University of Innsbruck

# Screening experts' distributional preferences

**Dominik Erharter**

**Working Papers in Economics and Statistics**

2013-27

Research Platform eeecon
Empirical and Experimental Economics
University of Innsbruck
http://eeecon.uibk.ac.at/

**University of Innsbruck**
**Working Papers in Economics and Statistics**

The series is jointly edited and published by

- Department of Economics
- Department of Public Finance
- Department of Statistics

Contact Address:
University of Innsbruck
Department of Public Finance
Universitaetsstrasse 15
A-6020 Innsbruck
Austria
Tel: + 43 512 507 7171
Fax: + 43 512 507 2970
E-mail: eeecon@uibk.ac.at

The most recent version of all working papers can be downloaded at
http://eeecon.uibk.ac.at/wopec/

For a list of recent papers see the backpages of this paper.

# Screening Experts' Distributional Preferences*

Dominik Erharter†

October 2013

**Abstract**

We study optimal direct mechanisms for a credence goods expert who can be altruistic or spiteful. The expert has private information about her distributional preferences and possibly also about her customer's needs. We introduce a method that allows the customer to offer separate contracts to different preference types and outline when separation is optimal. Furthermore, we demonstrate that the optimality of separating mechanisms is sensitive to minor changes of the customer's utility function. Additionally, we illustrate how our results extend to more than two preference types and discuss possible policy implications.

*Keywords:* other-regarding preferences; inequality aversion; credence good; principal-agent model; adverse selection; moral hazard.
*JEL Classification:* D63; D64; L13; L15; C72.

*This paper is a revised version of Chapter 3 of my dissertation at the University of Innsbruck. Many thanks to Rudolf Kerschbamer, Markus Walzl, Joel Sobel, Georg Kirchsteiger, Bram De Rock, Niall Flynn, Heiko Karle and Christian Waibel for helpful comments and discussions. Funding by the Austrian Academy of Sciences and the Vice Rector for Research at the University of Innsbruck and the hospitality of ECARES/ULB and the Department of Economics at the University of California at San Diego (UCSD) are gratefully acknowledged.

†University of Innsbruck, Department of Economics, SOWI Building, Universitätsstrasse 15, 6020 Innsbruck, Austria. Email: dominik.erharter@uibk.ac.at. Phone: +43 (0)512 507–7374. Fax: + 43 (0)512 507–2980.

# 1 Introduction

This paper studies optimal mechanisms for a credence goods expert who has private information about her distributional preferences. Credence goods come in different qualities and have the property that customers do not know which quality suits them best.[1] Experts can diagnose customers' needs (or 'problems') and can provide a treatment that satisfies these needs – or they can defraud customers. Economic theory suggests that customers can induce experts to be honest by paying a sufficiently large information rent. Yet, Dulleck et al. (2009) find that experts do not respond to monetary incentives as predicted and show that this can best be explained by the presence of distributional preferences, meaning that experts not only care for their own monetary payoff, but also for the monetary payoff of their customers. Motivated by this observation, Erharter (2012) derives optimal mechanisms for an expert with distributional preferences that are common knowledge. Compared to a selfish expert, the information rent paid to an expert is higher if the expert's utility decreases in the customer's monetary payoff (spiteful preferences) and lower if her utility increases in the customer's monetary payoff (altruistic preferences).

The present paper extends the author's work to unknown distributional preferences. We assume that the expert's utility can increase (altruistic preferences) or decrease (spiteful preferences) in the customer's monetary payoff. We consider the case where the expert's actions are constrained by liability and verifiability and investigate the one-dimensional screening problem where the customer knows his needs but not the expert's preference type and the two-dimensional screening problem where the customer neither knows his needs (the expert's 'problem type') nor the expert's preferences. In both cases, the customer has to decide whether to choose information rents that are optimal for the spiteful type (pooling mechanism) – and to forgo some payoff from the altruistic type – or to choose information rents which are optimal for the altruistic type (shut-down mechanism) – and to be untreated by the spiteful type. Alternatively, the customer can try to offer different information rents to both expert types (separating mechanism). In contrast to standard principal-agent problems where differences in effort costs or productivities can be exploited, separating preference types simply means to assign one type a lower share of the surplus than the other type. However, as long as the expert cares more for her own monetary payoff than the monetary payoff of the customer, even an altruistic expert should prefer a higher share of the pie to a lower share.[2] Nevertheless, the customer can induce the altruistic expert to accept a lower share by burning a part of the surplus generated by a spiteful expert.

We obtain similar results in the one-dimensional case and the two-dimensional case. If spiteful types are common, it is optimal to pool preference types and to separate problem types. If spiteful types are uncommon, it is optimal to separate

[1]Credence goods have been defined by Darby and Karni (1973) as extension to Nelson (1970)'s categorization of ordinary goods, search goods and experience goods. Dulleck and Kerschbamer (2006) provide an extensive survey of the literature.

[2]Experimental studies measuring distributional preferences find that that there are few subjects who put a higher (positive) weight on others' monetary payoff than on their own monetary payoff if they are behind in monetary terms (Andreoni and Miller, 2002; Fisman et al., 2007; Kerschbamer, 2010).

preference types and problem types. However, even though problem types are separated, all spiteful types end up with the same monetary payoff. Furthermore, we find that if the domain of transfers is unrestricted and separation is optimal, it is optimal for the customer to burn the entire surplus generated by a spiteful expert and to assign an altruistic expert her (negative) first-best payoff. Thus, the customer's expected payoff is equivalent to the expected payoff generated by a shut-down mechanism. Conversely, if transfers have to exceed expert's treatment costs, a mechanism that separates preference types can outperform a shut-down mechanism. A further inspection of this result reveals that the optimality of a shut-down mechanism is highly sensitive to changes in the customer's utility function. For example, a separating mechanism is also optimal if the customer is risk-averse or has a negative outside utility. In addition, we demonstrate how our results extend to more than two preference types and discuss the relevance of our findings for actual credence goods markets.

This paper is related to a growing literature on principal-agent problems and distributional preferences[3] and complements this literature in several ways: (i) while most papers only consider one preference type, we allow for two (or more) preference types; (iii) while almost all papers assume that distributional preferences are known, we assume that they are unknown, (iii) furthermore, this paper is the first that presents a setting where agents can be separated with regard to their distributional preferences; (iv) finally this papers extends the existing literature to credence goods markets.

The paper proceeds as follows. Section 2 discusses related literature; Section 3 introduces a parsimonious credence goods model proposed by Erharter (2012) and derives the optimal mechanism for the first-best case where the customer knows his problem and the expert's distributional preferences. Section 4 studies optimal mechanisms for the (intermediate) second-best case where the customer knows his problem but not the expert's distributional preferences. Section 5 derives optimal mechanisms for the two-dimensional screening problem arising when the customer knows neither his problem nor the expert's preferences. Section 6 discusses results and possible generalizations. Section 7 concludes.

# 2 Background and related literature

This paper builds on a model proposed by Erharter (2012) that preserves the informational structure arising in credence goods markets – a unique combination of adverse selection and moral hazard – but abstracts from the structural assumptions of existing models. In this model a self-regarding customer (he) incurs either a low or a high need. Low needs can be satisfied by low (cheap) or high (expensive) treatments, high needs can only be satisfied by high treatments. The customer consults a single expert (she) with distributional preferences, who has the ability to diagnose the customer's need at zero cost and can provide an appropriate treatment. The customer has the power to design a direct revelation mechanism (Myerson, 1982)

[3] Most recent contributions include Fehr et al. (2007), Englmaier and Wambach (2010), Neilson and Stowe (2010), von Siemens (2011) and Kosfeld and von Siemens (2011).

that specifies the rules of interaction between himself and the expert. In particular, this mechanism specifies the kind of treatment the expert has to carry out given her diagnosis and the transfer the expert receives for each kind of treatment. Erharter (2012) studies optimal mechanisms in the absence of institutions, under liability and/or verifiability[4] and discovers that distributional preferences have a large impact on the expert's behavior in the absence of institutions and in the verifiability case. In the liability case, the expert's action space is severely narrowed and her preferences over the remaining actions are more aligned regardless of the specificities of the model. Therefore, distributional preferences – along with other factors such as capacity constraints or reputation – have less impact on expert's behavior. Nevertheless, the customer can improve her expected monetary payoff in all of the above settings by taking the expert's distributional preferences into account.

The present paper contributes to the literature on principal-agent problems and distributional preferences. Itoh (2004), Demougin et al. (2006), Desiraju and Sappington (2007), Dur and Glazer (2008), Bartling and von Siemens (2010b), Englmaier and Wambach (2010) and Neilson and Stowe (2010) review optimal employment contracts when workers are inequality-averse. While Grund and Sliwka (2005) investigate inequality aversion in tournaments, Bartling and von Siemens (2010a) discuss the role of inequality aversion in partnerships. Cabrales and Calvò-Armengol (2008) and Kosfeld and von Siemens (2009, 2011) analyze labor market segregation when workers are inequality-averse. Although it is well established that distributional preferences are heterogeneous among subjects (Fehr and Schmidt, 1999), most authors assume that distributional preferences are homogeneous and observable. Fehr et al. (2007) are a notable exception. They consider a principal-agent model with moral hazard where both principal and agent can be either selfish or 'fair' (inequality-averse). They allow principals to choose between more and less explicit contracts and find that the presence of fair subjects leads to less explicit contracts.

To the best of our knowledge, von Siemens (2011) is the only author who studies a principal-agent setting where agents have heterogeneous distributional preferences. There is a risk-neutral, profit maximizing employer. The principal faces a continuum of workers who can be of high or low-productivity and are either inequality-averse or selfish. The inequality-averse agents have a piece-wise linear utility function as proposed by Fehr and Schmidt (1999). Agents choose their reference group endogenously. If they accept a contract by the principal, their reference group consists of other co-workers. If they do not accept a contract, they get an outside rent of zero and their reference group consists of other unemployed workers. Workers with the same productivity have identical preferences over contracts. Hence, it is impossible to screen for preferences. However, selfish and inequality-averse agents may have different participation constraints. High-productivity types get information rents that prevent them from mimicking low-productivity types. Inequality-averse agents suffer more from disadvantageous inequality than vice versa. Therefore, inequality-averse

[4]As shown by Dulleck and Kerschbamer (2006), liability and verifiability are sufficient to contain fraud on credence goods markets if experts are selfish. Moreover, the authors show that the large variations in fraud levels in various credence goods models can be traced back to the presence or absence of liability and/or verifiability.

agents with high productivity strictly prefer a high-productivity contract to a low productivity contract because they have a higher payoff and hence have a lower disutility from inequality.[5] If low-productivity types get no rents, an inequality-averse low-productivity type feels envious towards high-productivity types and therefore might prefer to stay unemployed. In order to overcome this incentive problem, the principal can either exclude inequality-averse low-productivity workers or he can give a positive rent to all low-productivity workers and increase the rents of high-productivity workers accordingly. In contrast to von Siemens (2011), we assume that the agent's (expert's) reference group is the principal (customer). Given the experimental evidence provided by Dulleck et al. (2009), this is a plausible assumption.

# 3 A model

## 3.1 Basic setup

The following credence goods market model is an extension to the model proposed by Erharter (2012). There is a customer (he) who has a low need (or 'problem') $\ell$ with probability $p_\ell$ and a high need $h$ with probability $p_h = 1 - p_\ell$. The customer consults an expert (she) who can diagnose the customer's need at zero cost. The expert has private information about her own distributional preferences. There is a spiteful preference type (type 1) that occurs with probability $q_1$ and an altruistic preference type (type 2) that occurs with probability $q_2 = 1 - q_1$. Together, expert's private information can be represented by a two-dimensional problem-preference type $\theta_{ik} \in \{\theta_{1\ell}, \theta_{1h}, \theta_{2\ell}, \theta_{2h}\}$. Note that as the customer's needs are independent of the expert's preferences, both type dimensions are drawn independently from each other.

The expert can provide a low treatment $x_\ell$ that can satisfy low needs or a high treatment $x_h$ that can satisfy both needs. The cost of low and high treatment be are $x_\ell$ and $x_h$ respectively, where $0 \leq x_\ell \leq x_h$. If the customer's need is satisfied, he receives a valuation of $v$ that is larger than $x_h$. If the customer's need is not satisfied, he receives a valuation of zero. This setup permits three types of fraud. The expert can 'undertreat' the customer by providing a low quality treatment when high quality is needed, and can 'overtreat' the customer by providing a high quality treatment when low quality would be sufficient to the customer's needs. Furthermore, the expert can 'overcharge' the customer by charging for a treatment she did not provide. Dulleck and Kerschbamer (2006) consider a self-regarding expert and suggest that two market institutions – verifiability of experts' treatment decisions and liability against undertreatment – can overcome fraud each on their own or in combination if they are accompanied with suitable monetary incentives. However, in Dulleck et al. (2009, 2011), the authors later find experimental evidence suggesting that the same is not true if the expert has distributional preferences. In the following we will assume that both liability and verifiability hold. As shown in Erharter (2012), this is the only constellation where undertreatment, overtreatment

[5] According to Fehr and Schmidt (1999), inequality averse agents suffer more from inequality if they are behind in monetary terms than if they are ahead.

and overcharging can be avoided simultaneously – even if the expert is spiteful. Liability entails that the expert cannot undertreat the customer – if the customer has the high problem, the expert can decide either to provide a high treatment or to reject service. On the other hand, if the customer has the low problem, the expert can either treat the customer appropriately by choosing $x_\ell$ or overtreat the customer by choosing $x_h$. Thus, liability implies that the customer always receives valuation $v$ when the expert provides a treatment. Verifiability means that the customer can observe the expert's actions and can condition transfers (and the decision to burn surplus) on this observation. This implies that the expert cannot overcharge the customer by providing a low treatment and claiming the transfer specified for a high treatment.

