that you had to report this particular income, guilt may be reduced since you had no intention to cheat.

⁸One example of unawareness could be the "d-type agents" in Hokamp and Pickhardt (2010), who want to be honest, but fail due to the complexity of the tax system. Alternatively, it could be more straightforward ignorance, such as if a poker player mistakenly assumes that his earnings are not subject to taxation. Allowing the unaware to evade with a certain probability less than one (but still quite high) would not alter the qualitative results.

⁹One could, in principle, even assume a negative payoff, $x < 0$. Such an extension complicates the analysis with very little added to the results, so we limit our analysis to the nonnegative values $x \in [0, f]$.

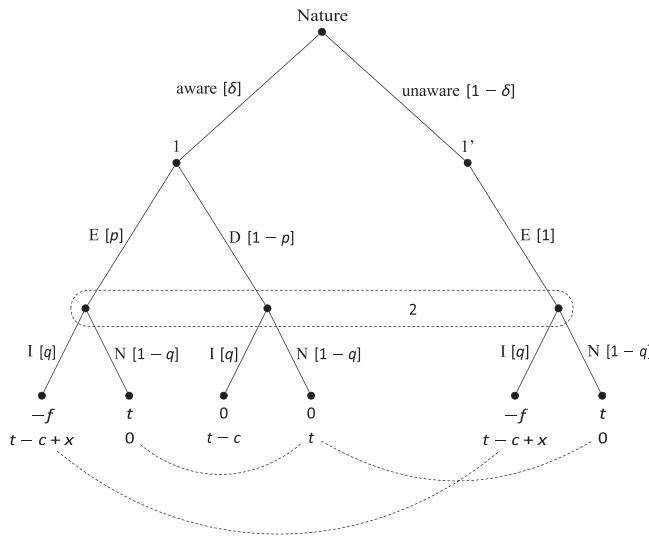


FIGURE 1 The tax compliance game

What about player 3's payoff? Despite that the neighbor is a dummy player, his payoff and information is crucial to our analysis. We assume that players 2 and 3 have the same payoff. The interpretation is that player 2's payoff reflects in a meaningful way public resources. Player 3, being a representative agent for the public, is assumed to care about those just like 2 does. Alternatively put, 2's payoffs and incentives are designed to align with the payoffs of the public. Since the payoff of players 2 and 3 coincide, we collapse them into one (the second) line in Figure 1.¹⁰

The information sets indicated at 3's payoffs at the end nodes describe what we assume that 3 knows about preceding choices. He does not have perfect information. He cannot tell whether his neighbor is aware or not, and if 2 does not inspect 3 cannot tell whether or not 1 evaded. The latter feature reflects an assumption that tax returns are private. Unless 1 is inspected, 3 cannot tell whether or not he committed a crime (although 3 will form beliefs). In Section 5, when we assume public tax returns, we consider a new game that modifies 3's information relative to Figure 1. The end-node information of 1, 1', and 2 is immaterial to our analysis, so we do not specify that.

2.2 | Welfare

To compare equilibria and evaluate policy, it is useful to define a notion of whether outcomes are good or bad. The game parameters do not offer clear guidance. For example, t reflects a transfer between players and it is not obvious why and how welfare is enhanced by such a payment. And while incurring an inspection cost c may be seen as pure waste, it is not a comprehensive account of resources spent on tax collection, as fixed costs are neglected. Yet another alternative would be to consider the expected payoff of player 2 (and 3), as this may

¹⁰Making player 3's payoff some fixed proportion of 2's would not alter the analysis in any qualitative way.

reflect public resources. However, as will be made clear along the analysis, equilibrium expected payoff will be constant and unaffected by most policies. Hence, such welfare criterion would give little guidance for policy evaluation. We propose that the relevant way to assess welfare rather goes via the probabilities p and q with which players 1 and 2 evade and inspect, respectively. The lower these numbers are the better, *ceteris paribus*:

- Our game concerns a society where a democratic process decided that citizens' payment of taxes is a duty, while tax evasion is criminal. The benefits of paid taxes likely include societal gain of tax-financed public goods. The costs of evasion may be more nebulous. Perhaps if tax evasion is widespread this undermines civic morale more generally, such that people litter, steal, or engage in corruption more willingly? We shall not try to quantify exactly these societal gains and losses, but rather just take the view that a lower p tends to be a good thing.
- Tax inspection is costly. Beyond the costs-per-inspection, c , a background institutional structure is needed: buildings, administrators, lawyers, prosecutors, policemen. The less inspection is going on, the less costly the needed apparatus. Hence, the lower q is the better.

2.3 | Solution with classical preferences

We solve for the equilibrium probabilities p^0 and q^0 with which players 1 and 2, respectively, evade and inspect. If $\delta = 1$ (2 encounters the aware 1 for sure), we have an inspection game with standard properties: If 2 knew that 1 declared it would not inspect, since $c > 0$. If 1 knew that 2 did not inspect, he would evade, since $t > 0$. If 2 knew that 1 evaded it would inspect, since $t - c + x > 0$. If 1 knew that 2 inspected, he would declare, since $0 > -f$ etc. Hence, there would be no Nash equilibrium in pure strategies, but we can solve for the (unique) equilibrium in mixed strategies. However, since we allow that $\delta \leq 1$, there is room for pure strategy Nash equilibria under some circumstances.

It will turn out that a crucial factor is how profitable an inspection is for the tax authority. Let $R = \frac{t-c+x}{t+x} < 1$ denote the “revenue ratio:” how large the net revenue of catching an evader is in relation to the gross revenue. The higher is R , the more “efficient” is the inspection of an evader.

To construct a pedagogical narrative, we present our results as a series of propositions. We intend to signal economic importance, not mathematical complexity. To enhance readability for those who wish to concentrate on implications more than logic, we relegate all proofs to the appendix.

Proposition 1.

- If $\delta < R$, the game described in this section has the unique equilibrium (D, I) for all parameter constellations.*
- If $\delta > R$, the game described has a unique mixed-strategies equilibrium for all parameter constellations:*

$$p^0 = 1 - \frac{R}{\delta}. \quad (1)$$



$$q^0 = \frac{t}{t+f}. \quad (2)$$

The feature that if $\delta < R$ the equilibrium is (D, I) is true for all specifications, irrespective of guilt aversion: With a sufficiently high proportion of unaware taxpayers who evade with certainty, 2 always finds it worthwhile to inspect. This is understood by 1 who, therefore, chooses to declare, an obvious implication. The reason why we incorporate δ in our model, however, is to highlight effects that are relevant when that case it not at hand. Therefore, in what follows, we mostly focus on case (ii) of Proposition 1, where $\delta > R$.¹¹

Keeping in mind the “classical results” from for example, Becker (1968) and Allingham and Sandmo (1972), Proposition 1 (ii) may appear counterintuitive. The cost to the taxpayer of getting caught, f , has no effect on the evasion probability, but on the authority’s audit probability. While this effect is not the usual in the tax-compliance literature, where the taxpayer is regarded as the only active part, it is a typical finding for inspection games as presented, for example, in Graetz et al. (1986).¹² When inspector and inspectee move simultaneously, the unique and mixed equilibrium implies that the two players, so to say, hold each other indifferent. Since a more severe punishment makes evasion less attractive, the inspection probability has to go down in order for the taxpayer to remain indifferent between evading and declaring. The effects of harsher punishment are manifested, not in terms of less crime, but rather in terms of less inspection.

Note that we have not mentioned player 3 at all. Rightly so, since 3 has no bearing on anything under classical preferences.

3 | INCORPORATING GUILT

The filing of tax returns is an example where guilt plausibly influences behavior. By withholding provision of public funds, tax evaders may hurt fellow citizens who expect compliance. Conscientious filers dislike that and may declare honestly to avoid guilt. This is in line with findings in psychology. In an influential study, (Baumeister et al., 1994, p. 247) explain that “If people feel guilt for hurting their partners ... and for failing to live up to their expectations, they will alter their behavior (to avoid guilt) in ways that seem likely to maintain and strengthen the relationship.”¹³ Note the link to others’ expectations. Using designs that elicit beliefs about beliefs, several experimental studies tested for such belief-dependent motivation and found support.¹⁴ B&D develop two models—*simple guilt* and *guilt from blame*—that describe how

¹¹If $\delta = R$, then equilibrium requires $p^0 = 0$ and $q^0 \in \left[\frac{t}{t+f}, 1 \right]$.

