Maximising Social Welfare in a Principal-agent Setting through Psychological Incentives
A Game Theoretical Approach

ERASMUS UNIVERSITY ROTTERDAM
Erasmus School of Economics
Department of Economics

Supervisor: Prof. dr. O.H. Swank

Name: Ronald van Bezu
Exam number: 331582
Email address: ronaldvanbezu@gmail.com

September 4, 2013

Abstract

This paper considers a game theoretical model similar to auditing models. A large group of agents must comply to a principal's rules, while having an incentive to cheat. The principal has an incentive to monitor behaviour and fine cheating agents. The current paper additionally considers a feeling of shame when an agent is caught, which reduces incentives to cheat and thus the need for monitoring efforts. Psychological incentives, such as feelings of shame, are closely related to norms and values. These could be influenced and, in theory, optimised on the long term by governments, through for example education or religion. This paper proves the existence of some positive, finite level of shame, such that the welfare loss of the necessity to monitor behaviour is minimised, under certain parameter conditions.

Keywords: Shame, Psychological incentives, Monitoring, Principal-agent, Auditing, Welfare, Game theory

Contents

1 Introduction

In this paper, I consider a situation where a large group of agents interacts with a principal. The agents have to comply to the principal's rules. However, the agents have an incentive not to comply, which in turn gives the principal an incentive to monitor their behaviour. Research into this type of situations is important for policy making, as it concerns understanding the behaviour of agents and minimising the costs of monitoring it.

In classical economical models, compliance of an individual agent in such situations would be based on monetary incentives. However, negative feelings from cheating and getting caught doing so can as well play an important role in understanding the agents' behaviour, as they would decrease an agent's incentive to cheat. Therefore, it is also interesting to consider such psychological incentives in an economical model, as it may decrease monitoring costs while at the same time being a cost itself. In this paper, I will model these psychological incentives as a feeling of shame of the agent towards other agents (and perhaps to itself) when getting caught cheating.

An additional component I will include in the model is a probability that agents who comply to the rules of the principal can still be fined. This could happen due to either errors in the system, or mistakes by the agent. While this adds a degree of realism to the model, it also ensures that an infinite amount of shame, which is obviously not realistic, would not be a solution for the necessity of costly monitoring efforts.

Although the model may be suitable for many types of situations, I will consider a setting of a railway line. Passengers are the agents in this model, required to comply by buying a train ticket. They can also 'dodge' by not buying a ticket, at risk of being fined by the railway company and experiencing disutility of feeling ashamed for it. This railway company is the principal, with the goal of profit maximisation. Although this may not be directly true for public organisations, its management is mostly judged by its financial performance or at least cost-effectiveness. This makes profit maximisation a credible assumption for public organisations as well.

I assume the railway line to only consist of two stations, so the trip has no stops in between. This way, no passengers can avoid a train conductor that checks tickets. Another assumption is that the conductor is unprejudiced when writing fines to passengers who do not have a ticket, although in practice a train conductor may act with some (perhaps acceptable) level of discrimination. Also, it may be possible to predict someones wealth by his appearance and adjust the monitoring strategy for that passenger accordingly. However, such practices would be considered unethical in practice, and are not applied in the model. Instead, the two assumptions lead to the important property that all passengers face the same monitoring strategy of the railway company.

Central to this paper is the described feeling of shame that an agent experiences when he is caught cheating. In the game theoretical model, it is treated as an exogeneous factor. However, such feelings of shame can be considered closely connected to the norms and values that exist in a society. Its government is able to exert influence on these norms and values, for example through education, religion or advertising campaigns on television. Thereby, the government is able to adjust, on the long term, the level of shame agents would experience when they are caught cheating. This may not only create a social and livable society, but can also serve an economical purpose. Therefore, I formulate the following research question:

> ***Is there some level of shame that minimises the welfare loss of the necessity to monitor behaviour?***

On the one hand, such shame when getting caught lowers the agents' incentives to cheat, which would require less monitoring and thus increases welfare. On the other hand, the shame itself also creates an additional psychological cost, decreasing welfare. This makes the effect of shame in the model unambiguous and interesting to consider.

One may argue that the additional psychological cost that a caught agent experiences is not in the interest of total welfare of a society, because it is his own fault. However, in this paper I consider the approach of utilitarianism and weigh the utility of each individual equally in calculation of social welfare.

In answering the research question, I will consider two types of games, with different timings of actions. The first is a simultaneous game, where both the agents and the principal move at the same time. The second is a sequential game, where the principal commits to a certain monitoring strategy and the agents make their move after observing this strategy. Additionally, I will consider the differences in outcomes of both games, in terms of welfare and monitoring efforts.

An interesting aspect is that the principal can benefit from fining many agents, especially since complying agents risk being erroneously fined as well.

Therefore, it could be possible that a principal prefers a high amount of cheating agents. However, I generally expect to find that the principal commits to higher monitoring efforts in the sequential game, to deter the agents of cheating.

The structure of the paper is as follows. After the related literature in the next section, I first discuss the setup of the model in section 3. In section 4.1, I analyse the simultaneous game theoretical model and provide proof for the existence of such an optimal level of shame, using several lemma's and propositions. In section 4.2, I discuss how the results for the sequential game differ from those of the simultaneous game. In section 5, I will conclude and provide additional remarks about the model and possible further research into the subject.

2 Related literature

The described model with dodging and checking of train tickets as an example has not yet received much attention in the literature. However, it shares a connection to the field of accounting and is closely related to the literature on auditing models. Similarly, in such models, an agent has to comply to some behaviour required by a principal, who has an incentive to monitor the agent's behaviour because it has an incentive not to comply.

Of auditing models, several applications are discussed in the literature, where the application to tax auditing is considered the largest application. It is quite similar to that of the current paper, as both concern compliance of a large group of individuals to the rules of a single organisation, where the individuals mostly risk a fine based on their earlier compliance decision.

Auditing models require an important assumption about the monitoring strategy of the principal. In the earlier literature, for example Allingham and Sandmo (1972), the monitoring strategy of the tax authority was assumed to be the same for each individual taxpayer, regardless of its characteristics.

Later on, however, auditing models became more based on the assumption that a principal's monitoring strategy towards each agent depends on that agent's characteristics. Reinganum and Wilde (1986) were one of the first to introduce the possibility of such heterogeneous monitoring strategies, in the application of tax auditing. They allowed for the possibility to differentiate in monitoring strategy between each taxpayer by their reported income, which is available information to the tax authority. Taxpayers with a higher income were found to underreport less and audited with a lower probability in their model.

Although the assumption is still applied in more recent literature, see for example Hokamp and Pickhardt (2010), I instead propose a model where the older assumption of a more simple monitoring strategy is appropriate. As described in the introduction, the more simple monitoring strategy results from the assumptions that no passenger can avoid the ticket checking and that the train conductor is not able to discriminate between passengers.