In Erharter (2012), we assume that the expert receives a transfer $t \in \mathbb{R}$ for her services. Note that this domain permits transfers to be below the expert's treatment costs. As the utility of an altruistic expert increases in the customer's monetary payoff, it is possible that the expert is willing to treat the customer at a monetary loss. Given that businesses that (constantly) incur losses should be forced to drop out of the market sooner or later, this is not an entirely convincing assumption. Therefore, we additionally consider the case of restricted transfers, where $t_{ik} \geq x_{ik}$ for $x_{ik} \in \{x_\ell, x_h\}$. As it will turn out, assumptions about the domain of transfers have important implications on the set of optimal mechanisms.

The customer can commit himself to 'burn' a part of the surplus generated by the expert's service. Given the expert provides treatment $x_{ik}$, the customer can destroy any amount $w_{ik} \in [0, v - x_{ik}]$ for $x_{ik} \in \{x_\ell, x_h\}$. Burning surplus allows the customer to generate trade-offs between surplus maximization and the expert's own monetary payoff that can be used to separate preference types. If the expert refuses treatment, both players get an outside utility of zero. If the expert provides a treatment, her monetary payoff is given by $m_{ik} = t_{ik} - x_{ik}$ and customer's monetary payoff is given by $o_{ik} = v - t_{ik} - w_{ik}$.

The customer designs a mechanism that specifies the set of messages the expert can send about her (two-dimensional) type. This commits the customer to recommend a certain treatment for every possible message, to pay the expert a certain transfer, and to burn a certain amount of surplus for every treatment the expert can provide.[6] As discussed in Erharter (2012), it is sufficient to consider direct revelation mechanisms where the expert's message space is equal to her type space and where the expert has an incentive to reveal her (two-dimensional) type truthfully and to follow the customer's treatment recommendation obediently. This entails that the customer's design has to be incentive-compatible. A direct revelation mechanism can be represented by a vector $\mu = (x_{ik}, t_{ik}, w_{ik})_{i=\ell,h}^{k=1,2}$ that specifies a treatment $x_{ik}$, a transfer $t_{ik}$ and an amount of surplus to be burned $w_{ik}$ for every type $\theta_{ik}$.[7] As in Erharter (2012), we apply the standard tie-breaking rule that the expert behaves honestly and obediently if indifferent between actions.

The timing of events is as follows: (i) nature draws the expert's (two-dimensional) type, (ii) the customer designs a direct revelation mechanism $\mu$, (iii) the expert

[6] This formulation takes into account that the customer can verify the expert's actions. See Erharter (2012) for a more general formulation.

[7] That is, $\mu = (x_{ik}, t_{ik}, w_{ik})_{i=\ell,h}^{k=1,2} = (\mu(\theta_{\ell 1}), \mu(\theta_{\ell 2}), \mu(\theta_{h1}), \mu(\theta_{h2}))$.

observes the mechanism, learns her type and decides whether to treat the customer or not. If she decides to treat the customer, she sends a message, otherwise the interaction ends. (iv) If the customer receives a message, he recommends the pre-specified treatment. Upon receiving the customer's recommendation, the expert provides this treatment, (v) the expert receives the pre-specified transfer, the pre-specified part of the surplus is burned and payoffs are realized.

The customer maximizes his own (expected) payoff. Hence, his utility from treatment $x_{ik}$, transfer $t_{ik}$ and burning $w_{ik}$ is

$$V_{ik} = v - t_{ik} - w_{ik}. \tag{1}$$

The expert has distributional preferences that can be represented by a utility function of the form

$$U_{ik} = t_{ik} - x_{ik} + b_k(v - t_{ik} - w_{ik}), \tag{2}$$

where $b_k \in \{b_1, b_2\}$ with $-1 < b_1 < 0 < b_2 < 1$ is the expert's valuation of the customer's monetary payoff. As $b_1 < 1$, type 1 is (linearly) spiteful (Levine, 1998). As $b_2 > 2$, type 2 is (linearly) altruistic (López-Pérez, 2008).

In Erharter (2012) we consider the more general class of piece-wise linear distributional preferences. These preferences have the property that the expert can have a different attitude towards the customer's payoff if she is ahead ($t_{ik} - x_{ik} > v - t_{ik} - w_{ik}$) than if she is behind ($t_{ik} - x_{ik} < v - t_{ik} - w_{ik}$). This class includes the most prominent models of distributional preferences used in experimental and behavioral economics, notably the inequality aversion model by Fehr and Schmidt (1999) and the quasi-maximin model proposed by Charness and Rabin (2002). However, given liability and verifiability, it can never be optimal for the customer to choose a transfer such that the expert is ahead in monetary terms (Erharter, 2012). Hence, our assumption is without loss of generality.[8]

## 3.2 Optimal mechanisms in the first-best case

In the first-best case where the expert's (two-dimensional) type is known and transfers are unrestricted, the customer's optimization problem is

$$\max_{x_{ik}, t_{ik}, w_{ik}, i \in \{\ell, h\}, k \in \{1,2\}} v - t_{ik} - w_{ik} \quad \text{s.t. } \forall i \in \{\ell, h\} \text{ and } \forall k \in \{1, 2\} \tag{$FB$}$$

$$t_{ik} - x_{ik} + b_k(v - t_{ik} - w_{ik}) \geq 0, \tag{$IR_{ik}$}$$

$$x_{ik} \geq x_i, \tag{$L_{ik}$}$$

$$w_{ik} \in [0, v - x_{ik}], \tag{$W_{ik}$}$$

where $IR_{ik}$ is expert type $\theta_{ik}$'s participation constraint, requiring that type $\theta_{ik}$'s utility from treating the customer is higher than her outside utility of zero. $L_{ik}$ is type $\theta_{ik}$'s liability constraint, requiring that undertreatment is infeasible and $W_{ik}$ specifies the lower and upper bound of $w_{ik}$. As the expert's (two-dimensional) type

[8] If the expert is behind in monetary terms, (the two-player version of) Fehr and Schmidt (1999) inequality aversion is equivalent to linear spite and (the two-player version of) Charness and Rabin (2002) quasi-maximin preferences are equivalent to linear altruism.

is known, there are no incentive constraints and different contracts can be offered to different expert types without paying information rents. The optimal mechanism for problem $(FB)$ is specified in Proposition 1.

**Proposition 1.** *The optimal mechanism for the first-best problem with unrestricted transfers* $(FB)$ *specifies contracts* $\mu^{FB}(\theta_{ik}) = \left(x_i, \frac{x_i - b_k v}{1-b_k}, 0\right)$ *for all* $i \in \{l, h\}$ *and all* $k \in \{1, 2\}$.

*Proof.* As participation constraint $IR_{ik}$ is tightened if the customer requests a high treatment when only a low treatment is needed, while the customer's valuation from treatment is unaffected, overtreatment cannot be optimal. As $v > x_h$, undertreatment cannot be optimal (however, this alternative is ruled out by liability constraint $L_{ik}$ anyway). Hence, appropriate treatment has to be optimal.

Burning surplus cannot be optimal: the utility of the spiteful expert increases by $(1 - b_1) > 1$ in transfers and increases only by $(-b_1) < 1$ in $w_{ik}$. Therefore, it is cheaper for the customer to increase transfers than to burn surplus. The utility of the altruistic expert decreases in $w_{ik}$ and thus cannot be optimal.

Finally, expert's participation constraint $IR_{ik}$ has to be binding. If it were not binding, the customer could increase his own payoff by decreasing transfers until it is binding without affecting other constraints. Hence, $x_{ik} = x_i$, $t_{ik} = (x_i - b_k v)/(1-b_k)$ and $w_{ik} = 0$ for all $\forall i \in \{\ell, h\}$ and $k \in \{1, 2\}$. ■

Proposition 1 implies that in the first-best case with unrestricted transfers and appropriate treatment is optimal and burning surplus is sub-optimal. Recall that $m = v - t$ is the expert's monetary payoff and that $o = v - t - w$ is the customer's monetary payoff. Thus, the optimal allocations $(m, o)$ following from Proposition 1 are

$$\left(-b_k \frac{v - x_i}{1 - b_k}, \frac{v - x_i}{1 - b_k}\right) \quad \forall i \in (\ell, h) \text{ and } \forall k \in (1, 2). \qquad (A^{FB}(\theta_{ik}))$$

Accordingly, the customer receives a fraction of $1/(1 - b_k) > 1/2$ of the surplus, while the expert receives a fraction of $b_k/(1 - b_k) < 1/2$ for $k \in \{1, 2\}$. In other words, the customer is ahead in monetary terms, as has been asserted in the previous subsection. If the expert is altruistic, the customer gets even more than the surplus generated from treatment – that is, $1/(1 - b_2) > 1$ – while the expert incurs a monetary loss.

The customer's optimization problem in the first-best case where the customer knows the expert's (two-dimensional) type and transfers are restricted is

$$(FB) \qquad \text{s.t. } \forall i \in \{\ell, h\} \text{ and } \forall k \in \{1, 2\} \qquad (FB_R)$$

$$t_{ik} \geq x_{ik} \qquad (R_{ik})$$

where constraint $R_{ik}$ requires that transfers be larger than the expert's treatment costs. The optimal mechanism in this case is specified in Proposition 2.

**Proposition 2.** *The optimal mechanism for the first-best problem with restricted transfers* $(FB_R)$ *specifies contracts* $\mu_R^{FB}(\theta_{\ell 1}) = \left(x_\ell, \frac{x_\ell - b_1 v}{1-b_1}, 0\right)$, $\mu_R^{FB}(\theta_{h1}) = \left(x_h, \frac{x_h - b_1 v}{1-b_1}, 0\right)$, $\mu_R^{FB}(\theta_{\ell 2}) = (x_l, x_l, 0)$ *and* $\mu_R^{FB}(\theta_{h2}) = (x_h, x_h, 0)$.

*Proof.* By the same arguments as in the proof of Proposition 1, it has to be optimal for the customer to request appropriate treatment and to burn no surplus for all expert types. If transfers for the spiteful expert type are lowered, participation constraints $IR_{\ell 1}$ and $IR_{h1}$ bind before the restriction on transfers becomes binding. Therefore, optimal transfers are the same as in Proposition 1. Conversely, if transfers for the altruistic expert type are decreased, constraints $R_{\ell 2}$ and $R_{h2}$ bind before participation constraints $IR_{\ell 2}$ and $IR_{h2}$. Consequently, transfers $t_{\ell 2} = x_\ell$ and $t_{h2} = x_h$ have to be optimal. ∎

Proposition 2 implies that given problem $(FB_R)$, appropriate treatment is optimal and burning surplus is sub-optimal. The optimal allocations $(m, o)$ following from Proposition 2 are

$$\left(-b_1 \tfrac{v-x_i}{1-b_1}, \tfrac{v-x_i}{1-b_1}\right) \quad \text{and} \qquad (A_R^{FB}(\theta_{i1}))$$

$$(0, v - x_i) \qquad (A_R^{FB}(\theta_{i2}))$$

for $i \in \{\ell, h\}$. If the expert is spiteful, the customer still receives a fraction of $1/(1-b_1) > 1/2$ of the surplus while the expert receives a fraction of $b_1/(1-b_1) < 1/2$). However, if the expert is altruistic, the customer now receives the entire surplus (and not more) and the expert receives a monetary payoff of zero.

# 4 Screening expert's distributional preferences

In this section we consider the (intermediate) second-best case where the customer knows his needs but not the expert's distributional preferences. For ease of exposition, we assume that the customer always requests and receives appropriate treatment and drop subscript $i \in \{l, h\}$. Thus, the customer solves the optimization problem

$$\max_{t_1,t_2,w_1,w_2} V = q_1(v - t_1 - w_1) + q_2(v - t_2 - w_2) \quad \text{s.t.} \qquad (SD)$$

$$t_1 - x + b_1(v - t_1 - w_1) \geq 0, \qquad (IR_1)$$

$$t_2 - x + b_2(v - t_2 - w_2) \geq 0, \qquad (IR_1)$$

$$t_1 - x + b_1(v - t_1 - w_1) \geq t_2 - x + b_1(v - t_2 - w_2), \qquad (IC_{12})$$

$$t_2 - x + b_2(v - t_2 - w_2) \geq t_1 - x + b_2(v - t_1 - w_1), \qquad (IC_{21})$$

$$w_1 \in [0, v - x] \qquad (W_1)$$

$$w_2 \in [0, v - x] \qquad (W_2)$$

where $IR_1$ and $IR_2$ are expert type 1's and 2's participation constraint while $W_1$ and $W_2$ specify the lower and upper bound of burned surplus $w_1$ and $w_2$. $IC_{12}$ is the incentive constraint requiring that expert type 1 prefers the contract specified for herself, $\mu^{SD}(\theta_1) = (x, t_1, w_1)$, to the contract specified for type 2, $\mu^{SD}(\theta_2) = (x, t_2, w_2)$. Accordingly, $IC_{21}$ requires that type 2 has to prefer contract $\mu^{SD}(\theta_2)$ to contract $\mu^{SD}(\theta_1)$.