¹²The broader and somewhat counterintuitive game-theoretic angle is that as regards mixed strategy equilibrium a change in one player’s payoff is predicted to only affect the other player’s behavior. Camerer (2003, pp. 139–140) discusses lab experiments that tested this prediction and reports that “This wacky prediction is surprisingly close to correct.” We also note that there is large experimental literature on -oetax evasion games- which largely does not shed light on the issue discussed here since most of the games studied are not inspection games but rather treat the inspection probability as a parameter; see Alm and Malézieux (2021, especially p. 713).

¹³Compare also Baumeister et al. (1995) and Tangney (1995).

¹⁴For example, Dufwenberg and Gneezy (2000) show for a dictator game that more is given by subjects who expect their co-players to expect a lot. See also Dufwenberg and studies have met some criticism—see Ellingsen et al. (2010) and Vanberg (2008)—and some follow-up defense—see for example, Khalmetski et al. (2015). Cartwright (2019) surveys much of the literature.

belief-dependent guilt affects interaction in games. We introduce, adapt, and apply these to our setting.

3.1 | Simple guilt

We solve for an equilibrium where 1 evades with probability p^{SG} and 2 inspects with probability q^{SG} . 2's payoff function is as before, but 1's utility is different. Before approaching our specific game, we reproduce the following:

3.1.1 | Elements from B&D

B&D consider finite extensive game forms. Consider such a game form and let Z be its set of endnodes. Players' material payoffs are given by functions $\mathbf{m}_i : Z \rightarrow \mathbb{R}$, $i \in N$. A typical material payoff is denoted by m_i as in $m_i = \mathbf{m}_i(z)$. Function \mathbf{m}_i need not represent i 's utility, which is given by a different and belief-dependent function as we will now explain. Let H_i be the set of player i 's information sets, including over endnodes. Conditional on $h \in H_i$, i holds a conditional first-order belief $\alpha_i(\cdot|h) \in \Delta(S_{-i})$ about co-players' strategies, where $S_{-i} = \times_{j \neq i} S_j$ with S_j player j 's set of pure strategies. $\alpha_i = (\alpha_i(\cdot|h))_{h \in H_i}$ is i 's system of first-order beliefs, and i holds a second-order belief $\beta_i(h)$ about the first-order belief system α_j of each $j \neq i$, a third-order belief $\gamma_i(h)$, etc. Assume that higher-order beliefs are degenerate point beliefs (anticipating their use in defining equilibrium) and so identify $\beta_i(h)$ with an array of conditional first-order beliefs $\alpha_{-i} = (\alpha_j(\cdot|h'))_{j \neq i, h' \in H_j}$, etc. The beliefs i would hold at $h \in H$ satisfy Bayes' rule and common certainty of that (cf. B&D). Given $s_j \in S_j$ and initial (= at the root, labeled h^0) first-order beliefs $\alpha_j(\cdot|h^0)$ player j forms a material payoff expectation: $\mathbb{E}_{s_j, \alpha_j}[m_j|h^0] = \sum_{s_{-j}} \alpha_j(s_{-j}|h^0) \mathbf{m}_j(\mathbf{z}(s_j, s_{-j}))$, where $\mathbf{z}(s_j, s_{-j})$ is the end node reached if (s_j, s_{-j}) is played. For any end node z consistent with s_j , define $D_j(z, s_j, \alpha_j) = \max \{0, \mathbb{E}_{s_j, \alpha_j}[m_j|h^0] - \mathbf{m}_j(z)\}$; this measures how much j is let down. If i knew z , $s_{-i} \in S_{-i}$, and $\alpha_j(\cdot|h^0)$, he could calculate how much of $D_j(z, s_j, \alpha_j)$ is due to his behavior: $G_{ij}(z, s_{-i}, \alpha_j) = D_j(z, s_j, \alpha_j) - \min_{s_i} D_j(\mathbf{z}(s_i, s_{-i}), s_j, \alpha_j)$. He is affected by simple guilt if his utility function u_i^{SG} has the form

$$u_i^{SG}(z, s_{-i}, \alpha_{-i}) = \mathbf{m}_i(z) - \sum_{j \neq i} \theta_{ij} G_{ij}(z, s_{-i}, \alpha_j); \quad (3)$$

s_{-i} is consistent with reaching z and $\theta_{ij} \geq 0$ is i 's guilt sensitivity wrt j .

The utility function u_i^{SG} is belief-dependent in that α_{-i} —the beliefs of other players—appears as an argument. At endnode z , player i receives utility both from his material payoff and the guilt associated with z . The material payoff $\mathbf{m}_i(z)$ is straightforward to interpret; it depends only on z . The guilt part is belief-dependent: i computes the guilt felt toward each player j at z as the gap (if positive and due to i) between what j initially expected to get and what j actually gets at z . Player i cannot observe j 's beliefs directly, but he will form beliefs about j 's beliefs and then maximize the expectation of u_i^{SG} .



3.1.2 | Our solution

In our application below, it is player 1 who computes the guilt felt toward player 3. In this connection, note that the gap between what 3 initially expected to get and what 3 actually gets may not be known by 3 himself, because 3 may not observe his own payoff (as we explained in Section 2). We propose that the way we compute 1's guilt nevertheless makes psychological sense, if 3's payoff gaps are interpreted as part of a mental exercise by 1 rather than something that is actually experienced by 3.¹⁵

We adopt the notions from B&D to our setting and (3) to u_1^{SG} , that is, player 1's utility. However, if (3) were taken literally, if $x > 0$ there would be let-down also at the end node reached by (D, I) : the neighbor would be disappointed and the taxpayer would feel guilty for denying the government a fine by not being dishonest! In a tax evasion context that is contrived, (3) describes the utility at each end node. We only apply it *at end nodes where 1 evaded*.¹⁶ With that, and assuming that $\theta_{13} = \theta \geq 0 = \theta_{ij} \forall (i, j) \neq (1, 3)$, the taxpayer has the following expected utility if he evades:

$$\begin{aligned}\mathbb{E}[u_1^{SG}(E)] &= q^{SG}[-f - \theta \max\{0, Rt - (t - c + x)\}] + (1 - q^{SG})[t - \theta Rt] \\ &= -q^{SG}f + (1 - q^{SG})[t - \theta Rt],\end{aligned}\tag{4}$$

where Rt is the expected payoff for player 3 (and 2).¹⁷ Hence, the only end-node where player 3 is let down by 1 who senses simple guilt (actual payoff is lower than expected) is where 1 evades and 2 does not inspect. Since no simple guilt is sensed when declaring, $u_1^{SG}(D) = 0$.

Proposition 2. *Assume that $\delta > R$. The game with simple guilt has a unique equilibrium for all relevant parameter constellations.*

- (i) *If $\theta < \frac{1}{R}$ the equilibrium will be in mixed strategies, where p^{SG} and q^{SG} are determined by (5) and (6), respectively:*

$$p^{SG} = p^0 = 1 - \frac{R}{\delta},\tag{5}$$

$$q^{SG} = \frac{t - \theta Rt}{t + f - \theta Rt}.\tag{6}$$

- (ii) *If $\theta > \frac{1}{R}$ the equilibrium is (D, N) .*

¹⁵We thank a referee for calling attention to and encouraging us to comment on this subtle issue.

¹⁶Should B&D be criticized for not capturing this in their original model? We do not think so. They presented and worked with what is already more than a mouthful: a large class of game forms including monetary payoffs. Applied work will couple any psychological game from B&D's class with additional economic assumptions, regarding whether an action is to be interpreted as tipping, or stealing, evasion, etc., and there may be relevant further information concerning, for example, torts, norms, property rights, and liability clauses. These details may matter for specifying the limits of guilt, and applied behavioral researchers need to be smart about this as they consider their specific application.

¹⁷See (A6) in Appendix A.