Another important assumption is the possibility to commit to a monitoring strategy. Both possibilities have been considered in the literature. For example, Baron and Besanko (1984), Border and Sobel (1987) and Townsend (1979) assume it is possible to commit to a monitoring strategy, while Reinganum and Wilde (1986) assume the opposite.

Others considered commitment unrealistic, since the principal ex post has less incentive to audit because more agents have complied, see for example Khalil (1997), Bolton and Scharfstein (1990) and Hart (1995). Nonetheless, the current paper concerns an environment where the agent can, in reality, observe deviations from the monitoring activities of the principal (although not modelled) and anticipate in future trips. I therefore consider both possibilities of strategy commitment.

In this respect, Ho and Wang (2013) found that being unable to commit to an auditing strategy in a credit market can result in lower welfare. Likewise, I investigate if the same is true for the model of the current paper.

3 Model setup

In this section, I will translate the proposed setting into a meaningful game theoretical model. The separate characteristics of the model concern the type of game, the players and their payoffs and strategies, and the appropriate solution concept. Additional to the outcomes of this solution concept to both types of players, I will discuss a third, broader type of outcome: total welfare for society.

3.1 Passengers

Let n be the amount of passengers that travel with the railway company. Each player i, $i = 1, ..., n$, gets benefit $B > 0$ from the trip, while facing the option to either *'buy'* a ticket at price $p > 0$, or *'dodge'* the ticket and pay nothing.

However, because the railway company may occasionally check tickets (see section 3.2), both dodging passengers and paying passengers who lose their ticket with probability $\ell \in (0, 1)$ risk getting caught. In that case, each has to pay a fine $f > 0$, beside (again) the regular price p, and suffers additional psychological costs of shame $S \geq 0$. For simplicity, I assume it is not possible for dodging passengers to 'find' a ticket that a buying passenger lost.

The probability of getting caught is denoted by $\pi \in [0, 1]$, and a direct result of the ticket checking activity of the railway company, as will be explained in section 3.2. In this model, I assume a world without discrimination, as discussed in the introduction. Therefore, all passengers face the same level of π.

As the parameters so far are identical for all passengers, the model would not result in realistic continuous reactions to different levels of π. For low values of π under a certain level, all passengers would dodge. For high levels of π above a certain level, everyone would buy a ticket. Therefore, I choose to differentiate between the passengers by a different valuation of money for each passenger i, denoted by $v_i \geq 0$, which follows a certain distribution over the population of passengers.

Let $\psi(v_i)$ be the probability density function and $\Psi(v_i)$ its cumulative distribution function. This function can be chosen in any way subject to $v_i > 0$, to reflect differences in society. For example, a beta distribution can be convenient, as its shape parameter can be used to model inequality in income and thus valuation of money.

Each passenger i now faces the decision between the options 'buy' and 'dodge', with the following expected utilities (EU), given the level of π,

$$\begin{aligned} EU_i(buy|\pi) &= B - v_i p - \pi\ell\Big(v_i(p+f)+S\Big) \\ EU_i(dodge|\pi) &= B - \pi\Big(v_i(p+f)+S\Big) \end{aligned} \tag{1}$$

From these expected utilities follows that there is some threshold level of v_i for which a passenger i is indifferent. Let this threshold level of v_i be denoted by v^*. Passengers with $v_i < v^*$ choose to buy a ticket, whereas passengers with $v_i > v^*$ value money enough to dodge the ticket.

Next, let δ be the proportion of passengers that have the strategy to 'dodge' the ticket. For the purpose of identification of individual strategies, let $v_i > v_j\ \forall i > j$ such that passengers with higher i have a higher tendency to 'dodge'. Then δ defines the individual strategy of each passenger i as 'dodge' if $i > n - \delta n$ and 'buy' if $i \leq n - \delta n$.

With n very large as a finite integer, analysis of the equilibria will be too complex. Therefore, I let $n \to \infty$ to maintain continuity in the (reaction) functions used in all further analysis. Although the amount of players must be finite for proper game theory, there is extensive research on such large games with infinite amount of players, also called atomic models. See for example Aumann (1964), Neyman (2002) and Jara-Moroni (2012) for further reference.

3.2 Railway company

The railway company chooses its ticket checking probability π as strategy. For any ticket checking, the railway company incurs deployment costs of the train personnel. These costs are scaled by the factor $c_\pi > 0$ and quadratic in π, to model that catching more passengers is increasingly more expensive. The resulting checking costs are $c_\pi \pi^2$, where checking all passengers ($\pi = 1$) costs c_π per passenger.

In cases where the railway company does not check tickets, it only receives price p from the paying passengers. If it does checks tickets, it receives the additional price p and fine f from the passengers without a ticket. However, writing a fine incurs administrational costs $c_f > 0$. Also, $c_f < f$ to ensure the railway company has an incentive to indeed write a fine when catching a passenger without a ticket.

The resulting profit function Π, given some level of dodging δ, is

$$\Pi(\pi|\delta) = \pi\left(p + \Big(\delta + (1-\delta)\ell\Big)(f - c_f)\right) + (1-\pi)(1-\delta)p - c_\pi \pi^2. \quad (2)$$

3.3 Timing of actions

In this paper, I will consider two games with different timings of the players' actions. The first is a simultaneous game, which is a static game with imperfect information, where all players choose their strategy simultaneously without knowing what the other players will do. The second is a sequential game, where the railway company commits to a certain level of π. The passengers have perfect information and learn this level of π and react to it. Both types of games have complete information and a Nash Equilibrium (NE) as solution concept.

3.4 Social welfare

With the combination of the levels of δ and π, social welfare can be defined as the benefit B minus the non-retrievable administrational- and checking costs, all expressed as an average per passenger. Let W be the social welfare per passenger, defined as

$$W = B - \pi\Big(\delta + (1-\delta)\ell\Big)(c_f + S) - c_\pi \pi^2. \quad (3)$$

I assume that the administrational- and checking costs do not flow back to society through, for example, salaries, and that lost ticket revenue for the railway company is equally well spent in society by the dodging passengers.

4 Analysis and results

In this section, I will further analyse both the simultaneous game and the sequential game, in terms of reaction functions and behaviour in the solution of the Nash Equilibrium. Furthermore, I will provide proof for the existence of an $S > 0$ such that social welfare is maximised, for the simultaneous game.

I will refer to a numerical example throughout this section, to illustrate characteristics of the game theoretical model. This example contains the parameters $B = 20$, $p = 7.5$, $f = 35$, $c_f = 5$, $c_\pi = 10$ and $\ell = 0.02$. For simplicity, v_i follows a uniform distribution between 0 and $V = 2.5$. I chose the parameters to be realistic for the described railway environment.

4.1 Simultaneous game

First, consider the choice of each passenger i between the expected utilities in (1). It chooses to dodge if $EU_i(buy|\pi) < EU_i(dodge|\pi)$, which simplifies to

$$v_i > \frac{\pi(1-\ell)S}{p - \pi(1-\ell)(p+f)} = v^* \qquad\qquad \text{if } v^* > 0 \tag{4}$$

Passenger i's valuation of money, v_i, must be larger to choose *'dodge'* when either shame or fine increases, or when the price decreases. When these changes occur, the reaction function of the total proportion of passengers dodging, δ, must decrease, because now less passengers value money more than the increased value of v^*.