Given problem $(SD)$, the customer can issue three kinds of mechanisms. The first possibility is to offer a *pooling mechanism* $\mu^P = (\mu^P(\theta_1), \mu^P(\theta_2))$, that assigns the same contract to both preference types. Then incentive constraints $IC_{12}$ and $IC_{21}$ hold trivially and only the constraints from the first-best problem remain. Hence, it cannot be optimal to burn surplus and it must be optimal to decrease transfers until one of the participation constraints $IR_1$ and $IR_2$ binds. As $b_1 < b_2$, this has to be participation constraint $IR_1$. So, both expert types are offered type 1's first-best contract $\mu^{FB}(\theta_1) = \left(x, \frac{x-b_1 v}{1-b_1}, 0\right)$. The allocations created by a pooling mechanism are therefore

$$\left(-b_1 \frac{v-x}{1-b_1}, \frac{v-x}{1-b_1}\right) \quad \text{for } k \in \{1,2\}. \tag{$A^P(\theta_k)$}$$

Thus, the expert always gets a fraction of $-b_1/(1-b_1)$ of the surplus. The customer always gets a fraction of $1/(1-b_1)$ of the surplus and has an expected payoff of $V = -b_1(v-x)/(1-b_1)$.

Alternatively, the customer can choose a *shut-down mechanism* $\mu^D = (\mu^D(\theta_1), \mu^D(\theta_2))$ that offers only contracts that are accepted by type 2, but are rejected by type 1. Again this implies that incentive constraints $IC_{12}$ and $IC_{21}$ hold trivially and it has to be optimal for the customer to decrease transfers until participation constraint $IR_2$ is binding. This implies that the customer assigns type 2's first-best contract $\mu^{FB}(\theta_2) = \left(x, \frac{x-b_2 v}{1-b_2}, 0\right)$ to both types. Thus, the allocations created by a shut-down mechanism are

$$(0,0) \quad \text{and} \tag{$A^D(\theta_1)$}$$

$$\left(-b_2 \frac{v-x}{1-b_2}, \frac{v-x}{1-b_2}\right). \tag{$A^D(\theta_2)$}$$

Thus, a spiteful expert gets no surplus[9], while an altruistic expert gets a (negative) 'fraction' of $-b_2/(1-b_2)$ of the surplus. The customer gets no surplus if he encounters a spiteful expert and a 'fraction' of $1/(1-b_2) > 1$ of the surplus if he encounters an altruistic expert. Thus, the customer's expected payoff is $V = q_2(v-x)/(1-b_2)$.

The customer's third possibility is to offer a *separating mechanism* $\mu^S = (\mu^S(\theta_1), \mu^S(\theta_2))$, where both preference types are active and both are assigned different contracts. Hence, incentive constraints $IC_{12}$ and $IC_{21}$ are no longer trivially satisfied. The optimal mechanisms for problem $(SD)$ are presented in Proposition 3.

**Proposition 3.** *The optimal mechanisms for problem* $(SD)$ *specify the following contracts*

- *if* $q_1 \in \left[0, \frac{b_2-b_1}{1-b_1}\right]$, $\mu^S(\theta_1) = (x, x, v-x)$ *and* $\mu^S(\theta_2) = \left(x, \frac{x-b_2 v}{1-b_2}, 0\right)$,
- *if* $q_1 \in \left(\frac{b_2-b_1}{1-b_1}, 1\right]$, $\mu^P(\theta_1) = \mu^P(\theta_2) = (x, \frac{x-b_1 v}{1-b_1}, 0)$.

[9] Recall that customer and expert both have an outside utility of zero. As the customer's and the expert's utility from allocation $(m, o) = (0,0)$ is zero as well, shutting down type 1 yields the same utility as if the customer would assign allocation $(0,0)$.

The proof is relegated to the appendix. Let $q_1^* = (b_2 - b_1)/(1 - b_1)$. Then the optimal allocations $(m, o)$ following from Proposition 3 are for $q_1 \leq q_1^*$

$$(0,0) \quad \text{and} \tag{$A^S(\theta_1)$}$$

$$\left(-b_2 \frac{v-x}{1-b_2}, \frac{v-x}{1-b_2}\right). \tag{$A^S(\theta_2)$}$$

For $q_1 > q_1^*$ the optimal allocations are $A^P(\theta_k)$, where $k \in \{1, 2\}$. Separation is achieved by burning a part of the surplus generated by the spiteful type (type 1) which creates a trade-off between surplus maximization and own monetary payoff for the altruistic type (type 2). In consequence, the customer's payoff from the spiteful type decreases, whereas the payoff from the altruistic type increases. Hence, separating preference types is optimal if the frequency of the spiteful type is relatively low ($q_1 \leq q_1^*$). Conversely, if the frequency of the spiteful type is relatively high ($q_1 > q_1^*$), it is more worthwhile to increase the payoff generated by the spiteful type and to forgo some potential payoff from the relatively infrequent altruistic type. Thus, the customer chooses pooling mechanism $\mu^P$ discussed above.

Note that due to the linear form of the customer's utility function only corner solutions can be optimal. Therefore, if it is optimal to separate preference types by destroying a part of type 1's surplus at all, it has to be optimal to destroy the entire surplus generated by type 1. That is, $w_1 = v - x$. Hence, the customer receives no surplus if he encounters a spiteful expert and a fraction of $1/(1 - b_2)$ of the surplus if he encounters an altruistic expert and has an expected payoff of $V = q_2(v - x)/(1 - b_2)$. If follows that the separating mechanism is equivalent to the shut-down mechanism discussed above and that the customer cannot gain from separating preference types.

To illustrate that separation of distributional preferences can benefit the customer, we will next consider the case where transfers have to exceed expert's treatment costs. The customer's optimization problem in this case is,

$$(SD) \qquad \text{s.t. } \forall k \in \{1, 2\} \tag{$SD_R$}$$

$$t_k \geq x \tag{$R_k$}$$

where constraint $R_k$ requires that transfers be larger than the expert's treatment costs. Note that the allocations implemented by a *pooling mechanism* $\mu_R^P$ are equivalent to the allocations implemented by $\mu^P$. That is, $A_R^P(\theta_k) = A^P(\theta_k)$. However, the allocations implemented by a shut-down mechanism $\mu_R^D$ are now

$$(0,0) \quad \text{and} \tag{$A_R^D(\theta_1)$}$$

$$(0, v-x). \tag{$A_R^D(\theta_2)$}$$

whereas the allocations implemented by a separating mechanism $\mu_R^S$ are (again) a priori unknown. The optimal mechanisms for problem ($SD_R$) are outlined in Proposition 4.

**Proposition 4.** *The optimal mechanisms for problem ($SD_R$) specify the following contracts*

– *if* $q_1 \in \left[0, \frac{b_2-b_1}{1-b_1}\right]$, $\mu_R^S(\theta_1) = \left(x, \frac{x-b_1v}{1-b_1} + b_1 g \frac{v-x}{1-b_1}, g(v-x)\right)$, *where* $g = -b_1 \frac{1-b_2}{b_2-b1} \in (0,1)$ *and* $\mu_R^S(\theta_2) = (x, x, 0)$,

– *if* $q_1 \in \left(\frac{b_2-b_1}{1-b_1}, 1\right]$, $\mu_R^P(\theta_1) = \mu_R^P(\theta_2) = \left(x, \frac{x-b_1v}{1-b_1}, 0\right)$.

The proof is given in the appendix. The optimal allocations $(m, o)$ following from Proposition 4 are for $q_1 \leq q_1^*$

$$\left(-b_1 \frac{1-g}{1-b_1}(v-x), \frac{1-g}{1-b_1}(v-x)\right) \quad \text{and} \qquad (A_R^S(\theta_1))$$

$$(0, v-x). \qquad (A_R^S(\theta_2))$$

For $q_1 > q_1^*$ the optimal allocations are $A_R^P(\theta_k)$, where $k \in \{1, 2\}$. As above, separating preference types is optimal if the frequency of the spiteful type is relatively low ($q_1 \leq q_1^*$) and pooling preference types is optimal if the frequency of the spiteful type is relatively high ($q_1 > q_1^*$). In the latter case, all expert types (again) receive $-b_1/(1-b_1)$ of the surplus and the customer receives $-1/(1-b_1)$ of the surplus and has an expected payoff of $V = -b_1(v-x)/(1-b_1)$.

If separation is optimal, the customer cannot gain by burning more surplus than necessary to hold the altruistic expert indifferent between choosing the spiteful expert's contract $\mu_R^{SD}(\theta_1)$ and giving up all the surplus to the customer. This is the case if $w = g(v-x)$. Thus, a spiteful expert receives a fraction of $-b_1(1-g)/(1-b_1)$ of the surplus while an altruistic expert receives no surplus at all. Furthermore, the customer receives a fraction of $(1-g)/(1-b_1)$ of the surplus if he encounters a spiteful expert and the entire surplus if he encounters an altruistic expert. It follows that the customer has an expected payoff of $V = (q_1(1-g)/(1-b_1) + q_2)(v-x)$. In contrast, a shut-down mechanism yields an expected payoff of $V = q_2(v-x)$, which is strictly lower if the spiteful type occurs with positive probability. Therefore, a separating mechanism clearly outperforms a shut-down mechanism.

Figure 1 shows allocations resulting from $(SD)$ and $(SD_R)$ in payoff space $(m, o)$. The surplus generated by treating the customer is $(v-x)$. The budget line, that is, the line satisfying $o = (v-x) - m = v - t$, puts an upper bound on the payoff the customer can get for himself. Note that the budget line is extended into the 2[nd] quadrant. This implies that the customer can (potentially) increase his payoff beyond $(v-x)$ by setting a transfer smaller than the expert's cost $x$. Due to liability, the expert cannot undertreat the customer. Hence, it can never be optimal for the customer to set transfers such that he gets an expected payoff below his outside utility of zero.

$I_1^0$ is the indifference curve where the spiteful expert (type 1) has a utility of zero. That is, $U_1 = m + b_1 o \equiv 0$ and hence the indifference curve is given by the function $o = (-m)/b_1$. As $b_1$ is smaller than zero, the indifference curves of the spiteful expert have a positive slope and better sets are towards the south-east of the figure. The customer cannot burn more than the entire surplus and cannot give up more than the entire surplus to the expert. Consequently, the set of possible allocations is bounded from below at $A^S(\theta_1) = (0, 0)$ and $A^m = (v-x, 0)$. The highest payoff the customer can get from expert type 1 is $(v-x)/(1-b_1)$, which is achieved if the customer implements allocation $A^P(\theta_k)$. This allocation can be determined graphically as the intersection between the budget line $o = (v - x) - m$ and $I_1^0$. Thus, the set of allocations type 1 will accept is contained in the triangle $\overline{A^S(\theta_1)A^P(\theta_k)A^m}$.

$I_2^0$ is the indifference curve where the altruistic expert (type 2) has a utility of zero. As $b_2 > 0$, type 2's indifference curves have a negative slope of $-1/b_2$ and better sets are towards the north-east. If transfers are unrestricted, allocation $A^S(\theta_2)$ yields the maximal payoff the customer can get from expert type 2. Thus, feasible allocations for type 2 are within the triangle $\overline{A^S(\theta_1)A^S(\theta_2)A^m}$. If the expert is not permitted to make losses, the maximal payoff the customer can get is the whole surplus in allocation $A_R^S(\theta_2)$. Thus, feasible allocations are in the triangle $\overline{A^S(\theta_1)A_S^R(\theta_2)A^m}$. Note that if offered $A^S(\theta_2)$ and $A^P(\theta_k)$ (or $A_R^S(\theta_2)$ and $A^P(\theta_k)$ if transfers are restricted), both a spiteful and an altruistic expert will prefer allocation $A^P(\theta_k)$. The reason for this is that for an altruistic expert, the indifference curve passing through this allocation is to the north-east of the indifference curves passing through $A^S(\theta_2)$ (that is, $I_2^0$) and $A_R^S(\theta_2)$ (that is, $I_2^R$).