While $p^{SG} = p^0$ irrespective of θ , we see that the inspection probability is decreasing in the guilt sensitivity θ in the mixed-strategy equilibrium (note that if $\theta = 0$ then $q^{SG} = q^0$). For high enough values of θ , the taxpayer will always prefer D to E irrespective of 2's actions and we reach the pure-strategy equilibrium with full compliance without any inspection.¹⁸

3.2 | Guilt from blame

We solve for an equilibrium where 1 evades with probability p^{GB} and 2 inspects with probability q^{GB} . 2's utility remains as before, but 1's is different. Before approaching our specific game, we adapt the following:

3.2.1 | Elements from B&D

Given s_i and initial beliefs $\alpha_i(\cdot|h^0)$ and $\beta_i(h^0)$, compute how much i expects to let j down:

$$G_{ij}^0(s_i, \alpha_i, \beta_i) = \mathbb{E}_{s_{-i}, \alpha_{-i}, \beta_{-i}}[G_{ij}|h^0] = \sum_{s_{-i}} \alpha_i(s_{-i}|h^0) G_{ij}(\mathbf{z}(s_i, s_{-i}), s_{-i}, \beta_{ij}^0(h^0)), \quad (7)$$

where $\beta_{ij}^0(h^0)$ denotes the initial (point) belief of i about $\alpha_j(\cdot|h^0)$. Suppose z is reached. $\mathbb{E}_{\alpha_j, \beta_j, \gamma_j}[G_{ij}^0|h]$, where $z \in h$, measures j 's inference regarding how much i intended to let j down, or how much j "blames" i conditional on $H_j(z)$. Player i is affected by guilt from blame if he dislikes being blamed: his preferences are represented by utility function u_i^{GB} of the form

$$u_i^{GB}(z, \alpha_{-i}, \beta_{-i}, \gamma_{-i}) = \mathbf{m}_i(z) - \sum_{j \neq i} \theta_{ij} \mathbb{E}_{\alpha_j, \beta_j, \gamma_j}[G_{ij}^0|H_j(z)]. \quad (8)$$

The utility function u_i^{GB} is belief-dependent in that β_{-i} – the second-order beliefs of other players – appears as an argument. Compare the intuition underlying u_i^{GB} with that of u_1^{SG} . Whereas i sensed simple guilt toward j if he actually let j down relative to j 's initial beliefs, with guilt from blame it rather matters to i what is j 's impressions regarding i 's willingness to let j down.

3.2.2 | Our solution

Again assuming that $\theta_{13} = \theta \geq 0 = \theta_{ij} \forall (i, j) \neq (1, 3)$, we extend these ideas to our setting. We make one adjustment relative to B&D. As seen above, the expectation $G_{ij}^0(s_i, \alpha_i, \beta_i)$ given by (7) is derived using $\alpha_i(\cdot|h^0)$ and $\beta_i(h^0)$; these are beliefs at the root rather than following nature's choice *aware*. Given our interpretation that 1 and 1' are different persons, it makes more sense to derive $G_{13}^0(s_1, \alpha_1, \beta_1)$ substituting $\alpha_1(\cdot|h)$ and $\beta_1(h)$ for $\alpha_1(\cdot|h^0)$ and $\beta_1(h^0)$, where h is the

¹⁸If $\theta = \frac{1}{R}$, then equilibrium requires $p^{SG} \in [0, p^0]$ and $q^{SG} = 0$.



information set where 1 makes his choice rather than the root. That is, we compute how much 1 expects to let 3 down, *when he makes his choice*.

As in Section 3.1, we maintain that 1 only lets 3 down when he evades without being caught and punished. In the end-node where 1 evades and 2 does not inspect the actual let-down is Rt , which in case of evasion, occurs with probability $(1 - q)$. However, 3 cannot be sure whether 1 chose E or D if 2 did not inspect. This matters under guilt from blame, which is sensed to the degree that 3 expects 1 to have the intention of letting him down.

Following B&D's formulas, in our context, we must calculate the probability that 1 is blameworthy according to 3. We can do this by noting that, as seen in Figure 1, player 3 has three information sets:

- $\{(aware, D, I)\}$ —player 2 inspects finding a declaring 1, so 3 knows that 1 did not intend to let 3 down (as 1 chose D). 1 is not blamed.
- $\{(aware, E, I), (unaware, E, I)\}$ —2 inspects catching a tax evader, who could either be aware or unaware; only the former is blamed by 3 (because 1' had no choice but E , and so could not have intended to let 3 down). In equilibrium, by Bayes' rule, choice E was made by blameworthy 1 (rather than 1') with probability $\mu = \frac{\delta p^{GB}}{\delta p^{GB} + (1 - \delta)}$.
- $\{(aware, D, N), (aware, E, N), (unaware, E, N)\}$ —2 does not inspect. The taxpayer could be an aware (player 1) or unaware (player 1') evader, or an aware declarer, only the first of whom is blameworthy. In equilibrium, by Bayes' rule, the probability of an aware (blameworthy) evader is $\lambda = \frac{\delta p^{GB}}{\delta p^{GB} + \delta(1 - p^{GB}) + (1 - \delta)} = \delta p^{GB}$.

The blame is the product of the expected let-down in case of evasion, $(1 - q)Rt$ and the probability that the taxpayer actually is a blameworthy evader (0 , λ , or μ depending on the information set of 3 as just described). The expected blame in case of no inspection is thus $\lambda(1 - q)Rt$ whether evading or not and in case of inspection $\mu(1 - q)Rt$ if 1 evaded. If 1 declares and is inspected there is no blame at all. Hence, if the taxpayer declares, his expected utility is

$$\mathbb{E}[u_1^{GB}(D)] = q^{GB}0 + (1 - q^{GB})[0 - \theta\lambda(1 - q^{GB})Rt], \quad (9)$$

and in case he evades it is

$$\mathbb{E}[u_1^{GB}(E)] = q^{GB}[-f - \theta\mu(1 - q^{GB})Rt] + (1 - q^{GB})[t - \theta\lambda(1 - q^{GB})Rt]. \quad (10)$$

Under simple guilt, if θ were high enough $(\theta > \frac{1}{R})$ then the taxpayer would not evade regardless of the authority's action, and the equilibrium would be (D, N) . A striking insight is that this result does not have a counterpart under guilt from blame:

Proposition 3. (D, N) cannot be an equilibrium under guilt from blame, regardless of θ .

Intuitively, if (D, N) were an equilibrium, then 3, on observing information set $\{(aware, D, N), (aware, E, N), (unaware, E, N)\}$, would infer that the probability of $(aware, E, N)$ equals 0. Hence 3 would not blame regardless of 1's choice, so 1 would be safe to evade.

The equilibrium under guilt from blame rather looks as follows.

Proposition 4. *Assume that $\delta > R$. The game with guilt-from-blame has a unique equilibrium for all parameter constellations, where p^{GB} and q^{GB} are determined by (11) and (12), respectively:*

$$p^{GB} = p^0 = 1 - \frac{R}{\delta}. \quad (11)$$

$$q^{GB} = \frac{t}{t + f + \theta t A (1 - q^{GB})}, \quad (12)$$

where $A = \frac{R[c - (t + x)(1 - \delta)]}{c} \leq R$.

Note that if $\theta = 0$, then $q^{GB} = q^0$. Moreover, if $\theta > 0$, then q^{GB} is defined implicitly, as it appears in each side of (12). As noted in the proof (in the appendix) it is straightforward to verify that (12) has a unique solution $q^{GB} \in (0, 1)$. Hence, with guilt from blame, no pure strategy equilibrium is attainable, no matter how strong the guilt sensitivity, θ . Under simple guilt, guilt is only sensed in (E, N) . With guilt from blame 1 instead senses guilt whenever 3 may have a reason to believe that 1 had the intention to let 3 down, whether or not he actually did so. Hence, the only end node where 1 senses no guilt from blame is when he is inspected and found to be declaring (strategy profile (aware, D, I)). If not inspected, 1 senses the same guilt whether or not he evaded since the inference about his intention is the same.¹⁹

3.3 | Comparisons

Comparing results, we note that the taxpayer evades with the same probability irrespective of the guilt sensitivity, that is, $p^{GB} = p^{SG} = p^0$. The difference induced by guilt aversion, is seen via the inspection probability, which is lower in the equilibria with guilt, and especially with simple guilt:²⁰

Proposition 5. *Whenever $\theta > 0$, then $q^0 > q^{GB} > q^{SG}$.*

Under guilt from blame the cause of a bad conscience is the perceived intention to let down rather than the actual let-down. If 2 chooses not to inspect, then 3 believes with probability λ that 1 evaded. Hence, unlike the case with simple guilt, the taxpayer cannot fully avoid a bad conscience by choosing D . Therefore, the inspection probability which keeps the taxpayer indifferent between evading and not is higher under guilt from blame than under simple guilt.²¹

¹⁹ Does 1' experience guilt from blame? We propose that this is the case if and only if 1' is inspected, since presumably that is the only outcome which makes 1' aware that he is suspected of intentional evasion. Note, however, that this is all irrelevant as regards how the game is solved, since 1' is not an active player.