Note also that, when

$$\pi \geq \frac{p}{(1-\ell)(p+f)}, \tag{5}$$

the denominator in (4) becomes negative. Buying a ticket is then always the worst option, despite the level of S or v_i. The reaction δ will then remain 0 when π increases further.

In general, the condition in (4) leads to the following reaction function:

$$\delta^{BR}(\pi) = 1 - \Pr(v_i \leq v^*).$$

Specifically, the conditions lead to:

$$\delta^{BR}(\pi) = \begin{cases} 0 & \text{if} \quad v^* < 0 \\ 1 - \Psi\left(\frac{\pi(1-\ell)S}{p-\pi(1-\ell)(p+f)}\right) & \text{if} \quad v^* \geq 0 \end{cases} \tag{6}$$

As expected, δ^{BR} is decreasing in S, π and f, and increasing in p and ℓ. For the special case that $S = 0$, the reaction function is:

$$\delta^{BR}(\pi) = \begin{cases} 0 & \text{if } \pi > \frac{p}{(1-\ell)(p+f)} \\ [0,1] & \text{if } \pi = \frac{p}{(1-\ell)(p+f)} \\ 1 & \text{if } \pi < \frac{p}{(1-\ell)(p+f)} \end{cases} \tag{7}$$

Note that the second condition in (7) can lead to a Nash Equilibrium with a mixed strategy, where δ can be anywhere between 0 and 1.

The above results in two lemma's that support the construction of proof for the existence of a level of $S > 0$ such that welfare is maximised.

Lemma 1 $\delta^{BR}(\pi)$ *is monotonically decreasing in* π*.*

Note that it is strictly decreasing in π, unless condition (5) is satisfied.

Lemma 2 $\delta^{BR}(\pi)$ *is monotonically decreasing in* S*.*

Here, $\delta^{BR}(\pi)$ is also strictly decreasing in S, unless condition (5) is satisfied.

Next, the reaction function of the railway company, $\pi^{BR}(\delta)$, can be obtained by the first order condition that maximises the expected profit function in (2), which yields

$$\frac{\partial \Pi(\pi|\delta)}{\partial \pi} = p + \Big(\delta + (1-\delta)\ell\Big)(f - c_f) - (1-\delta)p - 2c_\pi \pi = 0$$

$$\Longrightarrow \pi^{BR}(\delta) = \frac{\delta\Big(p + (1-\ell)(f-c_f)\Big) + \ell(f-c_f)}{2c_\pi} \tag{8}$$

This result supports the following lemma, which I will also prove using economic reasoning instead of mathematics.

Lemma 3 π^{BR} *is strictly increasing in* δ*.*

Proof If δ increases, then the marginal revenue (received price and fine) of checking more tickets increases, because more people are then caught because more people dodge. The marginal costs remain the same at the current level of π. In equilibrium, marginal revenue of increasing π must equal marginal costs. Because the marginal revenue is now higher, the railway company has an incentive to increase π. ■

Note that in the reaction function $\pi^{BR} = \frac{\delta[p+(1-\ell)(f-c_f)]+\ell(f-c_f)}{2c_\pi}$ it can be seen that if $\delta = 0$, $\pi^{BR} = \frac{\ell(f-c_f)}{2c_\pi}$. As this means that π will always remain positive in equilibrium, the following lemma can be proven as well.

Lemma 4 *δ^{BR} becomes infinitely small as S becomes infinity large.*

Proof As $\pi \geq \frac{\ell(f-c_f)}{2c_\pi} > 0$, it can be seen in (4) that $v^* \to \infty$ as $S \to \infty$. In order to have a proper probability distribution where total probability sums up to 1, the probability density function $\psi(v)$ must tend to 0 when $v \to \infty$. A necessary result is that $\lim_{v\to\infty} \Psi(v) = 1$. Therefore it must be true that $\lim_{S\to\infty} \delta^{BR}(\pi) = 1 - \lim_{v^*\to\infty} \Psi(v^*) = 0$. ■

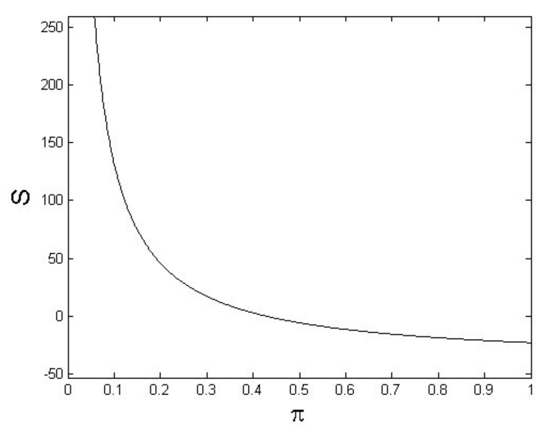
Consider again lemma 2. The following lemma considers the magnitude of the effect of S on $\delta^{BR}(\pi)$, i.e. the derivative $\frac{\partial \delta^{BR}}{\partial S}$.

Lemma 5 *The negative influence of S on δ^{BR} is largest when S is small, but not infinitely large.*

Proof When S is low, its total negative influence on δ^{BR} is low, so δ^{BR} will still be relatively large. Since π^{BR} is only influence by S through δ^{BR}, it follows from lemma 3 that π^{BR} is still relatively large as well.

Consider the utility of dodging in (2). When S is small such that the given level of π is still large, an increase in S has a relatively large influence on this utility because the probability that the passenger incurs these shame costs is high. As a result, the minimum valuation of money needed to dodge, v^* as in (4), increases more. This results in a larger decrease in $\delta^{BR}(\pi)$.

Figure 1: Combinations of given levels of S and π for the passengers, where v^* remains unchanged.



This can be illustrated by considering the trade-off between S and the given level of π, that is necessary for some passenger to remain indifferent between buying or dodging the ticket, i.e. when $EU_i(buy|\pi) = EU_i(dodge|\pi)$. Due to this indifference, a curve that connects such combinations of S and π concerns the indifferent passenger at the threshold v^*, where the level of δ^{BR} would not change. See figure 1 for an example of such a curve.