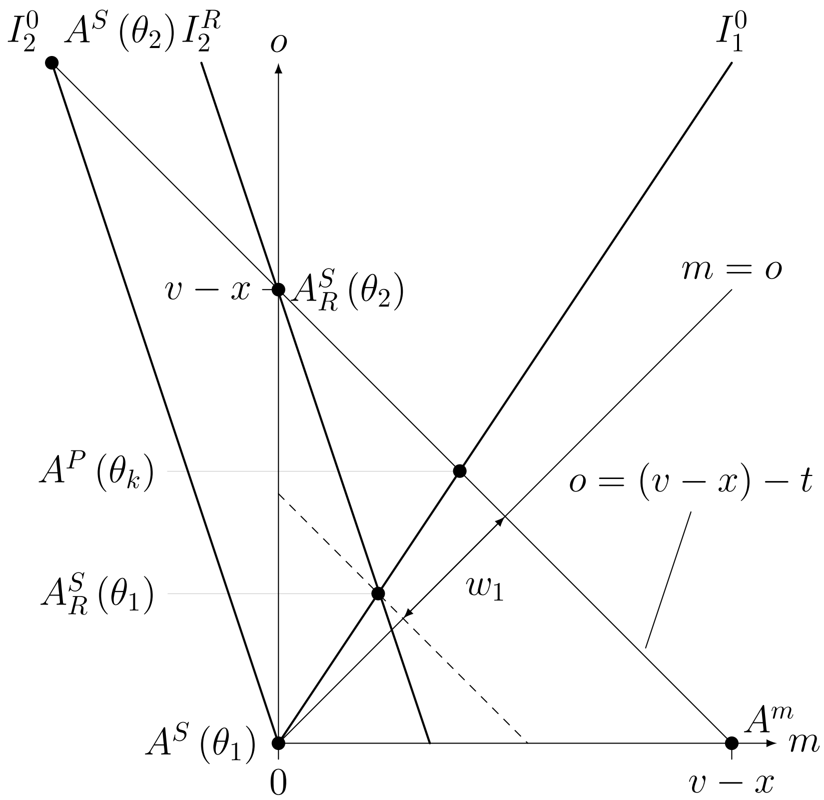


**Figure 1. Allocations and indifference curves in payoff space $(m, o)$ for $x_\ell = x_h = x$.**

By burning the entire surplus ($w_1 = v - x$), the customer can generate allocation $A^S(\theta_1)$ and induce the altruistic expert to choose $A^S(\theta_2)$. By burning a part of the surplus with size $w_1 = g(v - x)$, the customer can induce the altruistic expert to choose $A_R^S(\theta_2)$.

# 5 Screening expert's two-dimensional type

In the (proper) second-best case, the customer knows neither his needs nor the expert's distributional preferences. Thus, the customer solves the following optimization problem[10]:

$$\max_{x_{ik},t_{ik},w_{ik}} V = \sum_{i=\ell,h} \sum_{k=1,2} p_i q_k (v - t_{ik} - w_{ik}) \tag{$SB$}$$

$$\text{s.t.}\forall(i,j) \in \{\ell,h\} \text{ and } \forall(k,z) \in \{1,2\}$$
$$\text{with } i \neq j \text{ and/or } k \neq z,$$

$$t_{ik} - x_{ik} + b_k(v - t_{ik} - w_{ik}) \geq 0, \tag{$IR_{ik}$}$$

$$t_{ik} - x_{ik} + b_k(v - t_{ik} - w_{ik}) \geq t_{jz} - x_{jz} + b_k(v - t_{jz} - w_{jz}), \tag{$IC_{ik,jz}$}$$

$$x_{ik} \geq x_i, \tag{$L_{ik}$}$$

$$w_{ik} \in [0, v - x_{ik}]. \tag{$W_{ik}$}$$

As in previous sections, $IR_{ik}$ is type $\theta_{ik}$'s participation constraint. $IC_{ik,jz})$ is the incentive constraint requiring that type $\theta_{ik}$ prefers the contract specified for herself, $\mu(\theta_{ik})$, to the contract specified for type $\theta_{jz}$, $\mu(\theta_{jz})$, where $\forall(i,j) \in \{l,h\}$ and $\forall(k,z) \in \{1,2\}$ with $i \neq j$ and/or $k \neq z$. $L_{ik}$ is the liability constraint requiring that expert type $\theta_{ik}$ cannot provide low treatment if high treatment is needed. $W_{ik}$ specifies the lower and upper bound of $w_{ik}$. Note that due to liability, there are only 8 incentive constraints that can be potentially binding.

Now the customer has the possibility to offer one pooling contract for all types $\theta_{ik} \in \{\theta_{\ell 1}, \theta_{\ell 2}, \theta_{h1}, \theta_{h2}\}$, to separate all types by offering 4 different contracts or to pool some types but not others. Again, some extreme cases stand out. One case is a mechanism that is pooling preference types and separating problem types, $\mu^P = (\mu^P(\theta_{ik}))_{i=\ell h}^{k=1,2}$. In this case it cannot be optimal to burn surplus. Furthermore, it has to be optimal to decrease transfers until participation constraints $IR_{\ell 1}$ and $IR_{h1}$ are binding. Thus, types $\theta_{\ell 1}$ and $\theta_{\ell 2}$ are offered first-best contract $\mu^{FB}(\theta_{\ell 1})$ and types $\theta_{h1}$ and $\theta_{h2}$ are offered first-best contract $\mu^{FB}(\theta_{h1})$ and the mechanism creates allocations

$$\left(-b_1 \frac{v - x_i}{1 - b_1}, \frac{v - x_i}{1 - b_1}\right) \quad \forall i \in \{\ell, h\} \text{ and } \forall k \in \{1, 2\}. \tag{$A^P(\theta_{ik})$}$$

Therefore, this mechanism yields an expected payoff of $V = (v - p_l x_l - p_h x_h)/(1 - b_1)$. A second extreme case is a mechanism $\mu^D = (\mu^D(\theta_i k))_{i=\ell h}^{k=1,2}$ that shuts down spiteful types $\theta_{\ell 1}$ and $\theta_{h1}$ and assigns types $\theta_{\ell 2}$ and $\theta_{h2}$ their first-best contracts.[11] Thus, the allocations created by mechanism $\mu^D$ are for $i \in \{\ell, h\}$

$$(0, 0) \quad \text{and} \tag{$A^D(\theta_{i1})$}$$

$$\left(-b_2 \frac{v - x}{1 - b_2}, \frac{v - x}{1 - b_2}\right). \tag{$A^D(\theta_{i2})$}$$

[10]The expanded optimization problem is given in the Appendix.

[11]Again, it follows from Erharter (2012) that such a mechanism is incentive compatible.

This mechanism yields an expected payoff of $V = q_2(v - p_\ell x_\ell - p_h x_h))/(1 - b_2)$. The optimal mechanisms for the two-dimensional optimization problem with unrestricted transfers $(SB)$ are presented in Proposition 5.

**Proposition 5.** *The optimal mechanisms for problem $(SB)$ specify the following contracts*

- *if* $q_1 \in \left[0, \frac{b_2 - b_1}{1 - b_1}\right]$, $\mu^{SB}(\theta_{\ell 1}) = (x_\ell, x_\ell, v - x_\ell)$, $\mu^{SB}(\theta_{h1}) = (x_h, x_h, v - x_h)$, $\mu^{SB}(\theta_{\ell 2}) = \left(x_\ell, \frac{x_\ell - b_2 v}{1 - b_2}, 0\right)$ *and* $\mu^{SB}(\theta_{h2}) = \left(x_h, \frac{x_h - b_2 v}{1 - b_2}, 0\right)$,
- *if* $q_1 \in \left(\frac{b_2 - b_1}{1 - b_1}, 1\right]$, $\mu^{SB}(\theta_{\ell 1}) = \mu^{SB}(\theta_{\ell 2}) = \left(x_\ell, \frac{x_\ell - b_1 v}{1 - b_1}, 0\right)$ *and* $\mu^{SB}(\theta_{h1}) = \mu^{SB}(\theta_{h2}) = \left(x_h, \frac{x_h - b_1 v}{1 - b_1}, 0\right)$.

The proof is relegated to the Appendix. Recall that $q_1^* = (b_2 - b_1)/(1 - b_1)$. The optimal allocations $(m, o)$ following from Proposition 5 are for $q_1 \le q_1^*$ and $i \in \{\ell, h\}$

$$(0, 0) \quad \text{and} \qquad (A^S(\theta_{i1}))$$

$$\left(-b_2 \frac{v - x_i}{1 - b_2}, \frac{v - x_i}{1 - b_2}\right). \qquad (A^S(\theta_{i2}))$$

For $q_1 > q_1^*$ the optimal allocations are $A^P(\theta_{ik})$, for $i \in \{\ell, h\}$ and $k \in \{1, 2\}$. Proposition 5 states that it is optimal for the customer to separate preference types if the spiteful type is infrequent $(q_1 \le q_1^*)$ and to pool preference types if the spiteful type is frequent $(q_1 > q_1^*)$. If pooling is optimal, all expert types (again) receive $-b_1/(1 - b_1)$ of the surplus and the customer receives $1/(1 - b_1)$ of the surplus and has an expected payoff of $V = (v - p_\ell x_\ell - p_h x_h)/(1 - b_1)$.

If separation is optimal, the entire surplus generated by a spiteful type is burned and transfers for altruistic types are decreased until participation constraints $IR_{l2}$ and $IR_{h2}$ are binding. Thus, if $q_1 \le q_1^*$, an altruistic expert receives a (negative) fraction of $-b_2/(1 - b_2)$ of the surplus, while a spiteful expert receives nothing. If the customer is facing an altruistic expert he receives a fraction of $1/(1 - b_2)$ of the surplus, while if he encounters a spiteful expert he receives nothing. Thus, the customer's expected payoff is $V = q_2(v - p_\ell x_\ell - p_h x_h)/(1 - b_2)$ which is equivalent to the expected payoff generated by shut-down mechanism $(\mu^P)$. Hence, if transfers are unrestricted, a mechanism that separated preference types $(\mu^S)$ cannot outperform a mechanism that shuts down the spiteful type $(\mu^S)$, as was the case in the one-dimensional screening problem discussed in the previous section.

Note that mechanisms $\mu^S$ and $\mu^P$ both demand appropriate treatment. To see this, note that if the expert is spiteful and there is no liability, it might be optimal for the customer to demand undertreatment if the frequency of the high need is small and to demand overtreatment if the frequency of the high need is large, thus avoiding the worst-case scenario that the expert undertreats if the customer has a high need and overtreats if the customer has a low need. Due to liability, undertreatment is ruled out and accepting overtreatment as the lesser of two evils can no longer be optimal (Erharter, 2012). Nevertheless, in our two-dimensional setting it could be optimal for the customer to demand overtreatment from a spiteful expert in order to squeeze out a higher payoff from an altruistic expert. However, as discussed in the proof of Proposition 5 in the Appendix, it is cheaper for the customer to burn surplus directly. In both cases, the altruistic expert faces a trade-off between efficiency and her own monetary payoff. The difference is that if the customer demands overtreatment, he has to pay a higher transfer to the spiteful expert in order to keep her indifferent, while burning surplus allows the customer to decrease the transfer for the spiteful expert. If transfers are restricted, the customer's optimization problem is,

$$(SB) \qquad \text{s.t. } \forall i \in \{\ell, h\} \text{ and } \forall k \in \{1, 2\} \tag{$SB_R$}$$

$$t_{ik} \geq x_{ik} \tag{$R_{ik}$}$$

where constraint $R_{ik}$ requires that transfers be larger than the expert type $\theta_{ik}$'s treatment costs. Again, the allocations implemented by a *pooling mechanism* $\mu_R^P$ are equivalent to the allocations implemented by $\mu^D$. That is, $A_R^P(\theta_{ik}) = A^D(\theta_{ik})$. However, the allocations implemented by a shut-down mechanism $\mu_R^D$ are now

$$(0, 0) \quad \text{and} \tag{$A_R^D(\theta_{i1})$}$$

$$(0, v - x) \tag{$A_R^D(\theta_{i2})$}$$

for $i \in \{\ell, h\}$. The optimal mechanisms for problem $(SB_R)$ are outlined in Proposition 6.

**Proposition 6.** *The optimal mechanisms for problem $(SB_R)$ specify the following contracts for $i \in \{\ell, h\}$*

- *if* $q_1 \in \left[0, \frac{b_2 - b_1}{1 - b_1}\right]$, $\mu_R^S(\theta_{i1}) = \left(x_i, \frac{x_i - b_1 v}{1 - b_1} + b_1 g \frac{v - x_i}{1 - b_1}, g(v - x_i)\right)$, *where* $g = -b_1 \frac{1 - b_2}{b_2 - b1} \in (0, 1)$ *and* $\mu_R^S(\theta_{i2}) = (x_i, x_i, 0)$,
- *if* $q_1 \in \left(\frac{b_2 - b_1}{1 - b_1}, 1\right]$, $\mu_R^P(\theta_{i1}) = \mu_R^P(\theta_{i2}) = \left(x_i, \frac{x_i - b_1 v}{1 - b_1}, 0\right)$.