²⁰ Note that if $\theta = 0$ then $q^{GB} = q^{SG} = q^0$.

²¹ One extension would be to let player 1 receive the extra income with a probability less than one. While both q^0 and q^{SG} would remain unaffected by that assumption, q^{GB} would be lower than if player 1 receives the income with certainty. This is the reason: The probability by which the aware taxpayer receiving the income evades increases to

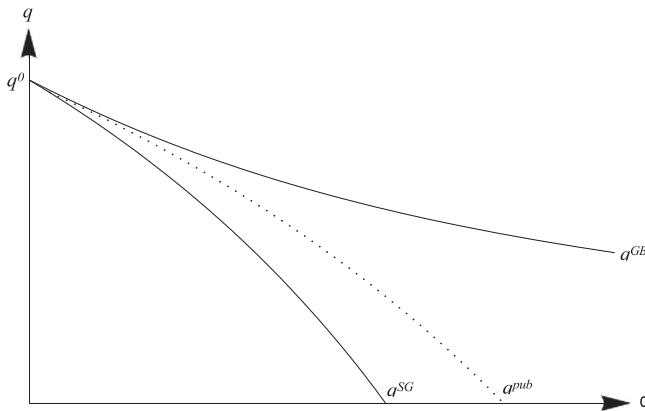


FIGURE 2 Comparison of inspection probabilities

In Figure 2, the solid lines show q^{GB} and q^{SG} as functions of θ . (The dotted line will be addressed in section 5.)

4 | A POPULATION INTERPRETATION

An arguably more realistic setting than the tax authority interacting with a single taxpayer is that there is a population of taxpayers. We now show that our results carry over to an economy populated by taxpayers heterogeneous in terms of guilt aversion. The taxpayer population is normalized to unity with a share $\delta > R$ being aware. We show that under a wide range of assumptions concerning the distribution of θ , our previous propositions still hold with the interpretation that the evasion probability p^0 (the probability that an arbitrary aware taxpayer evades) is population-wide. Perhaps this interpretation is more intuitive: The authority inspects with a probability that makes a certain share of taxpayers evade—a share that makes the authority indifferent between inspecting and not. In Section 4.1, we assume that everyone has the same type of guilt aversion (simple guilt or guilt from blame), but that θ differs between individuals. In Section 4.2 we let one share of the aware population be motivated by guilt from blame and the rest by simple guilt.

4.1 | A distribution of θ 's

Let us start with a class of distributions that makes the analysis especially straightforward. Assume that the taxpayers' θ 's are continuously distributed between 0 and $\frac{1}{R}$. Then there always exists a unique equilibrium, which is separating, where player 2 chooses an inspection probability $q \in (0, 1)$ such that there exists a $\theta^* \in (0, \frac{1}{R})$ where those with $\theta_i < \theta^*$ evade and those with $\theta_i > \theta^*$ declare with certainty.

keep the authority indifferent between inspecting and not. Hence, if an evader is caught, the probability μ that 3 blames 1 is increased and therefore q^{GB} is decreased.

We have previously shown that, for each $\theta \in [0, \frac{1}{R}]$, in equilibrium, there exists a unique inspection probability $q \in (0, \frac{t}{t+f})$ that makes an aware taxpayer indifferent between evading and declaring under simple guilt and guilt from blame, respectively (for q higher (lower) than the equilibrium value, the taxpayer would strictly prefer to declare (evade)).

In a population with continuously distributed θ 's, the value \hat{q} that makes a taxpayer with $\theta_i = \hat{\theta}$ indifferent between E and D , results in a share $F_\theta(\hat{\theta})$ evading and $1 - F_\theta(\hat{\theta})$ declaring, where F_θ denotes the cumulative density function of the guilt-aversion parameter.

How large a share of the aware taxpayers have to evade for the tax authority to be indifferent between inspecting and not? Analogously to Proposition 1(ii), the expected payoff from I has to equal that from N , that is, $t - c + x(1 - \delta + \delta F_\theta(\theta)) = t\delta(1 - F_\theta(\theta))$, for player 2 to be indifferent between I and N . This results in the equilibrium cdf, $F_\theta(\theta^*)$, which is analogous to probability p^0 in Proposition 1(ii):

$$F_\theta(\theta^*) = 1 - \frac{R}{\delta}. \quad (13)$$

Just as for one single taxpayer, this share evades in equilibrium, irrespective of the form of guilt aversion (simple guilt or guilt from blame).

To make the share $F_\theta(\theta^*) = 1 - \frac{R}{\delta}$ evade in equilibrium, q must be chosen so as to make a taxpayer holding $\theta_i = \theta^*$ indifferent between E and D . Under simple guilt, this inspection probability is analogous to the one in (6) in Proposition 2(i):

$$q_{het}^{SG} = \frac{t - \theta^* R t}{t + f - \theta^* R t}, \quad (14)$$

where $\theta^* \in (0, \frac{1}{R})$ is determined by (13). We know from Proposition 2 that $q_{het}^{SG} \in (0, \frac{t}{t+f})$ and since θ follows a continuous distribution (with zero mass at any specific $\hat{\theta}$), the equilibrium is separating where the share $F_\theta(\theta^*) = 1 - \frac{R}{\delta}$ evade and the rest declare.

As for guilt from blame, player 2 would choose q_{het}^{GB} implicitly given by

$$q_{het}^{GB} = \frac{t}{t + f + \theta^* t A (1 - q_{het}^{GB})}, \quad (15)$$

which is analogous to (12). Hence, as shown in Proposition 4, $q_{het}^{GB} \in (0, \frac{t}{t+f})$ is the unique inspection probability that makes taxpayers holding $\theta_i = \theta^*$ indifferent between E and D , implying that the share $F_\theta(\theta^*) = 1 - \frac{R}{\delta}$ evade.

Hence, most of the results previously derived for one taxpayer carry over to the situation with a population that differs in their strength of guilt aversion. Instead of one individual evading with probability $p^0 = 1 - \frac{R}{\delta}$, a share $F_\theta(\theta^*) = 1 - \frac{R}{\delta}$ of taxpayers evade and the rest declare. We have shown that this result holds for continuously distributed θ 's such that $\theta_i \in (0, \frac{1}{R}) \forall i$.

What if we loosen our restrictions that $\theta_i \in (0, \frac{1}{R})$ is continuously distributed?



It can easily be verified that the separating equilibria reached for continuous distributions may be reached for any distribution as long as θ^* fulfils the requirements that $p_\theta(\theta^*) = 0$ and $F_\theta(\theta^*) = 1 - \frac{R}{\delta}$. Such equilibria need not imply unique θ^* and q , though.

Proposition 2(ii) states that under simple guilt, there is an equilibrium in pure strategies between tax authority and a single taxpayer for $\theta > \frac{1}{R}$. What if we allow $\theta_i > \frac{1}{R}$ in our population-wide analysis? As long as the equilibrium $\theta^* < \frac{1}{R}$ the above equilibrium would not change. If $\theta^* \geq \frac{1}{R}$ we would, though, have an equilibrium where player 2 would not inspect and where those with $\theta_i < \frac{1}{R}$ would evade and those with $\theta_i > \frac{1}{R}$ would declare.²²

For a distribution with non-zero mass at θ^* there would not exist a separating equilibrium, but those with $\theta_i = \theta^*$ evade with some probability $p \in (0, 1)$, while those with $\theta_i < \theta^*$ evade with certainty and those with $\theta_i > \theta^*$ declare with certainty, yielding an over-all probability of evasion among aware taxpayers equal to $p^0 = 1 - \frac{R}{\delta}$. In all other respects, the above analysis is valid.