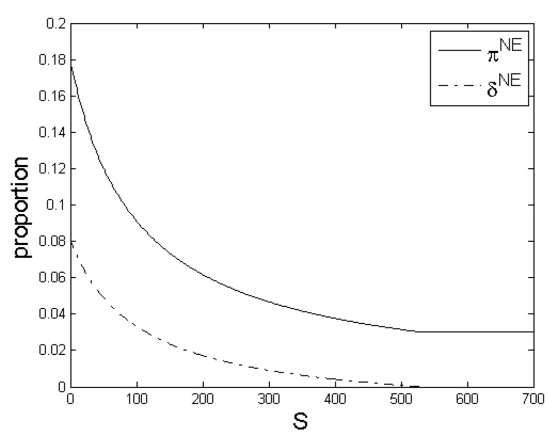
Consider the equation of v^* as in (4), with the purpose of keeping v^* constant. Expressing S in terms of π and taking the derivative yields

$$\frac{\partial S}{\partial \pi} = -\frac{v^* p}{(1-\ell)\pi^2} < 0.$$

This derivative's absolute value describes the rate of substitution between S and π necessary to remain indifferent between dodging and buying a ticket. This rate of substitution is smallest (or the derivative least negative) when S is small and the given π is large. This means that when S is close to 0, an increase of S would require a large decrease of the given π, for v^* and (thus) $\delta^{BR}(\pi)$ to remain the same as a result. The direct decrease in $\delta^{BR}(\pi)$ is then relatively large.

Also, note that the indifference curve in figure 1 has its horizontal asymptote below 0, at $S = -(p + f)$ as derived from (4). Therefore, because $S \in [0, \infty)$, the derivative $\frac{\partial \delta}{\partial S}$ cannot tend to $-\infty$ when S becomes very small. Additionally, because $\delta \in [0, 1]$, the derivative cannot tend to $-\infty$ on its domain and must be finite.

Figure 2: Equilibrium levels of δ and π, against changing S.



Although π^{BR} is only influenced indirectly by S through δ^{BR}, consider also how it adapts towards a Nash Equilibrium due to a change in S. When S increases, v^* increases, causing δ^{BR} to decrease. As a result, π^{BR} decreases indirectly, which further increases the rate of substitution between S and π at their new given levels. This again means that the influence of S on δ^{NE} gets smaller as it increases. This is confirmed by figure 2 of the equilibrium levels of π and δ in the numerical example, against changing S. ■

On the basis of the previous five lemma's, I propose the following proposition:

Proposition 1 *Equilibrium levels of ticket checking π^{NE} and ticket-dodging δ^{NE} are monotonically decreasing in the level of shame S. The decrease is largest when S is small.*

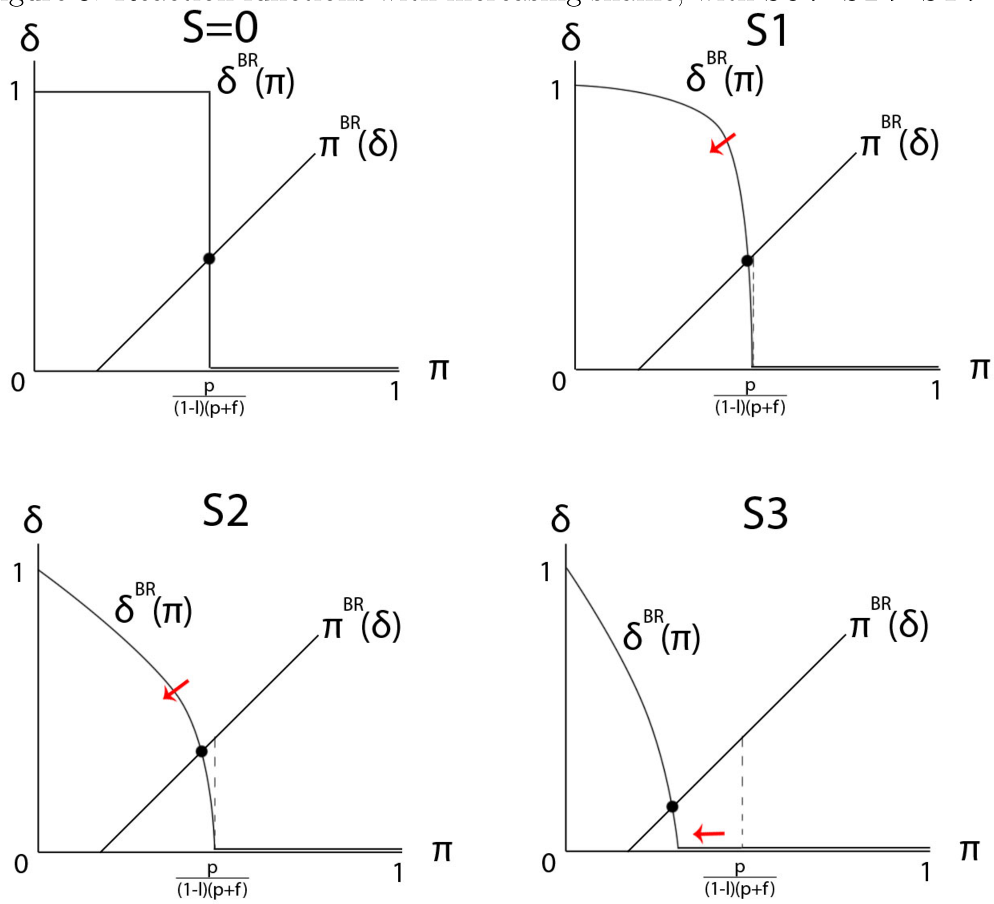
Proof From lemma 2 follows that δ^{BR} decreases. As a direct result, using lemma 3, π^{BR} decreases as well. Because S only has a direct effect on δ^{BR}, lemma 5 is sufficient to prove that the effect of S on δ^{NE} and, therefore, on π^{NE} is largest when S is small.

However, the decrease in π^{BR} causes δ^{BR} to increase again, making the total effect ambiguous. Therefore, assume instead by contradiction that δ^{NE} increases. π^{NE} then increases as well. However, passengers now face higher levels of both S and π. This means that δ^{NE} should be lower than before S increased, which is a contradiction. This proves that δ^{NE} must decrease, and so must π^{NE}. ■

To provide additional insight into the behaviour of δ^{NE} and π^{NE}, consider some example reaction functions with several values of S, in figure 3. The top left diagram shows the reaction functions with $S = 0$. The other three diagrams show increasing levels of S, with $S3 > S2 > S1 > 0$. The Nash Equilibrium in the plots shifts to the bottom left as S increases, with π^{NE} and δ^{NE} both decreasing. When S gets large enough, as in the last diagram with S3, $v^* > V$ occurs before π becomes too high as in condition (5). This results in the reaction function of the passengers intersecting the π-axis before the usual point $\pi = \frac{p}{(1-\ell)(p+f)}$.

At a certain level of S, the reaction function δ^{BR} reaches 0 before the point where the reaction function of the railway company intersects the π-axis. From this point, increasing S does not change the equilibrium levels π^{NE} and δ^{NE}.

Figure 3: Reaction functions with increasing shame, with $S3 > S2 > S1 > 0$



The next two propositions together support the theorem that there is some $S > 0$ such that welfare is maximised. First, proposition 2 states that welfare could be increased at all by introducing S in the model with some finite amount. Next, proposition 3 rules out the possibility that an infinitely high level of S benefits welfare, which would imply there is no optimum at all.