The proof is relegated to the Appendix. The optimal allocations $(m, o)$ following from Proposition 6 are for $q_1 \geq q_1^*$ and $i \in \{\ell, h\}$

$$(0, 0) \quad \text{and} \tag{$A_R^S(\theta_{i1})$}$$

$$\left(-b_2 \frac{v - x_i}{1 - b_2}, \frac{v - x_i}{1 - b_2}\right), \tag{$A_R^S(\theta_{i2})$}$$

and are for $q_1 > q_1^*$ given by $A_R^P(\theta_{ik})$ where all expert types (again) receive $-b_1/(1 - b_1)$ of the surplus and the customer receives $1/(1 - b_1)$ of the surplus and has an expected payoff of $V = (v - p_\ell x_\ell - p_h x_h)/(1 - b_1)$.

If $q_1 \leq q_1^*$, a spiteful expert receives a fraction of $-b_1(1 - g)/(1 - b_1)$ of the surplus while an altruistic expert receives no surplus at all. The customer receives a fraction of $(1 - g)/(1 - b_1)$ of the surplus if he encounters a spiteful expert and the entire surplus if he encounters an altruistic expert. Thus, the customer has an expected payoff of $V = (q_1(1-g)/(1-b_1)+q_2)(v-p_\ell x_\ell - p_h x_h)$, while a shut-down mechanism would only yield an expected payoff of $V = q_2(v - p_\ell x_\ell - p_h x_h)$. Therefore, a separating mechanism clearly outperforms a shut-down mechanism whenever spiteful experts occur with positive probability.

| problem | thresholds | $\boldsymbol{\theta_{\ell 1}}$ | $\boldsymbol{\theta_{\ell 2}}$ | $\boldsymbol{\theta_{h1}}$ | $\boldsymbol{\theta_{h2}}$ |
|---|---|---|---|---|---|
| $(FB)$ | | $A^P(\theta_{\ell k})$ | $A^S(\theta_{\ell 2})$ | $A^P(\theta_{hk})$ | $A^P(\theta_{h2})$ |
| $(FB_R)$ | | $A^P(\theta_{\ell k})$ | $A^S_R(\theta_{h1})$ | $A^P(\theta_{hk})$ | $A^S_R(\theta_{h2})$ |
| $(SB)$ | $q_1 \leq q_1^*$ | $A^S(\theta_{\ell 1})$ | $A^S(\theta_{\ell 2})$ | $A^S(\theta_{h1})$ | $A^S(\theta_{h2})$ |
| | $q_1 > q_1^*$ | $A^P(\theta_{\ell k})$ | $A^P(\theta_{\ell k})$ | $A^P(\theta_{hk})$ | $A^P(\theta_{hk})$ |
| $(SB_R)$ | $q_1 \leq q_1^*$ | $A^S_R(\theta_{\ell 1})$ | $A^S_R(\theta_{\ell 2})$ | $A^S_R(\theta_{h1})$ | $A^S_R(\theta_{h2})$ |
| | $q_1 > q_1^*$ | $A^P(\theta_{\ell k})$ | $A^P(\theta_{\ell k})$ | $A^P(\theta_{hk})$ | $A^P(\theta_{hk})$ |

**Table 1. Optimal allocations for problems $(FB)$,$(FB_R)$, $(SB)$ and $(SB_R)$.**

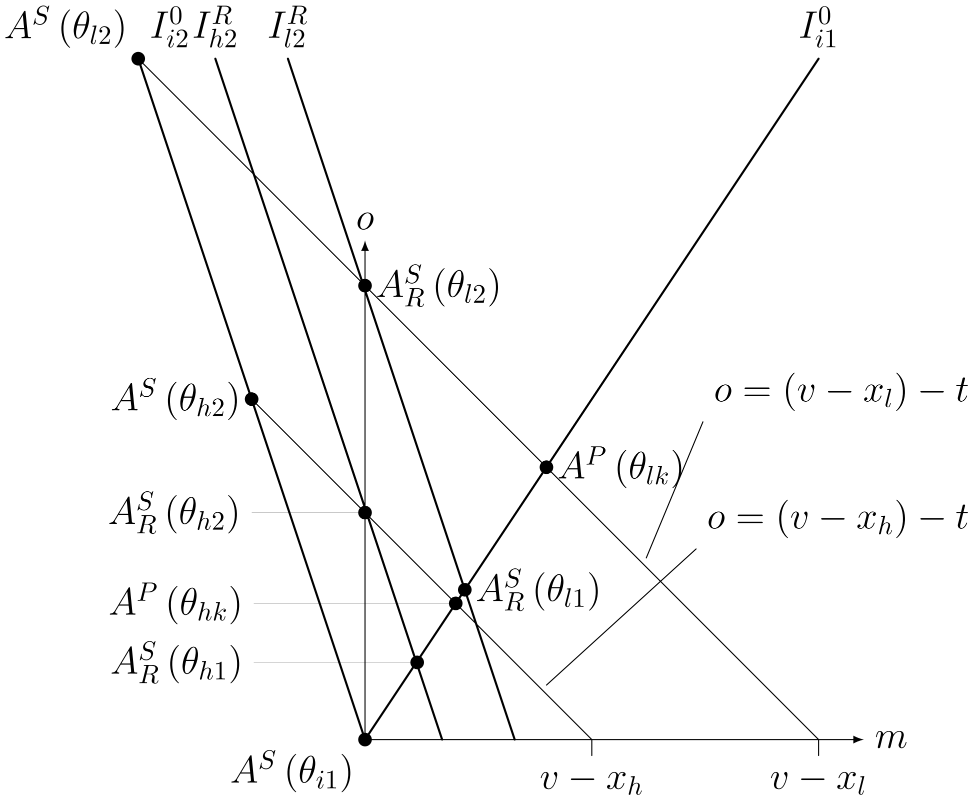


**Figure 2. Optimal allocations in payoff space.**

Optimal allocations for problems $(FB)$, $(FB_R)$, $(SB)$ and $(SB_R)$ are summarized in Table 1 and shown in Figure 2. $I^0_{i1}$ ($I^0_{i2}$) is the indifference curve where the spiteful (altruistic) expert has a utility of zero. $I^R_{\ell 2}$ ($I^R_{h2}$) is the indifference curve of an altruistic expert facing a low (high) need customer who has to give up the entire surplus (but not more than that) to the customer. If transfers are restricted and separation of preference types is optimal, allocations $A^S_R(\theta_{\ell 1})$ and $A^S_R(\theta_{h1})$ can be graphically derived as intersections of indifference curve $I^0_{i1}$ with indifference curves $I^R_{\ell 2}$ and $I^R_{h2}$ respectively.

# 6 Discussion and possible generalizations

In the previous sections we have illustrated that it is optimal to pool altruistic and spiteful experts if spiteful types are relatively frequent ($q_1 > q_1^*$) and that it is optimal to offer separate contracts to altruistic and spiteful types if spiteful types are sufficiently infrequent ($q_1 \leq q_1^*$). Threshold $q_1^* = (b_2 - b_1)/(1 - b_1)$ is identical for the one-dimensional screening problems discussed in section 4 and the two-dimensional screening problems discussed in section 5. This paper builds on a model framework introduced by Erharter (2012). This framework assumes (among other things) that transfers are unrestricted. Therefore, a mechanism that separates preferences types is always equivalent to a (simpler) mechanism that shuts down spiteful types. However, slight changes in these assumptions – such as a restriction of transfers – can render shut-down mechanisms sub-optimal to separating mechanisms.

Note however, that the optimality of a shut-down mechanism is highly sensitive to other assumptions as well. One typical assumption in the credence goods literature is that the customer is risk-neutral and maximizes his own monetary payoff. This implies that the customer's utility is linear in transfers $t$ and in burned surplus $w$. Therefore, derivatives with respect to t and w are zero and only corner solutions are optimal. As shown in previous sections, it is either optimal to burn the whole surplus of the spiteful expert, or no surplus at all. To the contrary, if the customer is risk-averse or if the customer cares for the expert's monetary payoff, his utility is non-linear in $t$ and $w$ and intermediate values of $w$ can be optimal. Another typical assumption is that the customer has an outside utility of zero. If this outside utility were below zero, the customer would get a negative utility from shutting down a spiteful type. Yet, in case of a separating contract, the customer would still get a monetary payoff – and hence a utility – of zero. Therefore, a separating mechanism would now outperform a shut-down mechanism even if transfers were unrestricted. On the other hand, if the customer's outside utility were above zero, a separating mechanism would yield a lower expected payoff than a shut-down mechanism.

Note that our framework implies that separation of preference types can only be optimal if at least one expert type is spiteful and one is altruistic. To see that one expert has to be spiteful, suppose again that the customer maximizes his expected monetary payoff, both players receive outside utilities of zero and that transfers have to be above the expert's treatment costs. If both expert types are altruistic and/or selfish (that is, if $b_2 > b_1 \geq 0$), both expert types are willing to give up the entire surplus to the customer. As this is the most the customer can get, burning surplus cannot be optimal. Thus, separation cannot be optimal. To see that one expert has to be altruistic, assume that both expert types are spiteful and/or selfish (that is, $b_1 < b_2 \leq 0$). If $b_2 < 0$, the customer has to burn the entire surplus of the more spiteful type and still does not get the entire surplus from the less spiteful type. Thus, constraints $R_{i2}$ with $i \in \{\ell, h\}$ are non-binding (transfers are above the expert's treatment costs) and therefore a separating mechanism cannot outperform a mechanism that shuts down the more spiteful type. If $b_2 = 0$, the customer still cannot get more than the entire surplus and constraints $R_{i2}$ are satisfied. Again, a separating mechanism cannot outperform a shut-down mechanism. As a corollary, if there are more than two types of distributional preferences, it can at most be optimal to separate two preference types. However, depending on the frequency of types, it can well be the case that a less altruistic type is pooled with a spiteful type, while a more altruistic type receives a different contract.

As outlined in Erharter (2012), it is important to ask how distributional preferences affect real world credence goods markets. Our analysis shows that it is important for the customer to consider differences in the expert's distributional preferences if altruistic types are sufficiently frequent. However, given the standard assumptions discussed in Erharter (2012), the customer can issue a shut-down mechanism instead of a more complicated mechanism that offers separate contracts to different preference types. In contrast, a regulator who aims to maximize total surplus cannot find it optimal to destroy surplus. Therefore, a regulator can simply offer a pooling mechanism fine-tuned to the (most) spiteful expert type and does not have to consider separating or shutting down preferences types.

# 7 Conclusion

In this paper we have derived optimal contracts for a credence goods expert who can either be altruistic or spiteful and who has private information about her preferences. We consider a one-dimensional screening problem, where a customer knows his needs but not the expert's preferences and a two-dimensional screening problem where the customer neither knows his own needs nor the expert's preferences. We show that in both cases, pooling preference types is optimal if spiteful types are frequent and separating preference types is optimal if spiteful types are infrequent. However, the optimal separating mechanism may be equivalent to a mechanism that shuts down spiteful types. However, slight changes of the customer's optimization problem can render shut-down mechanisms suboptimal to separating mechanisms. Independent of that, it is never optimal to offer identical contracts to different problem types in the two-dimensional screening problem. We point out how our results extend to more than two preference types and discuss the relevance of our results on actual credence goods markets.

# References


Andreoni, J. and Miller, J. (2002). Giving according to garp: An experimental test of the consistency of preferences for altruism. *Econometrica*, pages 737–753.

Bartling, B. and von Siemens, F. (2010a). Equal sharing rules in partnerships. *Journal of Institutional and Theoretical Economics JITE*, 166(2):299–320.

Bartling, B. and von Siemens, F. (2010b). The intensity of incentives in firms and markets: Moral hazard with envious agents. *Labour Economics*, 17(3):598–607.

Cabrales, A. and Calvò-Armengol, A. (2008). Interdependent preferences and segregating equilibria. *Journal of Economic Theory*, 139(1):99–113.

Charness, G. and Rabin, M. (2002). Understanding social preferences with simple tests. *Quarterly Journal of Economics*, 117(3):817–869.

Darby, M. and Karni, E. (1973). Free competition and the optimal amount of fraud. *The Journal of Law and Economics*, 16(1):67.

Demougin, D., Fluet, C., and Helm, C. (2006). Output and wages with inequality averse agents. *Canadian Journal of Economics/Revue canadienne d'économique*, 39(2):399–413.

Desiraju, R. and Sappington, D. (2007). Equity and adverse selection. *Journal of Economics & Management Strategy*, 16(2):285–318.

Dulleck, U. and Kerschbamer, R. (2006). On doctors, mechanics and computer specialists: The economics of credence goods. *Journal of Economic Literature*, 44(1):5–42.

Dulleck, U., Kerschbamer, R., and Sutter, M. (2009). The impact of distributional preferences on markets for credence goods. *University of Innsbruck*.

Dulleck, U., Kerschbamer, R., and Sutter, M. (2011). The economics of credence goods: An experiment on the role of liability, verifiability, reputation, and competition. *The American Economic Review*, 101(2):526–555.