4.2 | Heterogeneity in type of guilt aversion

Now, assume all aware taxpayers have the same $\theta \in (0, \frac{1}{R})$, but a share β are motivated by guilt from blame and $(1 - \beta)$ by simple guilt. Proposition 5 showed that, for any $\theta > 0$, the inspection probability making a taxpayer indifferent between evading and declaring, $q^{GB} > q^{SG}$. However, player 2 can only choose one q , which means that the two types will act differently in equilibrium. Denote p^{SG} and p^{GB} the probabilities by which the two types evade. It still holds that the over-all evasion probability has to be $p^0 = 1 - \frac{R}{\delta}$ for player 2 to be indifferent between inspecting and not, which implies that $\beta p^{GB} + (1 - \beta)p^{SG} = 1 - \frac{R}{\delta}$. Since a taxpayer will only evade with a probability $p \in (0, 1)$ if indifferent between E and D , at most one of the two types will choose $p \in (0, 1)$ and at least one will choose $p \in \{0, 1\}$. There are then three potential equilibria, and which is chosen depends on the size of β :

- i) 2 chooses $q = q^{SG}$ as derived in (6) to keep the simple-guilt individuals indifferent. Then $p^{GB} = 1$ and $p^{SG} = 1 - \frac{R}{\delta(1-\beta)} < p^0$ where $p^{SG} > 0$ requires that $\beta < 1 - \frac{R}{\delta}$.
- ii) 2 chooses $q = q^{GB}$ as derived in (12) to keep the guilt-from-blame individuals indifferent. Then $p^{SG} = 0$ and $p^{GB} = (1 - \frac{R}{\delta})/\beta > p^0$. $p^{GB} < 1$ requires that $\beta > 1 - \frac{R}{\delta}$.²³
- iii) 2 chooses some $q \in (q^{SG}, q^{GB})$. This gives a separating equilibrium where all guilt-from-blame individuals evade and all with simple guilt declare, that is, $\beta 1 + (1 - \beta)0 = 1 - \frac{R}{\delta}$. Obviously, this separating equilibrium will arise iff $\beta = 1 - \frac{R}{\delta}$.

²²Under guilt from blame, Proposition 3 states that we never encounter an equilibrium in pure strategies, so allowing for $\theta > \frac{1}{R}$ would not alter the analysis and we would always be able to find a q_{het}^{GB} such that one share of the aware taxpayers evade and the rest declare with certainty.

²³When regarding taxpayers motivated by guilt from blame, we need to reconsider the inference that neighbors make concerning the intention to let them down, μ and λ , when they make inference in a population context. On the one hand, the probability that an aware taxpayer evades with a positive probability is reduced (β as compared to 1), but on the other hand, the probability by which evasion takes place is higher ($\frac{p^0}{\beta}$ rather than p^0). In equilibrium, the two effects cancel out leaving μ and λ unchanged and thereby q^{GB} is the same as in (12).

5 | PRIVATE VERSUS PUBLIC RETURNS

With guilt in the picture, novel policy implications may emerge. In this section we consider what happens if one makes tax returns public. We again assume that the tax authority interacts with one single taxpayer, but keep in mind that the results can easily be given a population interpretation.

Lately, scholars have taken interest in the potential effects from publicly disclosing tax-returns from individuals and firms (see, e.g., Bø et al., 2015; Hoopes et al., 2018). Up till now, we treated the neighbor (player 3) as unable to distinguish whether or not the taxpayer chose to evade, unless there is an inspection. In countries where income-tax returns are not public information this is probably a fair assumption. Neighbors then only have the information which is disclosed by the authorities and as long as a certain tax return is not inspected, potential evasion remains a secret. In some countries, however, there is a principle of public access to official records, which also applies to tax returns.²⁴ This allows anyone to get information about incomes declared and taxes paid by anyone else, even if the tax authority does not inspect. Hence, a neighbor who observes someone's fancy car or luxurious lifestyle can check whether they have actually declared this income and if not, they can draw the inference that they are likely living next door to a tax evader.

What difference does it make whether neighbors (to whom one may sense guilt) can retrieve information about the income declared and can we determine whether public or private tax returns are preferred from a welfare point of view? The point with public access to tax returns is that people would be more reluctant to evade when their neighbors are able to check up on them. For this to matter, taxpayers need to care about what others think about them. Neither with classical preferences nor under simple guilt is this the case. However, guilt from blame depends on 3's inference about 1's intentions to let 3 down, which depends on 3's information. Hence, the notion of private versus public tax returns is highly relevant, and the rest of the section deals with this case. We mostly focus on the inspection probability since in the mixed-strategy equilibrium the probability of evasion is still determined by (1), unaffected by guilt or the neighbor's information.

Under the principle of public access, player 3 has four information sets, as shown in Figure 3. Comparing with Figure 1, the difference is that player 2 does not have to inspect in order for 3 to be able to learn whether player 1 evaded or not. Hence, for 3, the only unknown is whether an evader is aware or unaware, that is, whether the evasion is blameworthy or not.²⁵ Since there is no uncertainty regarding whether evasion actually took place, the uninspected honest taxpayer senses no guilt from blame toward 3. This is the crucial difference to the case with private tax returns, analyzed in Section 3.2. With public tax returns, the taxpayer's expected utilities are different. Utility when declaring is 0, just as under simple guilt, since 3 does not suspect evasion. If evading, the expected utility is different from that with private returns in (10). Since 3 knows that 1 evades also when 2 does not inspect, probability μ is assigned to 1 being blameworthy. When 2 inspects with probability q^{pub} , 1's expected utility in case of evasion thus becomes

²⁴Norway, Sweden, Finland, and Iceland all have public disclosure of personal tax returns to some extent. In Norway the tax authority provides the information online, free to access by anyone. This policy has been studied by, for example, Bø et al. (2015) and Perez-Truglia (2020). In Sweden, anyone can call the authority to get the information for free. There are also private actors who sell online information and the "taxation calendar," where incomes and taxes of ordinary people, high-income earners and celebrities are listed, is a yearly bestseller.

²⁵Remember that the unaware does not intend to let anyone down, but evades due to lack of knowledge.

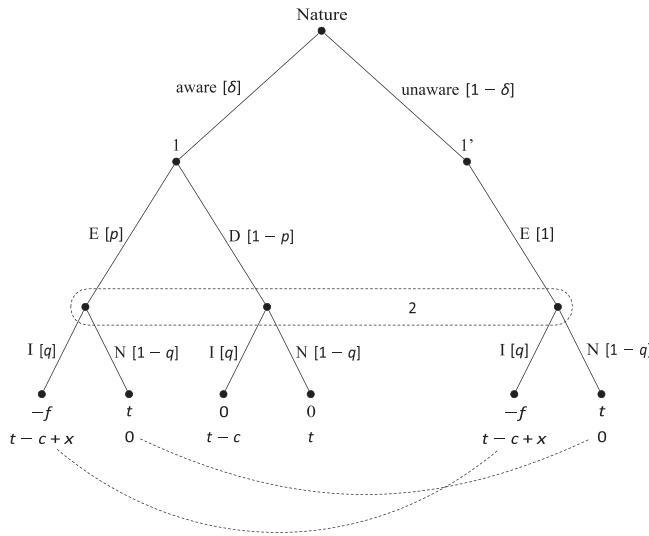


FIGURE 3 The information sets with public returns

$$\mathbb{E}[u_1^{pub}(E)] = q^{pub}[-f - \theta\mu(1 - q^{pub})Rt] + (1 - q^{pub})[t - \theta\mu(1 - q^{pub})Rt]. \quad (16)$$

The equilibrium probability q^{pub} that keeps 1 indifferent between evading and not when 3 has full information about evasion is now the explicit function

$$q^{pub} = \frac{t - \theta t A}{t + f - \theta t A}, \quad (17)$$

where $A = \frac{R[c - (t + x)(1 - \delta)]}{c} < R$ for $\delta < 1$.