Proposition 2 *Welfare increases when increasing shame S with some finite amount from $S = 0$, under certain conditions.*

Proof Since we know from proposition 1 that both π^{NE} and δ^{NE} will decrease, three welfare effects in the welfare function (3) can be identified. Consider these three effects in more detail, expressed as an average per passenger in the model.[1] It is important to consider that π, δ and $\pi\delta$ are functions of S in equilibrium.

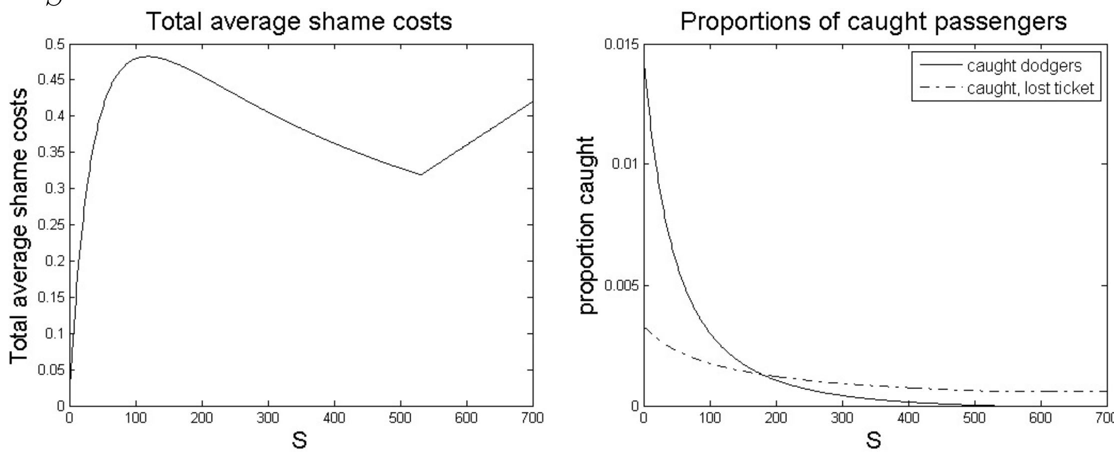
1. **Total average costs of ticket checking decrease**
   The derivative of the total average costs of ticket checking, as in equation (3), to S is
   $$\frac{\partial c_\pi \pi^2}{\partial S} = 2c_\pi \pi \frac{\partial \pi}{\partial S} < 0.$$
   As S increases, π^{NE} decreases and $\frac{\partial \pi}{\partial S} < 0$ becomes less steep as a result of lemma's 2, 3 and 5. The derivative, and therefore the effect of S on these costs, becomes less negative when S is larger.

2. **Total average administration costs of writing fines decrease**
   The derivative of the total average costs of writing fines, as in equation (3), to S is
   $$\frac{\partial \pi \left(\delta + (1-\delta)\ell\right) c_f}{\partial S} = \frac{\partial(\pi\delta)}{\partial S}(1-\ell)c_f + \frac{\partial \pi}{\partial S}\ell c_f < 0,$$
   where both $\frac{\partial(\pi\delta)}{\partial S}$ and $\frac{\partial \pi}{\partial S}$ are negative and become less steep as S increases. As a result, this second positive effect becomes smaller as well, as S increases.

3. **Total average incurred shame costs per caught passenger increase or decrease**
   The derivative of the total average shame costs, as in equation (3), to S is
   $$\frac{\partial \pi \left(\delta + (1-\delta)\ell\right) S}{\partial S} = \left(\pi\delta + \frac{\partial(\pi\delta)}{\partial S} S\right) + \ell\left(\pi(1-\delta) + \left(\frac{\partial \pi}{\partial S} - \frac{\partial(\pi\delta)}{\partial S}\right) S\right)$$

[1] Superscript notations of δ and π are mostly ommitted.

Figure 4: Shame costs and proportions of caught passengers, against level of S



From the first terms in both separate brackets follows that the currently caught dodgers ($\pi\delta$) and caught passengers who lost their tickets ($\ell\pi(1-\delta)$) incur more shame costs. The groups of caught dodgers shrinks by $\frac{\partial(\pi\delta)}{\partial S}$, while the group of caught passengers who lost their tickets could either grow or shrink by $\ell\left(\frac{\partial\pi}{\partial S} - \frac{\partial(\pi\delta)}{\partial S}\right)$. When considering the decrease of the proportion of all caught passengers,

$$\frac{\partial\pi\left(\delta + (1-\delta)\ell\right)}{\partial S} = (1-\ell)\frac{\partial\pi\delta}{\partial S} + \ell\frac{\partial\pi}{\partial S} < 0,$$

it becomes clear that less passengers are caught in total when S increases.

However, it still remains unclear how this decrease of the group of caught passengers weighs up against the shame costs created by the increase of S itself. Therefore, this third welfare effect, as well as the total welfare effect of an increase in S, is ambiguous.

Figure 4 shows that the total average shame costs indeed increase very sharply for low levels of S, although both proportions, of caught dodgers and passengers who lost their tickets, decreased sharply as well. Around $S = 100$, the proportion of caught passengers becomes sufficiently low and decreases sufficiently, such that the increase in S decreases total average shame costs.

It is clear that higher levels of parameters c_π, c_f and ℓ directly increase the magnitude of the three effects, respectively. However, since the three effects also strongly depend on the changes in proportions of δ^{NE}, and π^{NE}, it is important to consider the influence of the parameters on these changes.

Therefore, consider $\frac{\partial\delta}{\partial S}$. Its overall magnitude strongly depends on the shape and width of the distribution of the valuation of money. If one would observe that a relatively large proportion of passengers chooses to buy a ticket instead of dodging when S increases, δ^{BR} would decrease a lot, so π^{BR} would as well. Positive effects 1 and 2 will then be large. Concerning effect 3, shame costs may even decline when a lot less people are caught compared to the additional costs of a higher level of S. If, however, it takes a big increase of S to induce even a single passenger to consider buying a ticket, positive effects 1 and 2 may not compensate for effect 3, which would clearly be negative.

Also, when ℓ is large, an increase of S has relatively little effect on the tendency to cheat less, because instead buying a ticket still entails high risk of getting caught. Beside its influence on $\frac{\partial\delta}{\partial S}$, a high level of ℓ also implies that more passengers are caught in equilibrium, not only because fair passengers are caught more often, but also because more passengers dodge (and are caught with higher probability than when they bought a ticket). This means that an increase in S has a higher negative effect on welfare.

See figures 5 and 6 for an illustration of the levels of welfare for both the simultaneous and the sequential game, for the parameters of the numerical example. Again, these parameters are $B = 20$, $p = 7.5$, $f = 35$, $c_f = 5$, $c_\pi = 10$ and $\ell = 0.02$, and v_i uniformly distributed between 0 and $V = 2.5$. I vary the levels of V, ℓ, c_f and c_π in order to examine the sensitivity of the social welfare curves to these parameters.

The figures show that both games have a potential increase in welfare. However, in the simultaneous game for almost all parameter variations, the negative effect of shame costs dominated for small values of S. This is because π^{NE} is then still high, so more passengers are caught and incur those shame costs S. For higher values of S, welfare could increase above the $S = 0$ level.