Dur, R. and Glazer, A. (2008). Optimal contracts when a worker envies his boss. *Journal of Law, Economics, and Organization*, 24(1):120.

Englmaier, F. and Wambach, A. (2010). Optimal incentive contracts under inequity aversion. *Games and Economic Behavior*, 69(2):312–328.

Erharter, D. (2012). Credence goods markets, distributional preferences and the role of institutions.

Fehr, E., Klein, A., and Schmidt, K. (2007). Fairness and contract design. *Econometrica*, 75(1):121–154.

Fehr, E. and Schmidt, K. (1999). A theory of fairness, competition, and cooperation. *Quarterly Journal of Economics*, 114(3):817–868.

Fisman, R., Kariv, S., and Markovits, D. (2007). Individual preferences for giving. *American Economic Review*, 97(5):1858–1876.

Grund, C. and Sliwka, D. (2005). Envy and compassion in tournaments. *Journal of Economics & Management Strategy*, 14(1):187–207.

Itoh, H. (2004). Moral hazard and other-regarding preferences. *Japanese Economic Review*, 55(1):18–45.

Kerschbamer, R. (2010). A parsimonious experimental test for the type and intensity of distributional preferences at the individual level. *Mimeo, University of Innsbruck*.

Kosfeld, M. and von Siemens, F. (2009). Worker self-selection and the profits from cooperation. *Journal of the European Economic Association*, 7(2-3):573–582.

Kosfeld, M. and von Siemens, F. (2011). Competition, cooperation, and corporate culture. *The RAND Journal of Economics*, 42(1):23–43.

Levine, D. (1998). Modeling altruism and spitefulness in experiments. *Review of Economic Dynamics*, 1(3):593–622.

López-Pérez, R. (2008). Aversion to norm-breaking: A model. *Games and Economic Behavior*, 64(1):237–267.

Myerson, R. B. (1982). Optimal coordination mechanisms in generalized principal–agent problems. *Journal of mathematical economics*, 10(1):67–81.

Neilson, W. S. and Stowe, J. (2010). Piece-rate contracts for other-regarding workers. *Economic Inquiry*, 48(3):575–586.

Nelson, P. (1970). Information and consumer behavior. *Journal of Political Economy*, 78(2):311.

von Siemens, F. (2011). Heterogeneous social preferences, screening, and employment contracts. *Oxford Economic Papers*, 63(3):499.

# Appendix A The expanded optimization problem in the second-best case *(SB)*

Note that liability constraints $L_{ik}$ imply that $x_{h1} = x_{h2} = x_h$ and that constraints $IC_{h1,\ell1}$, $IC_{h1,\ell2}$, $IC_{h2,\ell1}$ and $IC_{h2,\ell2}$ are non-binding. Taking this into account, the (slightly simplified) expanded (proper) second best problem ($SB$) is

$$\max_{x_{ik},t_{ik},w_{ik},i\in\{\ell,h\},k\in\{1,2\}} V = p_\ell\left(q_1(v-t_{\ell1}-w_{\ell1}) + q_2(v-t_{\ell2}-w_{\ell2})\right) + \\ +p_h\left(q_1(v-t_{h1}-w_{h1}) + q_2(v-t_{h2}-w_{h2})\right) \quad \text{s.t.} \tag{SB}$$

$$t_{\ell1} - x_{\ell1} + b_1(v - t_{\ell1} - w_{\ell1}) \geq 0, \tag{$IR_{\ell1}$}$$
$$t_{\ell2} - x_{\ell2} + b_2(v - t_{\ell2} - w_{\ell2}) \geq 0, \tag{$IR_{\ell2}$}$$
$$t_h - x_{h1} + b_1(v - t_{h1} - w_{h1}) \geq 0, \tag{$IR_{h1}$}$$
$$t_h - x_{h2} + b_2(v - t_{h2} - w_{h2}) \geq 0, \tag{$IR_{h2}$}$$

$$t_{\ell1} - x_{\ell1} + b_1(v - t_{\ell1} - w_{\ell1}) \geq t_{\ell2} - x_{\ell2} + b_1(v - t_{\ell2} - w_{\ell2}), \tag{$IC_{\ell1,\ell2}$}$$
$$t_{\ell1} - x_{\ell1} + b_1(v - t_{\ell1} - w_{\ell1}) \geq t_{h1} - x_{h1} + b_1(v - t_{h1} - w_{h1}), \tag{$IC_{\ell1,h1}$}$$
$$t_{\ell1} - x_{\ell1} + b_1(v - t_{\ell1} - w_{\ell1}) \geq t_{h2} - x_{h2} + b_1(v - t_{h2} - w_{h2}), \tag{$IC_{\ell1,h2}$}$$

$$t_{\ell2} - x_{\ell2} + b_2(v - t_{\ell2} - w_{\ell2}) \geq t_{\ell1} - x_{\ell1} + b_2(v - t_{\ell1} - w_{\ell1}), \tag{$IC_{\ell2,\ell1}$}$$
$$t_{\ell2} - x_{\ell2} + b_2(v - t_{\ell2} - w_{\ell2}) \geq t_{h1} - x_{h1} + b_2(v - t_{h1} - w_{h1}), \tag{$IC_{\ell2,h1}$}$$
$$t_{\ell2} - x_{\ell2} + b_2(v - t_{\ell2} - w_{\ell2}) \geq t_{h2} - x_{h2} + b_2(v - t_{h2} - w_{h2}), \tag{$IC_{\ell2,h2}$}$$

$$t_{h1} - x_{h1} + b_1(v - t_{h1} - w_{h1}) \geq t_{h2} - x_{h2} + b_1(v - t_{h2} - w_{h2}), \tag{$IC_{h1,h2}$}$$
$$t_{h2} - x_{h2} + b_2(v - t_{h2} - w_{h2}) \geq t_{h1} - x_{h1} + b_2(v - t_{h1} - w_{h1}), \tag{$IC_{h2,h1}$}$$

$$w_{\ell1} \in [0, v - x_{\ell1}], \tag{$W_{\ell1}$}$$
$$w_{\ell2} \in [0, v - x_{\ell2}], \tag{$W_{\ell2}$}$$
$$w_{h1} \in [0, v - x_{h1}], \tag{$W_{h1}$}$$
$$w_{h2} \in [0, v - x_{h2}]. \tag{$W_{h2}$}$$

# Appendix B Proofs of Propositions 3–6

*Proof of Proposition 3.* Due to the linearity of problem ($SD$), the optimal mechanism is necessarily a corner solution that can be derived most efficiently by reducing the number of potentially binding constraints. Note that participation constraint $IR_1$ can be rearranged to $t_1 \geq (x - b_1 v + b_1 w_1)/(1 - b_1)$, $IR_2$ to $t_2 \geq (x - b_2 v + b_2 w_2)/(1 - b_2)$, $IC_{12}$ to $t_1 \geq t_2 + b_1(w_1 - w_2)/(1 - b_1)$ and $IC_{21}$ to $t_2 \geq t_1 - b_2(w_1 - w_2)/(1 - b_2)$. The problem can be simplified as follows:

(1) Suppose for the moment that $IC_{12}$ is satisfied.

(2) Transfer $t_1$ appears in (relevant) constraints $IR_1$, $IC_{21}$ and $W_2$. As $t_1$ tightens $IC_{2,1}$ and reduces the customer's expected payoff, it should be as small as possible and can be decreased until participation constraint $IR_1$ is binding. Thus, $t_1 = (x - b_1 v + b_1 w_1)/(1 - b_1)$ has to be optimal.

(3) Parameter $w_2$ appears in relevant constraints $IR_2$, $IC_{21}$ and $W_2$. As $w_2$ tightens constraints $IR_2$ and $IC_{21}$ and decreases the customer's expected utility, it should be as small as possible and can be decreased until $W_2$ is binding. Thus, $w_2 = 0$ has to be optimal.

(4) Transfer $t_2$ appears in (relevant) constraints $IR_2$ and $IC_{21}$. As this decreases the customer's expected utility, it should be decreased until one of the constraints is binding. Consider the (weak) inequality $t_1 - b_2/(1 - b_2)\ w_1 \geq (x - b_2 v + b_2 w_2)/(1 - b_2)$ where the left-hand side is derived from $IC_{21}$ and the right-hand side is derived from $IR_1$. This weak inequality can be transformed to $(x - b_1 v)/(1 - b_1) + b_1/(1 - b_1) w_1 \geq (x - b_2 v)/(1 - b_2) + w_1 b_2/(1 - b_2)$, which reduces to $v - x \geq w_1$. Due to $W_1$, this condition necessarily holds. Therefore, the binding constraint is $IC_{21}$ and $t_2 = t_1 - w_1 b_2/(1 - b_2)$ has to be optimal.

(5) Inserting transfers into participation constraint $IC_{12}$, we obtain $t_1 \geq t_1 - b_2/(1 - b_2)w_1 + b_1/(1 - b_1)w_1$ or $(b_1/(1 - b_1) - b_2/(1 - b_2))w_1 \geq 0$, which holds necessarily as $b_1 < 0 < b_2$ and $w_1 \in [0, v - x]$. Therefore, $IC_{12}$ is indeed satisfied.

Plugging transfers into the objective function of problem $(SD)$ yields the reduced problem

$$\max_{w_1} \quad \frac{v - x}{1 - b_1} + \left( \frac{b_2}{1 - b_2} - \frac{b_1}{1 - b_1} - \frac{q_1}{1 - b_2} \right) \text{ s.t.} \tag{$SD'$}$$

$$w_1 \in [0, v - x]. \tag{$W_1$}$$

As $w_1$ enters the customer's expected payoff linearly, only corner solutions can be optimal. If $b_2/(1 - b_2) - b_1/(1 - b_1) - q_1/(1 - b_2) > 0$ or $q_1 \leq (b_2 - b_1)/(1 - b_1)$, burning surplus (weakly) increases the customer's payoff and $w_1 = v - x$ has to be optimal. In contrast, if $q_1 > (b_2 - b_1)/(1 - b_1)$, $w_1 = 0$ has to be optimal. ■

*Proof of Proposition 4.* Note that observations (1)—(5) from the proof of Proposition 3 apply in the case with restricted payoffs as well. In particular, it is still the case that participation constraint $IR_1$ and incentive constraint $IC_{21}$ have to be binding in the optimal contract. However, as type 2 is altruistic and would be willing to accept a negative monetary payoff, constraint $R_2$, requiring that $t_2 \geq x$, has to be binding. Plugging $t_2 = x$ into $IC_{21}$ and solving for $w_1$ gives an upper bound of $w_1^* = -b_1 \frac{1 - b_2}{b_2 - b_1}(v - x) \in (0, v - x)$. For future reference, let $g = -b_1(1 - b_2)/(b_2 - b_1) \in (0, 1)$. It cannot be optimal for the customer to increase $w_1$ beyond $w_1^*$ regardless of the frequency of types, because this would decrease his payoff from type 1 without increasing his payoff from type 2. Plugging transfers into the customers objective function and setting an upper bound of $w_1^*$ in constraint $W_1$ yields the reduced problem

$$\max_{w_1} \quad \frac{v-x}{1-b_1} + \left(\frac{b_2}{1-b_2} - \frac{b_1}{1-b_1} - \frac{q_1}{1-b_2}\right) \text{ s.t.} \tag{$SD'_R$}$$

$$w_1 \in [0, g(v-x)]. \tag{$W_1$}$$

Again, $w_1$ enters the customer's utility linearly and only corner solutions can be optimal. If $q_1 \leq (b_2 - b_1)/(1-b_1)$, burning surplus increases the customer's payoff and $w_1 = g(v-x)$ has to be optimal. Plugging $w_1$ into transfers yields $t_1 = (x - b_1 v)/(1-b_1) + b_1 g(v-x)/(1-b_1)$ for the spiteful expert type and $t_2 = x$ for the altruistic expert types. In contrast, if $q_1 > (b_2 - b_1)/(1-b_1)$, $w_1 = 0$ has to be optimal and transfers are given by $t_1 = t_2 = (x - b_1 v)/(1-b_1)$. ■

*Proof of Proposition 5.* Again, the optimal mechanism has to be a corner solution that can be derived most efficiently by simplifying problem ($SB$).

(1) For the time being, suppose that incentive constraints $IC_{\ell 1,l2}$, $IC_{\ell 1,h2}$ and $IC_{h1,h2}$ are satisfied.

(2) Decreasing $t_{h1}$ until participation constraint $IR_{h1}$ is binding (weakly) increases the customer's expected payoff and relaxes incentive constraints $IC_{\ell 1,h1}$, $IC_{\ell 2,h1}$ and $IC_{\ell 2,h2}$ without affecting any other constraint. Therefore, $t_{h1} = (x_h - b_1 v + b_1 w_{h1})/(1-b_1)$ has to be optimal.

(3) Decreasing $t_{\ell 1}$ until participation constraint $IR_{\ell 1}$ and $IC_{\ell 1,h1}$ are (simultaneously) binding (weakly) increases the customer's expected payoff and relaxes incentive constraint $IC_{\ell 2,\ell 1}$. Therefore, $t_{\ell 1} = (x_{\ell 1} - b_1 v + b_1 w_{\ell 1})/(1-b_1)$ has to be optimal.