Proposition 6. *Under the principle of public access, the equilibrium inspection probability under guilt from blame, q^{pub} , is lower than when tax returns are private. However, the probability is still higher than under simple guilt whenever $\delta < 1$, that is, $q^{GB} > q^{pub} > q^{SG}$. Moreover, under the principle of public access, there will be a pure-strategy equilibrium (D, N) for $\theta > \frac{1}{A} > \frac{1}{R}$, also under guilt from blame.*

When neighbors can freely access tax returns evasion is less rewarding, so less formal inspection is needed in equilibrium. In this sense, the authority's and the social enforcement become substitutes.²⁶ In Figure 2, q^{pub} is represented by the dotted curve. The structure of q^{pub} in (17) is reminiscent more of q^{SG} in (6) than of q^{GB} in (12). Recall that the implicit structure of q^{GB} was due to 3's uncertainty about evasion in case of no inspection. Under the principle of public access, this uncertainty is gone. This is also the reason why the “good” equilibrium,

²⁶On a societal level, social norms in a broad sense and legal enforcement have been found to sometimes be substitutes, sometimes complements (see, e.g., Bowles & Polania-Reyes, 2012; Batrancea et al., 2019). However, when studying social enforcement as such in an experimental setting, Kube and Traxler (2011) found a clear substitutability between legal and social enforcement.

(D, N) is attainable under guilt from blame when tax returns are public. As long as $\delta < 1$ player 3 is, however, uncertain whether an evader is aware or unaware and thereby not blameworthy. Therefore, 2 inspects with a higher probability than under simple guilt.²⁷

6 | CONCLUDING REMARKS

Classify our contribution as “applied behavioral theory.” Drawing on Battigalli and Dufwenberg’s (2007) (B&D) psychological games based approach, we explore how two forms of guilt may shape behavior in an inspection game. The analysis may have multiple interpretation although we chose to highlight and focus on a simple context involving tax evasion and the need for audits. The analysis highlights several novel, and perhaps unexpected, effects.

Intuition may suggest that if tax evaders feel guilty then they are more likely to comply. However, if a tax authority anticipates compliance that reduces the incentives for inspection, which in turn makes evasion more lucrative. The overall equilibrium effect is in most cases null for the taxpayer, while the authority inspects less. Guilt is beneficial, not by making people more honest but by reducing the need for audits.

B&D define two forms of guilt: simple guilt, under which decision makers internalize a concern such that they do not want to let others down, and guilt from blame, under which decision makers instead care about others impressions regarding their intention to let others down. We contrasted the two, and demonstrated (a particular way) that the effect mentioned in the previous paragraph is stronger for simple guilt than for guilt from blame.

The presence of guilt may have idiosyncratic policy implications. We considered a particular case where the presence of guilt from blame makes a crucial difference: the issue of private versus public tax returns. This is a choice that merely affects the information-structure across a game form’s end nodes. That aspect would be irrelevant had we explored a standard rather than a psychological game. However, the issue affects conclusions in our context. In equilibrium, the probability of evasion does not depend on whether tax returns are private or public. However, in the latter case, when neighbors can freely access tax returns, evasion is less rewarding, so less formal inspection is needed in equilibrium.

We have reached our conclusions for a highly stylized and special setting. Yet, it is rich enough to realistically capture several crucial strategic and informational features that are financially and psychologically relevant: We consider an inspection game where each party’s optimal behavior depends on that of another party. We accounted for the presence of unaware decision makers as well as the extent to which a citizens reporting behavior is observable to neighbors, aspects which would be irrelevant with conventional preferences but which turn out to matter crucially with the belief-dependent effect of guilt in the picture.

ACKNOWLEDGMENTS

We have benefited from valuable comments by Geir Asheim, Pierpaolo Battigalli, Stefanos Tsikas, the editor Rabah Amir, from anonymous reviewers, and several participants in seminars and conferences. Financial support from the Swedish Research Council, project nos. 2016-01485-3 and 348-2014-3770 is gratefully acknowledged.

²⁷In the special case where $\delta = 1$, all uncertainty is removed when tax returns are public, and $q^{pub} = q^{SG}$.

**ORCID**

- Martin Dufwenberg <http://orcid.org/0000-0001-8014-2110>
Katarina Nordblom <https://orcid.org/0000-0002-5587-0543>

REFERENCES

- Allingham, M. G., & Sandmo, A. (1972). Income tax evasion: A theoretical analysis. *Journal of Public Economics*, 1, 323–338.
- Alm, J. (2019). What motivates tax compliance? *Journal of Economic Surveys*, 33(2), 353–388.
- Alm, J., & Malézieux, A. (2021). 40 years of tax evasion games: A meta-analysis. *Experimental Economics*, 24(3), 699–750.
- Andreoni, J., Erard, B., & Feinstein, J. (1998). Tax compliance. *Journal of Economic Literature*, XXXVI, 818–860.
- Avenhaus, R., von Stengel, B., & Zamir, S. (2002). Inspection games. In R. Aumann, & S. Hart (Eds.), *Handbook of Game Theory* (Vol. 3).
- Batrancea, L., Nichita, A., Olsen, J., Kogler, C., Kirchler, E., Hoelzl, E., Weiss, A., Torgler, B., Fooken, J., Fuller, J., Schaffner, M., Banuri, S., Hassanein, M., Alarcón-García, G., Aldemir, C., Apostol, O., Bank Weinberg, D., Batrancea, I., Belianin, A., ... Zukauskas, S. (2019). Trust and power as determinants of tax compliance across 44 nations. *Journal of Economic Psychology*, 74, 102191.
- Battigalli, P., & Dufwenberg, M. (2007). Guilt in games. *American Economic Review: Papers & Proceedings*, 97(2), 170–176.
- Battigalli, P., & Dufwenberg, M. (2009). Dynamic psychological games. *Journal of Economic Theory*, 144(1), 1–35.
- Battigalli, P., & Dufwenberg, M. (2021). Belief-dependent motivations and psychological game theory. *Journal of Economic Literature*, forthcoming.
- Baumeister, R., Stillwell, A., & Heatherton, T. (1994). Guilt: An interpersonal approach. *Psychological Bulletin*, 115, 243–267.
- Baumeister, R. F., Stillwell, A. M., & Heatherton, T. F. (1995). Personal narratives about guilt: Role in action control and interpersonal relationships. *Basic and Applied Social Psychology*, 17(1–2), 173–198.
- Becker, G. S. (1968). Crime and punishment: An economic approach. *Journal of Political Economy*, 72, 169–217.
- Bø, E. E., Slemrod, J., & Thoresen, T. O. (2015). Taxes on the internet: Deterrence effects of public disclosure. *American Economic Journal: Economic Policy*, 7(1), 36–62.
- Bosco, L., & Mittone, L. (1997). Tax evasion and moral constraints: Some experimental evidence. *Kyklos*, 50(3), 297–324.
- Bowles, S., & Polania-Reyes, S. (2012). Economic incentives and social preferences: Substitutes or complements? *Journal of Economic Literature*, 50(2), 368–425.
- Bruckner, C. (2016). Shortchanged: The tax compliance challenges of small business operators driving the on-demand platform economy. Available at SSRN 2784243.
- Camerer, C. (2003). *Behavioral game theory*. Russell Sage Foundation.
- Cartwright, E. (2019). A survey of belief-based guilt aversion in trust and dictator games. *Journal of Economic Behavior & Organization*, 167, 430–444.
- Charness, G., & Dufwenberg, M. (2006). Promises and partnership. *Econometrica*, 74(6), 1579–1601.
- Charness, G., & Dufwenberg, M. (2011). Participation. *American Economic Review*, 101(4), 1211–1237.
- Coricelli, G., Joffily, M., Montmarquette, C., & Villeval, M. C. (2010). Cheating, emotions, and rationality: An experiment on tax evasion. *Experimental Economics*, 13(2), 226–247.
- Dufwenberg, M., Gächter, S., & Hennig-Schmidt, H. (2011). The framing of games and the psychology of play. *Games and Economic Behavior*, 73(2), 459–478.
- Dufwenberg, M., & Gneezy, U. (2000). Measuring beliefs in an experimental lost wallet game. *Games and Economic Behavior*, 30(2), 163–182.
- Duleck, U., Fooken, J., Newton, C., Ristl, A., Schaffner, M., & Torgler, B. (2016). Tax compliance and psychic costs: Behavioral experimental evidence using a physiological marker. *Journal of Public Economics*, 134, 9–18.
- Eisenhauer, J. G. (2008). Ethical preferences, risk aversion, and taxpayer behavior. *The Journal of Socio-Economics*, 37(1), 45–63.