As expected, the graphs of the variations of parameters V and ℓ show that higher values may lack a potential increase in welfare. Both parameters appear to have a relatively large influence on the welfare curves, considering their rather mildy chosen level of variation.

Although high levels of c_f and c_π increase the first two positive cost-effects as described above, higher cost-levels do show a potential increase in welfare in figure 6, where lower cost-levels do not. This results from the fact that such higher costs provide a higher potential decrease in monitoring costs, compared to the (incurred) shame costs required to decrease dodging incentives.

Therefore, under the condition that ℓ and the spread in the distribution of v_i are sufficiently low, and c_f and c_π are sufficiently high, a potential increase in welfare due to some increase in S exists. ■

Proposition 3 *In the limit, increasing S results in an infinite decrease of welfare.*

Proof In other words, $\lim_{S\to\infty} W = -\infty$.

When S becomes sufficiently large, it follows from lemma 4 that $\delta^{BR} \to 0$. As a result, it follows from (8) that $\pi^{BR} \to \frac{\ell(f-c_f)}{2c_\pi} > 0$. The only passengers that are caught are those that lost their ticket, of which the proportion tends to $\pi \cdot \ell = \frac{\ell^2(f-c_f)}{2c_\pi}$.

As the equilibrium does not change in the limit, further increasing S does not decrease monitoring costs or administrational costs, but merely increases the shame costs the unfairly caught passengers experience. Therefore, $\lim_{S\to\infty} W = -\infty$. ■

These propositions lead to the following main theorem that concerns the research question.

Theorem 1 *There can be an optimal level of S for maximum welfare in a simultaneous game, under certain conditions.*

Proof The existence of an optimal S follows from propositions 2 and 3, under the conditions discussed in the former proposition.

Consider the lowest $S > 0$, for which welfare has increased as in proposition 2. From this point, welfare cannot increase towards infinity for some *finite* increase of S, because δ and π in (3) are bounded between 0 and 1 and are the only changing variables besides S. With an *infinite* increase of S, welfare must instead tend to minus infinity. Therefore, there must be some finite S in between where the welfare function has a global maximum. ■

Figure 5: Welfare levels for both the simultaneous and the sequential game, for some parameter variations.

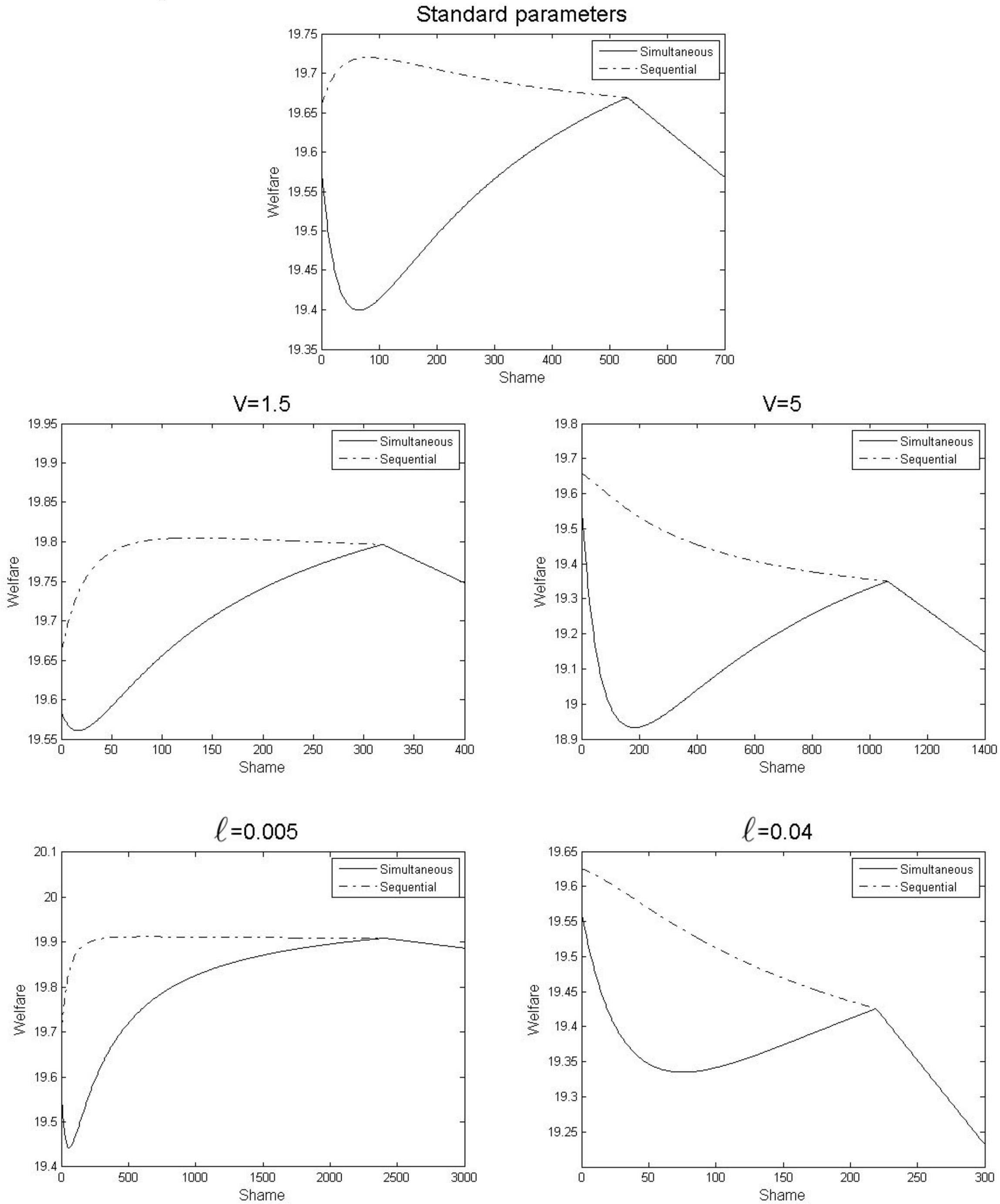
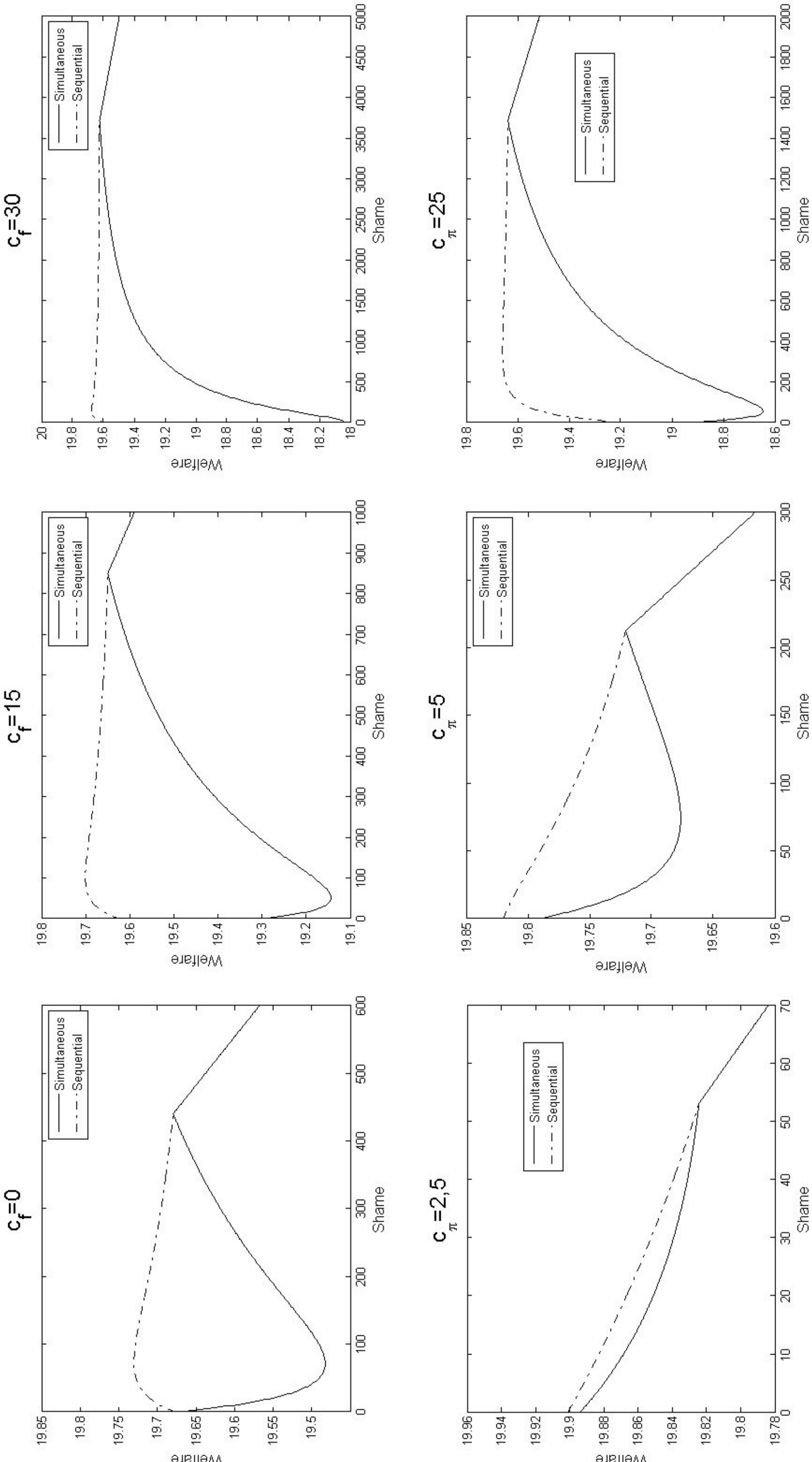