(4) Increasing $x_{\ell 2}$ beyond $x_\ell$ (weakly) decreases the customer's expected payoff and tightens participation constraint $IR_{\ell 2}$ and incentive constraints $IC_{\ell 2,\ell 1}$, $IC_{\ell 2,h1}$ and $IC_{\ell 2,h2}$. Therefore, $x_{\ell 2} = x_\ell$ has to be optimal.

(5) Increasing $x_{\ell 1}$ beyond $x_\ell$ (weakly) decreases the customer's expected payoff. However, as increasing $x_{\ell 1}$ relaxes incentive constraint $IC_{\ell 2,\ell 1}$, it is less straightforward whether appropriate treatment (i.e. $x_{\ell 1} = x_\ell$) is optimal in this case. Solving $IC_{\ell 2,\ell 1}$ for $t_{\ell 2}$ yields $t_{\ell 2} \geq t_{\ell 1} + \frac{x_\ell - x_{\ell 1}}{1-b_2} - \frac{b_2}{1-b_2}(w_{\ell 1} - w_{\ell 2})$. Replacing $t_{\ell 1}$ and rearranging yields $t_{\ell 2} \geq \frac{b_1 - b_2}{(1-b_1)(1-b_2)}(x_{\ell 1} - w_{\ell 1}) - \frac{b_1 v}{1-b_1} + \frac{x_\ell - b_2 w_{\ell 2}}{1-b_2}$.[12] As $(b_1 - b_2) < 0$, increasing $x_{\ell 1}$ clearly relaxes incentive constraint $IC_{\ell 2,\ell 1}$ – exactly to the same magnitude as increasing $w_{\ell 1}$. However, while increasing $w_{\ell 1}$ relaxes participation constraint $IR_{\ell 1}$ and incentive constraints $IC_{\ell 1,h1}$ and $IC_{\ell 1,h2}$, increasing $x_{\ell 1}$ tightens these constraints. In other words, it is cheaper for the customer to burn surplus than to demand overtreatment and $x_{\ell 1} = x_l$ is indeed optimal.

[12] Note that $t_{\ell 2} \geq \left(\frac{1}{1-b_1} - \frac{1}{1-b_2}\right) x_{\ell 1} + \left(\frac{b_1}{1-b_1} - \frac{b_2}{1-b_2}\right) w_{\ell 1} - \frac{b_1 v}{1-b_1} + \frac{x_\ell - b_2 w_{\ell 2}}{1-b_2}$ can be simplified to $t_{\ell 2} \geq \frac{1-b_2-1+b_1}{(1-b_1)(1-b_2)} + \frac{b_1 - b_1 b_2 - b_2 + b_1 b_2}{(1-b_1)(1-b_2)} w_{\ell 1} - \frac{b_1 v}{1-b_1} + \frac{x_\ell - b_2 w_{\ell 2}}{1-b_2} = \frac{b_1 - b_2}{(1-b_1)(1-b_2)}(x_{\ell 1} - w_{\ell 1}) - \frac{b_1 v}{1-b_1} + \frac{x_\ell - b_2 w_{\ell 2}}{1-b_2}$.

(6) Decreasing $w_{\ell 2}$ until $W_{\ell 2}$ is binding (weakly) increases the customer's expected payoff and relaxes participation constraint $IR_{\ell 2}$ and incentive constraints $IC_{\ell 2,\ell 1}$, $IC_{\ell 2,h1}$ and $IC_{\ell 2,h2}$ without affecting any other (relevant) constraint. Therefore, $w_{\ell 2} = 0$ has to be optimal.

(7) Decreasing $w_{h2}$ until $W_{h2}$ is binding (weakly) increases the customer's expected payoff and relaxes $IR_{h2}$ and $IC_{h2,h1}$. Furthermore, $IC_{h2,h1}$ implies that $IC_{\ell 2,h2}$ is unaffected by $w_{h2}$. To see this, note that $IC_{h2,h1}$ and $IC_{\ell 2,h2}$ can be rearranged to $t_{h2} \geq t_{h1} - \frac{b_2}{1-b_2}(w_{h1} - w_{h2})$ and $t_{\ell 2} \geq t_{h2} + \frac{x_h - x_\ell}{1-b_2} - \frac{b_2}{1-b_2} w_{h2}$ respectively. Combining both constraints yields $t_{\ell 2} \geq t_{h1} - \frac{b_2}{1-b_2}(w_{h1} - w_{h2}) + \frac{x_h - x_\ell}{1-b_2} - \frac{b_2}{1-b_2} w_{h2}$ and $w_{\ell 2}$ cancels out, so that $t_{\ell 2} \geq t_{h1} + \frac{x_h - x_l}{1-b_2} - \frac{b_2}{1-b_2} w_{h1}$. Therefore, $w_{\ell 2} = 0$ has to be optimal.

(8) Note furthermore that $IC_{h2,h1}$ and $IC_{\ell 2,h2}$ imply that $IC_{\ell 2,h1}$ has to be satisfied. To see this, note that $IC_{\ell 2,h1}$ can be rearranged to $t_{\ell 2} \geq t_{h1} - (x_h - x_\ell)/(1 - b_2) - b_2/(1 - b_2) w_{h1}$, which is equivalent to the combined constraints of $IC_{h2,h1}$ and $IC_{l2,h2}$ derived in observation (7).

(9) To sum up the results so far, we have found that $x_{\ell 1} = x_{\ell 2} = x_\ell$, $w_{\ell 2} = w_{h2} = 0$, while $t_{\ell 1} = \frac{x_\ell - b_1 v + b_1 w_{\ell 1}}{1-b_1}$ and $t_{h1} = \frac{x_h - b_1 v + b_1 w_{h1}}{1-b_1}$. Moreover, the number of relevant constraints has been reduced to participation constraints $IR_{\ell 2}$ and $IR_{h2}$, incentive constraints $IC_{\ell 2,\ell 1}$, $IC_{\ell 2,h2}$ and $IC_{h2,h1}$ and constraints $W_{\ell 1}$ and $W_{h1}$.

(10) Decreasing $t_{\ell 2}$ until $IC_{\ell 2,\ell 1}$ is binding weakly increases the customer's expected payoff, while $W_{\ell 1}$ implies that participation constraint $IR_{\ell 2}$ is still satisfied. To see this, note that $IC_{\ell 2,\ell 1}$ can be reformulated to $t_{\ell 2} \geq \frac{x_\ell - b_1 v}{1-b_1} + \left(\frac{b_1}{1-b_1} - \frac{b_2}{1-b_2}\right) w_{\ell 1}$, while $IR_{\ell 2}$ can be rearranged to $t_{\ell 2} \geq \frac{x_\ell - b_2 v}{1-b_2}$. Furthermore, note that $\frac{x_\ell - b_1 v}{1-b_1} + \left(\frac{b_1}{1-b_1} - \frac{b_2}{1-b_2}\right) w_{\ell 1} \geq \frac{x_\ell - b_2 v}{1-b_2}$ as $(b_2 - b_1)(v - x_\ell) \geq (b_2 - b_1) w_{\ell 1}$ according to $W_{\ell 1}$. Therefore, $t_{\ell 2} = \frac{x_\ell - b_1 v}{1-b_1} + \left(\frac{b_1}{1-b_1} - \frac{b_2}{1-b_2}\right) w_{\ell 1}$ has to be optimal.

(11) Decreasing $t_{h2}$ until $IC_{h2,h1}$ is binding weakly increases the customer's expected payoff, while $W_{h1}$ implies that participation constraint $IR_{\ell 2}$ is still satisfied. To see this, note that $\frac{x_h - b_1 v}{1-b_1} + \left(\frac{b_1}{1-b_1} - \frac{b_2}{1-b_2}\right) w_{h1} \geq \frac{x_h - b_2 v}{1-b_2}$ as $(b_2 - b_1)(v - x_h) \geq (b_2 - b_1) w_{h1}$ according to $W_{h1}$. Therefore, $t_{h2} = \frac{x_h - b_1 v}{1-b_1} + \left(\frac{b_1}{1-b_1} - \frac{b_2}{1-b_2}\right) w_{h1}$ has to be optimal.

(12) $IC_{\ell 2,h2}$ can be reformulated to $t_{\ell 2} \geq t_{h2} + \frac{x_h - x_\ell}{1-b_2}$. Plugging in $t_{\ell 2}$ and $t_{h2}$ as derived in observations (10) and (11) reveals that $IC_{\ell 2,h2}$ is satisfied if $w_{h1} \geq w_{\ell 1} - (x_h - x_\ell)$.

(13) To sum up once more, we now know that $x_{\ell 1} = x_{\ell 2} = x_\ell$, $w_{\ell 2} = w_{h2} = 0$, $t_{\ell 1} = \frac{x_\ell - b_1 v + b_1 w_{l1}}{1-b_1}$, $t_{h1} = \frac{x_h - b_1 v + b_1 w_{h1}}{1-b_1}$, $t_{\ell 2} = \frac{x_\ell - b_1 v}{1-b_1} + \left(\frac{b_1}{1-b_1} - \frac{b_2}{1-b_2}\right) w_{\ell 1}$ and $t_{h2} = \frac{x_h - b_1 v}{1-b_1} + \left(\frac{b_1}{1-b_1} - \frac{b_2}{1-b_2}\right) w_{h1}$. Parameter constraint $W_{\ell 1}$ is still relevant, while constraint $W_{h1}$ has to be reformulated to constraint $W'_{h1}$, requiring that $w_{h1} \in [w_{\ell 1} - (x_h - x_\ell), v - x_h]$.

(14) Finally, note that incentive constraints $IC_{\ell1,\ell2}$, $IC_{\ell1,h2}$ and $IC_{h1,h2}$ are indeed satisfied. $IC_{\ell1,\ell2}$ can be rearranged to $t_{\ell1} \geq t_{\ell2} + \frac{b_1}{1-b_1} w_{\ell1}$. Plugging in transfers yields $\frac{x_\ell - b_1 v + b_1 w_{\ell1}}{1-b_1} \geq \frac{x_\ell - b_1 v + b_1 w_{\ell1}}{1-b_1} - \frac{b_2 w_{\ell1}}{1-b_2}$, an expression that can be simplified to $w_{\ell1} \geq 0$, which is necessarily the case. $IC_{\ell1,h2}$ can be rearranged to $t_{\ell1} \geq t_{h2} - \frac{x_h - x_l}{1-b_1} + \frac{b_1}{1-b_1} w_{\ell1}$. Plugging in transfers yields $\frac{x_l - b_1 v + b_1 w_{\ell1}}{1-b_1} \geq \frac{x_h - b_1 v}{1-b_1} + \left(\frac{b_1}{1-b_1} - \frac{b_2}{1-b_2}\right) w_{h1} - \frac{x_h - x_\ell}{1-b_1} + \frac{b_1}{1-b_1} w_{\ell1}$, an expression that simplifies to $w_{h1} \geq 0$, which is always satisfied. Finally, $IC_{h1,h2}$ can be rearranged to $t_{h1} \geq t_{h2} + \frac{b_1}{1-b_1} w_{h1}$ an expression that again simplifies to $w_{h1} \geq 0$.

Plugging transfers into the objective function of problem $(SB)$ and replacing condition $W_{h1}$ by $W'_{h1}$ yields the reduced problem

$$\max_{w_{\ell1}, w_{h1}} \quad p_\ell \left( \frac{v - x_\ell}{1-b_1} + \left( \frac{b_2}{1-b_2} - \frac{b_1}{1-b_1} - \frac{q_1}{1-b_2} \right) w_{\ell1} \right) + \\ + p_h \left( \frac{v - x_h}{1-b_1} + \left( \frac{b_2}{1-b_2} - \frac{b_1}{1-b_1} - \frac{q_1}{1-b_2} \right) w_{h1} \right) \text{ s.t.} \tag{$SB'$}$$

$$w_{\ell1} \in [0, v - x_\ell]. \tag{$W_{\ell1}$}$$

$$w_{h1} \in [w_{\ell1} - (x_h - x_\ell), v - x_h]. \tag{$W'_{h1}$}$$

Again, the customer's payoff is linear in $w_{\ell1}$ and $w_{h1}$. Thus, if $\frac{b_2}{1-b_2} - \frac{b_1}{1-b_1} - \frac{q_1}{1-b_2} \geq 0$ or $q_1 \leq \frac{b_2 - b_1}{1-b_1}$, burning surplus (weakly) increases the customer's expected payoff and $w_{\ell1} = v - x_{\ell1}$ and $w_{h1} = v - x_{h1}$ have to be optimal. Note that this implies that types $\theta_{\ell1}$ and $\theta_{h1}$ receive transfers of $t_{\ell1} = x_\ell$ and $t_{h1} = x_h$, while types $\theta_{\ell2}$ and $\theta_{h2}$ are assigned their (possibly negative) first-best transfers, $t_{\ell2} = \frac{x_\ell - b_2 v}{1-b_2}$ and $t_{h2} = \frac{x_h - b_2 v}{1-b_2}$.