- Ellingsen, T., Johannesson, M., Tjøtta, S., & Torsvik, G. (2010). Testing guilt aversion. *Games and Economic Behavior*, 68(1), 95–107.
- Erard, B., & Feinstein, J. (1994). The role of moral sentiments and audit perceptions in tax compliance. *Public Finance/Finances Publiques*, 49(Supplement), 70–89.
- Gahvari, F., & Micheletto, L. (2020). Wage endogeneity, tax evasion, and optimal nonlinear income taxation. *Journal of Public Economic Theory*, 22(3), 501–531.
- Geanakoplos, J., Pearce, D., & Stacchetti, E. (1989). Psychological games and sequential rationality. *Games and Economic Behavior*, 1(1), 60–79.
- Graetz, M. J., Reingenanum, J. F., & Wilde, L. L. (1986). The tax compliance game: Toward an interactive theory of law enforcement. *Journal of Law, Economics, and Organization*, 2(1), 1–32.
- Hashimzade, N., Myles, G., & Tran-Nam, B. (2013). Applications of behavioural economics to tax evasion. *Journal of Economic Surveys*, 27, 941–977.
- Hokamp, S., & Pickhardt, M. (2010). Income tax evasion in a society of heterogeneous agents—Evidence from an agent-based model. *International Economic Journal*, 24, 541–553.
- Hoopes, J. L., Robinson, L., & Slemrod, J. (2018). Public tax-return disclosure. *Journal of Accounting and Economics*, 66(1), 142–162.
- Jung, H. M., Liang, M.-Y. & Yang, C. (2021). How much should we fund the IRS? *Journal of Public Economic Theory*, forthcoming.
- Khalmetski, K., Ockenfels, A., & Werner, P. (2015). Surprising gifts: Theory and laboratory evidence. *Journal of Economic Theory*, 159(Part A), 163–208.
- Kleven, H. J., Knudsen, M., Kreiner, C. T., Pedersen, S., & Saez, E. (2011). Unwilling or unable to cheat? Evidence from a tax audit experiment in Denmark. *Econometrica*, 79, 651–692.
- Kube, S., & Traxler, C. (2011). The interaction of legal and social norm enforcement. *Journal of Public Economic Theory*, 13(5), 639–660.
- Luttmer, E. F. P., & Singhal, M. (2014). Tax morale. *Journal of Economic Perspectives*, 28(4), 149–68.
- Meunier, G., & Schumacher, I. (2020). The importance of considering optimal government policy when social norms matter for the private provision of public goods. *Journal of Public Economic Theory*, 22(3), 630–655.
- Perez-Truglia, R. (2020). The effects of income transparency on well-being: Evidence from a natural experiment. *American Economic Review*, 110(4), 1019–1054.
- Reingenanum, J., & Wilde, L. (1985). Income tax compliance in a principal-agent framework. *Journal of Public Economic*, 26, 1–18.
- Reuben, E., Sapienza, P., & Zingales, L. (2009). Is mistrust self-fulfilling? *Economics Letters*, 104(2), 89–91.
- Slemrod, J. (2019). Tax compliance and enforcement. *Journal of Economic Literature*, 57(4), 904–954.
- Tangney, J. P. (1995). Shame and guilt in interpersonal relationships. In J. P. Tangney, & K. W. Fisher (Eds.), *Self-conscious emotions—The psychology of shame, guilt, embarrassment, and pride*. (1132–1145). The Guilford Press.
- Thayer, J. D. (2020). The gig worker: The growth of the online platform, the need for universal worker classification, and a reduction of the tax compliance burden. *Institute for Global Business Research Conference Proceedings*, 4(1), 118–123.
- Vanberg, C. (2008). Why do people keep their promises? An experimental test of two explanations. *Econometrica*, 76(6), 1467–1480.

How to cite this article: Dufwenberg, M., & Nordblom, K. (2022). Tax evasion with a conscience. *J Public Econ Theory*, 24, 5–29. <https://doi.org/10.1111/jpet.12547>



APPENDIX A: PROOFS

1.1 | Proof of Proposition 1

Proof. (i)

Player 2's expected payoff from I is $t - c + x(\delta p + 1 - \delta)$ and from N is $\delta(1 - p)t$. Assume $\delta < R = \frac{t - c + x}{t + x}$. Then I is preferred to N also if $p = 0$:

$$t - c + x(1 - \delta) > \delta t \Leftrightarrow t - c + x > \delta(t + x).$$

Hence, 2 will always find it more profitable to inspect, and 1's best response is to declare (so $p = 0$).

(ii)

Instead assume $\delta > R$. In equilibrium, if 2 mixes it must be indifferent between inspecting and not, that is, p must satisfy

$$\delta(1 - p)t = t - c + x(\delta p + 1 - \delta) \quad (\text{A1})$$

which implies that 1 must evade with probability

$$p^0 = 1 - \frac{t - c + x}{\delta(t + x)} = 1 - \frac{R}{\delta}. \quad (\text{A2})$$

$\delta > R$ assures that $p^0 \in (0, 1)$.

Similarly, 1's expected payoffs of D and E are, respectively, 0 and $q(-f) + (1 - q)t$. In equilibrium, if 1 mixes he is indifferent between D and E , so

$$0 = q(-f) + (1 - q)t \quad (\text{A3})$$

which implies that in equilibrium, 2 inspects with probability

$$q^0 = \frac{t}{t + f}. \quad (\text{A4})$$

Hence, $q^0 \in (0, 1)$. □

1.2 | Proof of Proposition 2

Proof. (i) $\theta < \frac{1}{R}$ Could we have an equilibrium where 2 chooses I ? If so, 1 responds with D (this maximizes material payoff, and there is no guilt by assumption) and only 1' is caught evading. This is an equilibrium iff 2's payoff from I is no lower than that of N , that is, $t - c + x(1 - \delta) \geq \delta t$, or $\delta \leq R$. Hence, assuming that $\delta > R$, (D, I) cannot be an equilibrium.

Could we have an equilibrium where 2 chooses N ?

Then $\mathbb{E}_E[u_1]^{SG} = (t - \theta R t) > 0$, so 1 responds with E , which in turn would cause 2 to choose I . Hence, when $\theta < \frac{1}{R}$, 2 cannot choose N in equilibrium.

Instead consider the possibility that $q^{SG} \in (0, 1)$. If so, 2 must be indifferent between inspecting and not. Using the same logic as in the proof of Proposition 1 we see that p^{SG} must satisfy

$$p^{SG} = p^0 = 1 - \frac{R}{\delta} \in (0, 1). \quad (\text{A5})$$

and we infer that 1 must be indifferent between D and E . Plugging p^{SG} into (either side of) (A1) one furthermore sees that player 3's expected payoff

$$\mathbb{E}_{s_3, \alpha_3}[m_3|h^0] = Rt > 0. \quad (\text{A6})$$

1 is indifferent between D and E , and since D gives utility 0, so must E . Recall that let-down and thereby simple guilt only occurs when E is chosen. Using notation $[a]^+ = \max\{a, 0\}$ we get

$$\begin{aligned} 0 &= \\ &(1 - q^{SG}) \left(t - \theta [\mathbb{E}_{s_3, \alpha_3}[m_3|h^0] - 0]^+ \right) \\ &\quad + q^{SG} \left(-f - \theta [\mathbb{E}_{s_3, \alpha_3}[m_3|h^0] - (t - c + x)]^+ \right) \\ &= (1 - q^{SG})(t - \theta Rt) - q^{SG}f. \end{aligned} \quad (\text{A7})$$

Simplifying this, we get

$$q^{SG} = \frac{t - \theta Rt}{t + f - \theta Rt}. \quad (\text{A8})$$

(ii)

When $\theta > \frac{1}{R}$, $\mathbb{E}_E[u_1]^{SG} = (1 - q^{SG})(t - \theta Rt) - q^{SG}f < 0$, so the taxpayer always prefers D to E also if not inspected. Given that, 2 will prefer not to inspect since $\delta t > \delta(t - c) + (1 - \delta)(t - c + x)$. Hence, the equilibrium will be in pure strategies (D, N) . \square

1.3 | Proof of Proposition 3

Proof. If (D, N) were an equilibrium then $p^{GB} = q^{GB} = 0$ and (as seen in the bullets of Section 3.2) $\lambda = \mu = 0$. Plugging this into (9) and (10) we get, regardless of θ , $\mathbb{E}_D[u_1]^{GB} = 0 < \mathbb{E}_E[u_1]^{GB} = t$, so player 1 would be able to profitably deviate, contradicting that (D, N) is an equilibrium. \square

1.4 | Proof of Proposition 4

Proof. Could we have an equilibrium where 2 chooses I under guilt from blame? If so, 1 responds with D (this maximizes material payoff, and as explained in Section 3.2 there is no blame or guilt once 3's information set $\{(aware, D, I)\}$ is reached) and only 1' is caught



evading. As before, this is an equilibrium if 2's payoff from I is no lower than that of N , that is, $t - c + x(1 - \delta) \geq \delta t$ which is false under the assumption that $\delta > R$.