Figure 6: Welfare levels for both the simultaneous and the sequential game, for some parameter variations.



4.2 Sequential game

Now, the railway company commits to a certain level of π beforehand, anticipating on the reaction of the passengers. Therefore, it is interesting to consider the railway company's incentives to induce either higher or lower levels of δ^{NE}, through commitment to a monitoring strategy of π.

Taking the equilibrium of the simultaneous game as reference point, the railway company faces certain benefits (more revenue in fines) and costs (more checking- and administrational costs) from increasing π any further, given the level of δ. Since it was an equilibrium, these benefits and costs are necessarily equal. With commitment in the sequential game, the railway company would anticipate that the passengers dodge less when π increases, and have an incentive to do so if this increases profits.

Therefore, consider the partial derivative of the railway company's profits in (2) to δ:

$$\begin{aligned}\frac{\partial EU_{RC}(\pi|\delta)}{\partial\delta} &= \pi(1-\ell)(f-c_f)-(1-\pi)p \\ &= \pi\Big((1-\ell)(f-c_f)+p\Big)-p\end{aligned}$$

Most naturally, one would expect that a company is always aversive to cheating. The derivative would be negative, such that cheating hurts profits. Rewriting $\frac{\partial EU_{RC}(\pi|\delta)}{\partial\delta}<0$ shows that this is the case when the condition

$$\pi^{NE} < \frac{p}{(1-\ell)(f-c_f)+p} \tag{9}$$

is satisfied. Observe that when writing a fine provides a large benefit of $f-c_f$ for the railway company, or when price p decreases such that a dodging passenger causes a smaller loss of revenue, the threshold in (9) decreases. As expected, this increases the tendency to instead prefer higher levels of dodging, compared to the equilibrium of the simultaneous game. The same happens when ℓ decreases, because fair passengers become less interesting for potential revenue of writing unfair fines, making dodging passengers relatively more profitable.

Note that all equilibria must satisfy $\pi^{NE}<\frac{p}{(1-\ell)(p+f)}$ (see also condition (5)). So, first, it can be seen that the railway company *always* commits to a higher level of π, and thus favors lower levels of dodging, when all these possible equilibria satisfy condition (9), i.e. when

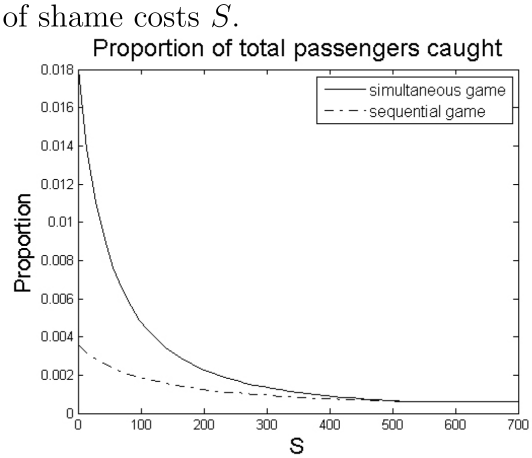
$$\frac{p}{(1-\ell)(f-c_f)+p} > \frac{p}{(1-\ell)(p+f)} \qquad \Longleftrightarrow \qquad (1-\ell)c_f - \ell p > 0 \quad (10)$$

is satisfied.

Consider also the case when condition (10) is not satisfied[2]. When π^{NE} is now sufficiently high, above the threshold in (9), the railway company instead favors *higher* levels of dodging and will commit to a *lower* level of π. This is the case for low levels of S, which follows from lemma 1. If, however, S increases and thus π^{NE} decreases, such that it is lower than that threshold in (9), the additional benefit of a dodging passenger decreases because it is caught with less probability. Profit is then hurt by dodging passengers, which provides the railway company an incentive to commit to a higher level of π than in the simultaneous game.

In the numerical example of the model, condition (10) is satisfied, such that the railway company always commits to a higher level of π and the levels of dodging are reduced as a result. Also, figure 7 confirms that less passengers are caught in total in the numerical example. Therefore, the railway company has lower administrational costs and the passengers incur less shame costs on average, at the expense of higher ticket checking costs.