We have thus ruled out that $q^{GB} = 1$. Could it be that $q^{GB} = 0$? No: If so, 1's best response would be E (this maximizes material payoff, and although some guilt-from-blame may be involved it is the same regardless of whether D or E is chosen as seen by noting that in each case 3 would apportion blame based on the same information set $\{(aware, D, N), (aware, E, N), (unaware, E, N)\}$), and 2's best response would be I , which is a contradiction. Hence, it must hold that $q^{GB} \in (0, 1)$, so that 2 must be indifferent between inspecting and not. Using the same logic as in the proofs of Propositions 1 and 2 we see that p^{GB} must satisfy

$$p^{GB} = p^0 = 1 - \frac{R}{\delta}. \quad (\text{A9})$$

Since $R > \delta$ we get $p^{GB} \in (0, 1)$ and we infer that 1 must be indifferent between D and E , a property which hence q^{GB} must induce. Equating the expected utilities associated with D and E , respectively, we get

$$\begin{aligned} -(1 - q^{GB})\theta\lambda(1 - q^{GB})Rt &= q^{GB}[-f - \theta\mu(1 - q^{GB})Rt] \\ &\quad + (1 - q^{GB})[t - \theta\lambda(1 - q^{GB})Rt]. \end{aligned} \quad (\text{A10})$$

Note that, although we assume no let-down when D is chosen, 1 may still be blamed if not inspected. This, in turn implies that 2 has to inspect with probability

$$q^{GB} = \frac{t}{t + f + \theta\mu(1 - q^{GB})Rt}. \quad (\text{A11})$$

Without guilt or with simple guilt, the probabilities p and q and p^{SG} and q^{SG} were unrelated. With guilt from blame, however, p^{GB} affects the probability that the taxpayer will be blamed and therefore affects his expected utility. Hence, q^{GB} becomes a function of p^{GB} via μ .²⁸ However, in equilibrium, $p^{GB} = p^0$ and still determined in (1) and we can write the inspection probability:

$$q^{GB} = \frac{t}{t + f + \theta t(1 - q^{GB})A}. \quad (\text{A12})$$

where $A = \frac{R[c - (t + x)(1 - \delta)]}{c} \leq R$. We have already verified that q^{GB} must take a value in $(0, 1)$. As the RHS of (12) is increasing in q^{GB} , we know that (12) has a unique solution $q^{GB} \in (0, 1)$. \square

²⁸This is because 3 cannot tell whether an inspected evader is aware or unaware, that is, blameworthy or not. In case also 3 started the game in node t^1 , the equilibrium inspection probability would have been slightly different:

$$q^{GB} = \frac{t}{t + f + \theta(1 - q^{GB})Rt}$$
 independent of p .

1.5 | Proof of Proposition 5

Proof. Subtract (12) from (2)

$$q^0 - q^{GB} = \frac{\theta t^2 A (1 - q^{GB})}{[t + f + \theta t A (1 - q^{GB})](t + f)} > 0. \quad (\text{A13})$$

Hence, $q^0 > q^{GB} \forall \theta > 0$.

Next, let us show that $q^{GB} > q^{SG} \forall \theta > 0$. We know that when $\theta = 0$, $q^{GB} = q^{SG}$. Moreover, we know from Proposition 4 that $q^{GB} > 0 \forall \theta \geq 0$ and from Proposition 2 ii) that $q^{SG} = 0$ when $\theta > \frac{1}{R}$. Thus, for $\theta > \frac{1}{R}$, $q^{GB} > q^{SG}$. If the two inspection probabilities never cross in the interval $\theta \in (0, \frac{1}{R}]$, we have thus shown that $q^{GB} > q^{SG} \forall \theta > 0$.

Assume that q^{GB} and q^{SG} cross somewhere in the interval $\theta \in (0, \frac{1}{R}]$. Then, the expressions in (6) and (12) are equal, that is,

$$\frac{t - \theta R t}{t + f - \theta R t} = \frac{t}{t + f + \theta A t (1 - q)} \quad (\text{A14})$$

$$\Rightarrow At(1 - q)(1 - \theta R) = Rf. \quad (\text{A15})$$

In a potential crossing, $q = q^{SG} = \frac{t - \theta R t}{t + f - \theta R t}$, which makes (A15) equivalent to

$$t(A - R)(1 - \theta R) = Rf. \quad (\text{A16})$$

Since $A \leq R$, this cannot be true, so q^{SG} and q^{GB} cannot cross. Hence $q^{GB} > q^{SG} \forall \theta > 0$. \square

1.6 | Proof of Proposition 6

Proof. Whenever $\theta \geq (1/A)$, the right-hand side of (17) is negative. Since q^{pub} must be non-negative, this means that player 2 cannot reduce q sufficiently to keep player 1 indifferent between E and D , but player 1 will always prefer D , irrespective of the action of player 2, who will therefore choose N . As $\frac{1}{A} > \frac{1}{R}$ for $\delta < 1$, the threshold value of θ is higher under guilt from blame when tax returns are public than under simple guilt.

We know that when $\theta = 0$, $q^{pub} = q^{GB}$. Moreover, we know from Proposition 4 that $q^{GB} > 0 \forall \theta$ and from the reasoning above that $q^{pub} = 0$ when $\theta > \frac{1}{A}$. Thus, for $\theta > \frac{1}{A}$, $q^{GB} > q^{pub}$. If the two inspection probabilities never cross in the interval $\theta \in (0, \frac{1}{A}]$, we have thus shown that $q^{GB} > q^{pub} \forall \theta > 0$.

Assume that q^{GB} and q^{pub} cross somewhere in the interval $\theta \in (0, \frac{1}{A}]$. Then, the expressions in (17) and (12) are equal, that is,

$$\frac{t - \theta A t}{t + f - \theta A t} = \frac{t}{t + f + \theta A t (1 - q)} \quad (\text{A17})$$



$$\Rightarrow t(1 - q)(1 - \theta A) = f. \quad (\text{A18})$$

In a potential crossing, $q = q^{pub} = \frac{t - \theta At}{t + f - \theta At}$, which makes (A18) equivalent to

$$tf(1 - \theta A) = f(t + f - t\theta A) \Leftrightarrow f = 0. \quad (\text{A19})$$

Since $f > 0$, this cannot be true, so q^{pub} and q^{GB} cannot cross. Hence $q^{GB} > q^{pub} \forall \theta > 0$.

Now, compare the inspection probability, q^{pub} in (17) with q^{SG} determined in (6).

We know that when $\theta = 0$, then $q^{pub} = q^{SG} = q^0$. Moreover $q^{SG} = 0 \forall \theta > \frac{1}{R}$ and $q^{pub} = 0 \forall \theta > \frac{1}{A} > \frac{1}{R}$, that is, q^{SG} reaches zero for a lower value of θ than q^{pub} . It is obvious that both inspection probabilities are negative functions of θ and thus never cross.²⁹ Hence, $q^{pub} > q^{SG} \forall \theta > 0$. \square

²⁹ $\frac{\partial q^{SG}}{\partial \theta} = -\frac{fIR}{[f + t(1 - \theta R)]^2} < 0$ and $\frac{\partial q^{pub}}{\partial \theta} = -\frac{fitA}{[f + t(1 - \theta A)]^2} < 0$.