Figure 7: Proportions of total caught passengers for both types of games, against the level of shame costs S.



[2]Note that the possibility of $\ell > 0$ is crucial for this to occur.

To see how this influences the difference in levels of welfare among both games, consider again figures 5 and 6. First, the levels of welfare are higher in the sequential game than in the simultaneous game. The opposite would be true when instead applying models with parameters that do not satisfy condition (10), such that the railway company commits to a *lower* level of π.

Second, beside these differences in absolute levels of welfare among both types of games, the commitment also changes the influence of the level of S on the change in welfare. Figures 5 and 6 show that welfare in the sequential game mostly increases directly when S is introduced in the model. This results from a decrease in both administrational- and ticket checking costs. The passenger's tendency to dodge less can be fully utilised by the railway company to prevent costly monitoring activity. In the numerical example, I even observed that dodging was mostly reduced to $\delta^{NE} = 0$ in the sequential game. At the same time, beside the decrease in monitoring costs, total average shame costs do not increase as much as in the simultaneous game, because less passengers are caught and incur shame costs S.

5 Conclusion and discussion

In this paper, I discussed a model where multiple agents must comply to the rules of a principal, who has to costly monitor the agents because they have an incentive not to comply. I showed that introducing the experience of shame for caught agents has the potential to increase social welfare, with the existence of a global maximum for a finite level of shame. Although this creates a psychological burden on the agents, such psychological costs of shame decrease the agents' incentives to cheat. Therefore, monitoring efforts and costs can be reduced as well, which could result in a higher social welfare.

However, certain conditions may affect the existence of an optimal level of shame. First, the heterogeneity among the group of agents in tendency to cheat, in this paper modeled by valuation of money, must not be too large. A more homogeneous group requires a smaller level of shame to change behaviour of the same amount of agents. Second, if the administrational- and monitoring costs are too small, the potential decrease in monitoring costs does not outweigh the introduced shame costs that caught agents experience. Third, when relatively many agents are caught unfairly when monitored, for example due to losing their ticket, a feeling of shame has less impact on cheating tendency, since compliance still entails a high risk of getting caught.

I considered the model in both a simultaneous game and a sequential game where the principal commits to a monitoring strategy. In the sequential game, the principal generally commits to a higher level of monitoring, resulting in higher welfare than in the simultaneous game. The opposite may occur in case of a high probability that an agents is caught unfairly when monitored. Also, because the principal can fully utilise the decrease in cheating tendency of the agents to prevent cheating, welfare mostly increases directly when shame is introduced into the model, whereas welfare at first decreases sharply in the simultaneous game.

Through the existence of an optimal level of shame, this paper demonstrates the economical importance of norms and values in a society, as these influence psychological feelings such as shame on the long term. It therefore shows the important role the government plays in the described situation, by being able to create and maintain norms and values to its citizens, by for example religion or education. These do not only make a social and liveable society, but can also serve an economical purpose when chosen optimally.

However, this paper considers social welfare according to the utilitarian approach, weighing the utility of each individual equally. Therefore, it does not take into account any ethical considerations, as one might for example believe that a cheating person deserves to be ashamed. Nonetheless, there are many situations where any average person may be tempted to cheat when it is very easy to do so or requires additional action to comply, such as mildly breaking the speed limit, 'forgetting' to report all of your taxes or not handing in the hotel's bathrobe after your stay.

A limitation of this paper is that both the fine and the price are treated as exogeneous factors in the model. In practice, however, a principal may have an incentive to adapt these parameters as well. Further research is necessary to see if considering a model with these parameters as endogeneous still produces similar results. However, one would have to overcome difficulties with integrating a demand function for the price parameter and a participation constraint due to a tendency towards very high fines.

Since the decisions of the agents contain an element of chance, an alternative to the valuation of money would be to differentiate between passengers on their risk averseness. First, this can be accomplished by a utility function with an overall exponent parameter, which is differentiated upon between agents, similar to v_i. Second, it is possible to apply prospect theory instead of expected utility theory, by multiplying utilities with a function of the probability of each event. This could be used to let agents assign greater weight to large losses (being fined), again with a different function for each agent.

When modelling risk in such ways, it is also possible to differentiate on wealth, in a simpler way than the multiplication factor v_i, by adding an amount m_i in the utilities (before the exponent) as money endowment or income. In the current model without risk exponents, such a parameter would be cancelled out when comparing utilities as in (4).

Another possibility for further research is to take a step back from the (infinitely) large amount of agents, as it may lack conformity to proper game theory. Single agents in the model may act to represent differentiated groups in society, for example a wealthy agent and a poor agent, or a risk-loving agent and a risk-averse agent, or combinations of these four together. This would prevent the problem that a single agent in the model creates outcomes with all agents either complying or cheating around a monitoring threshold, which would be unrealistic. Instead, as multiple (fixed) proportions of cheating agents are possible in the equilibria of the game, the model remains more realistic, while it should still be possible to reproduce the results of the current paper.

6 References

Allingham, M.G. and Sandmo, A. (1972), 'Income tax evasion: a theoretical analysis', *Journal of Public Economics* **1**(3-4), 323-338.

Aumann, R.J. (1964). 'Markets with a Continuum of Traders', *Econometrica*, **32**(1/2), 39-50

Baron, D.P. and D. Besanko (1984). 'Regulation, Asymmetric Information, and Auditing', *RAND Journal of Economics* **15**, 447-470.

Bolton, P. and Scharfstein, D.S. (1990). 'A Theory of Predation Based on Agency Problems in Financial Contracting', *American Economic Review* **80**, 93-106.

Border, K.C. and Sobel, J. (1987). 'Samurai accountant: A theory of auditing and plunder', *The Review of Economic Studies* **54**(4), 525-540.

Hart, O. (1995). 'Firms, Contracts, and Financial Structure', *Oxford: Oxford University Press.*

Ho, W.-H. and Y. Wang (2013). 'Asymmetric Information, Auditing Commitment, and Economic Growth', *Canadian Journal of Economics* **46**(2), 611-633.

Hokamp, S. and M. Pickhardt (2010). 'Income Tax Evasion in a Society of Heterogeneous Agents - Evidence from an Agent-based Model', *International Economic Journal* **24**(4), 541-553.

Jara-Moroni, P. (2012). 'Rationalizability in games with a continuum of players',*Games and Economic Behaviour* **75**, 668-684.

Khalil, F. (1997). 'Auditing without Commitment', *The RAND Journal of Economics* **28**(4), 629-640.

Neyman, A. (2002). 'Values of games with infinitely many players', in R. Aumann and S. Hart, eds., *Handbook of Game Theory.* North-Holland.

Reinganum, J.F. and L.L. Wilde (1986). 'Equilibrium verification and reporting policies in a model of tax compliance', *International Economic Review* **27**(3), 739-760.

Townsend, R.M. (1979). Optimal contracts and competitive markets with costly state verification, Journal of Economic Theory **21**, 265-293.