To the contrary, if $q_1 > \frac{b_2 - b_1}{1-b_1}$, burning surplus decreases the customer's expected payoff and $w_{\ell1} = w_{h1} = 0$ has to be optimal. This implies that types $\theta_{\ell1}$ and $\theta_{\ell2}$ assigned transfers $t_{\ell1} = t_{\ell2} = \frac{x_\ell - b_1 v}{1-b_1}$, while types $\theta_{h1}$ and $\theta_{h2}$ are assigned transfers $t_{h1} = t_{h2} = \frac{x_h - b_1 v}{1-b_1}$. ■

*Proof of Proposition 6.* Note that observations (1)—(14) from the proof of Proposition 5 equally apply in the case with restricted payoffs. However, as types $\theta_{\ell2}$ and $\theta_{h2}$ would be willing to accept a negative monetary payoff if the customer is willing to burn a large enough chunk of the surplus, constraints $R_{\ell2}$ and $R_{h2}$ are potentially binding. Recall from the proof of Proposition 4 that $g = -b_1 \frac{1-b_2}{b_2-b_1} \in (0,1)$. Plugging $t_{\ell2} = x_\ell$ into $IC_{\ell2,\ell1}$ and solving for $w_{\ell1}$ yields an upper bound of $w^*_{\ell1} = g(v - x_\ell)$. Analogously, plugging $t_{h2} = x_h$ into $IC_{h2,h1}$ yields an upper bound of $w^*_{h1} = g(v - x_h)$. It cannot be optimal for the customer to increase $w_{\ell1}$ ($w_{h1}$) beyond $w^*_{\ell1}$ ($w^*_{h1}$) regardless of the frequency of types, because this would decrease his payoff from type $\theta_{\ell1}$ ($\theta_{h1}$), without increasing his payoff from type $\theta_{\ell2}$ ($\theta_{h2}$). Plugging transfers into the customers objective function and setting an upper bound of $w^*_{\ell1}$ in constraint

$W_{\ell 1}$ and an upper bound of $w^*_{h1}$ in $W'_{h1}$ yields the reduced problem

$$\max_{w_{\ell 1}, w_{h1}} \quad p_\ell \left( \frac{v - x_\ell}{1 - b_1} + \left( \frac{b_2}{1 - b_2} - \frac{b_1}{1 - b_1} - \frac{q_1}{1 - b_1} \right) w_{\ell 1} \right) + \qquad (SB'_R)$$

$$+ p_h \left( \frac{v - x_h}{1 - b_1} + \left( \frac{b_2}{1 - b_2} - \frac{b_1}{1 - b_1} - \frac{q_1}{1 - b_1} \right) w_{h1} \right) \text{ s.t.}$$

$$w_{\ell 1} \in [0, g(v - x_\ell)]. \qquad (W'_{\ell 1})$$

$$w_{h1} \in [w_{\ell 1} - (x_h - x_\ell), g(v - x_h)]. \qquad (W''_{h1})$$

Again, $w_{\ell 1}$ and $w_{h1}$ enter the customer's utility function linearly and only corner solutions can be optimal. If $q_1 \leq (b_2 - b_1)/(1 - b_1)$, burning surplus increases the customer's payoff and $w_{\ell 1} = g(v - x_\ell)$ and $w_{h1} = g(v - x_h)$ have to be optimal. Plugging $w_{\ell 1}$ and $w_{h1}$ into transfers yields $t_{\ell 1} = \frac{x_\ell - b_1 v}{1 - b_1} + b_1 g \frac{v - x_\ell}{1 - b_1}$ and $t_{h1} = \frac{x_h - b_1 v}{1 - b_1} + b_1 g \frac{v - x_h}{1 - b_1}$ for the spiteful expert types and $t_{\ell 2} = x_\ell$ and $t_{h2} = x_h$ for the altruistic expert types. If $q_1 > \frac{b_2 - b_1}{1 - b_1}$, burning surplus decreases the customer's expected payoff and (as in the proof of Proposition 5) $w_{\ell 1} = w_{h1} = 0$ has to be optimal, resulting in transfers $t_{\ell 1} = t_{\ell 2} = \frac{x_\ell - b_1 v}{1 - b_1}$ for the low problem types and $t_{h1} = t_{h2} = \frac{x_h - b_1 v}{1 - b_1}$ for the high problem types. ■

**University of Innsbruck - Working Papers in Economics and Statistics**
**Recent Papers** can be accessed on the following webpage:

http://eeecon.uibk.ac.at/wopec/

2013-27 **Dominik Erharter:** Screening experts' distributional preferences

2013-26 **Loukas Balafoutas, Rudolf Kerschbamer, Matthias Sutter:** Second-degree moral hazard in a real-world credence goods market

2013-25 **Rudolf Kerschbamer:** The geometry of distributional preferences and a non-parametric identification approach

2013-24 **Nadja Klein, Michel Denuit, Stefan Lang, Thomas Kneib:** Nonlife ratemaking and risk management with bayesian additive models for location, scale and shape

2013-23 **Nadja Klein, Thomas Kneib, Stefan Lang:** Bayesian structured additive distributional regression

2013-22 **David Plavcan, Georg J. Mayr, Achim Zeileis:** Automatic and probabilistic foehn diagnosis with a statistical mixture model

2013-21 **Jakob W. Messner, Georg J. Mayr, Achim Zeileis, Daniel S. Wilks:** Extending extended logistic regression to effectively utilize the ensemble spread

2013-20 **Michael Greinecker, Konrad Podczeck:** Liapounoff's vector measure theorem in Banach spaces *forthcoming in Economic Theory Bulletin*

2013-19 **Florian Lindner:** Decision time and steps of reasoning in a competitive market entry game

2013-18 **Michael Greinecker, Konrad Podczeck:** Purification and independence

2013-17 **Loukas Balafoutas, Rudolf Kerschbamer, Martin Kocher, Matthias Sutter:** Revealed distributional preferences: Individuals vs. teams

2013-16 **Simone Gobien, Björn Vollan:** Playing with the social network: Social cohesion in resettled and non-resettled communities in Cambodia

2013-15 **Björn Vollan, Sebastian Prediger, Markus Frölich:** Co-managing common pool resources: Do formal rules have to be adapted to traditional ecological norms?

2013-14 **Björn Vollan, Yexin Zhou, Andreas Landmann, Biliang Hu, Carsten Herrmann-Pillath:** Cooperation under democracy and authoritarian norms

- 2013-13 **Florian Lindner, Matthias Sutter:** Level-k reasoning and time pressure in the 11-20 money request game *forthcoming in Economics Letters*
- 2013-12 **Nadja Klein, Thomas Kneib, Stefan Lang:** Bayesian generalized additive models for location, scale and shape for zero-inflated and overdispersed count data
- 2013-11 **Thomas Stöckl:** Price efficiency and trading behavior in limit order markets with competing insiders *forthcoming in Experimental Economics*
- 2013-10 **Sebastian Prediger, Björn Vollan, Benedikt Herrmann:** Resource scarcity, spite and cooperation
- 2013-09 **Andreas Exenberger, Simon Hartmann:** How does institutional change coincide with changes in the quality of life? An exemplary case study
- 2013-08 **E. Glenn Dutcher, Loukas Balafoutas, Florian Lindner, Dmitry Ryvkin, Matthias Sutter:** Strive to be first or avoid being last: An experiment on relative performance incentives.
- 2013-07 **Daniela Glätzle-Rützler, Matthias Sutter, Achim Zeileis:** No myopic loss aversion in adolescents? An experimental note
- 2013-06 **Conrad Kobel, Engelbert Theurl:** Hospital specialisation within a DRG-Framework: The Austrian case
- 2013-05 **Martin Halla, Mario Lackner, Johann Scharler:** Does the welfare state destroy the family? Evidence from OECD member countries
- 2013-04 **Thomas Stöckl, Jürgen Huber, Michael Kirchler, Florian Lindner:** Hot hand belief and gambler's fallacy in teams: Evidence from investment experiments
- 2013-03 **Wolfgang Luhan, Johann Scharler:** Monetary policy, inflation illusion and the Taylor principle: An experimental study
- 2013-02 **Esther Blanco, Maria Claudia Lopez, James M. Walker:** Tensions between the resource damage and the private benefits of appropriation in the commons
- 2013-01 **Jakob W. Messner, Achim Zeileis, Jochen Broecker, Georg J. Mayr:** Improved probabilistic wind power forecasts with an inverse power curve transformation and censored regression
- 2012-27 **Achim Zeileis, Nikolaus Umlauf, Friedrich Leisch:** Flexible generation of e-learning exams in R: Moodle quizzes, OLAT assessments, and beyond

2012-26 **Francisco Campos-Ortiz, Louis Putterman, T.K. Ahn, Loukas Balafoutas, Mongoljin Batsaikhan, Matthias Sutter:** Security of property as a public good: Institutions, socio-political environment and experimental behavior in five countries

2012-25 **Esther Blanco, Maria Claudia Lopez, James M. Walker:** Appropriation in the commons: variations in the opportunity costs of conservation

2012-24 **Edgar C. Merkle, Jinyan Fan, Achim Zeileis:** Testing for measurement invariance with respect to an ordinal variable *forthcoming in Psychometrika*

2012-23 **Lukas Schrott, Martin Gächter, Engelbert Theurl:** Regional development in advanced countries: A within-country application of the Human Development Index for Austria

2012-22 **Glenn Dutcher, Krista Jabs Saral:** Does team telecommuting affect productivity? An experiment

2012-21 **Thomas Windberger, Jesus Crespo Cuaresma, Janette Walde:** Dirty floating and monetary independence in Central and Eastern Europe - The role of structural breaks

2012-20 **Martin Wagner, Achim Zeileis:** Heterogeneity of regional growth in the European Union

2012-19 **Natalia Montinari, Antonio Nicolo, Regine Oexl:** Mediocrity and induced reciprocity

2012-18 **Esther Blanco, Javier Lozano:** Evolutionary success and failure of wildlife conservancy programs

2012-17 **Ronald Peeters, Marc Vorsatz, Markus Walzl:** Beliefs and truth-telling: A laboratory experiment

2012-16 **Alexander Sebald, Markus Walzl:** Optimal contracts based on subjective evaluations and reciprocity

2012-15 **Alexander Sebald, Markus Walzl:** Subjective performance evaluations and reciprocity in principal-agent relations

2012-14 **Elisabeth Christen:** Time zones matter: The impact of distance and time zones on services trade

2012-13 **Elisabeth Christen, Joseph Francois, Bernard Hoekman:** CGE modeling of market access in services

2012-12 **Loukas Balafoutas, Nikos Nikiforakis:** Norm enforcement in the city: A natural field experiment *forthcoming in European Economic Review*

2012-11 **Dominik Erharter:** Credence goods markets, distributional preferences and the role of institutions

2012-10 **Nikolaus Umlauf, Daniel Adler, Thomas Kneib, Stefan Lang, Achim Zeileis:** Structured additive regression models: An R interface to BayesX

2012-09 **Achim Zeileis, Christoph Leitner, Kurt Hornik:** History repeating: Spain beats Germany in the EURO 2012 Final

2012-08 **Loukas Balafoutas, Glenn Dutcher, Florian Lindner, Dmitry Ryvkin:** The optimal allocation of prizes in tournaments of heterogeneous agents

2012-07 **Stefan Lang, Nikolaus Umlauf, Peter Wechselberger, Kenneth Harttgen, Thomas Kneib:** Multilevel structured additive regression

2012-06 **Elisabeth Waldmann, Thomas Kneib, Yu Ryan Yu, Stefan Lang:** Bayesian semiparametric additive quantile regression

2012-05 **Eric Mayer, Sebastian Rueth, Johann Scharler:** Government debt, inflation dynamics and the transmission of fiscal policy shocks *forthcoming in Economic Modelling*

2012-04 **Markus Leibrecht, Johann Scharler:** Government size and business cycle volatility; How important are credit constraints? *forthcoming in Economica*

2012-03 **Uwe Dulleck, David Johnston, Rudolf Kerschbamer, Matthias Sutter:** The good, the bad and the naive: Do fair prices signal good types or do they induce good behaviour?

2012-02 **Martin G. Kocher, Wolfgang J. Luhan, Matthias Sutter:** Testing a forgotten aspect of Akerlof's gift exchange hypothesis: Relational contracts with individual and uniform wages

2012-01 **Loukas Balafoutas, Florian Lindner, Matthias Sutter:** Sabotage in tournaments: Evidence from a natural experiment *published in Kyklos*

University of Innsbruck

Working Papers in Economics and Statistics

2013-27

Dominik Erharter

Screening experts' distributional preferences

**Abstract**
We study optimal direct mechanisms for a credence goods expert who can be altruistic or spiteful. The expert has private information about her distributional preferences and possibly also about her customer's needs. We introduce a method that allows the customer to offer separate contracts to different preference types and outline when separation is optimal. Furthermore, we demonstrate that the optimality of separating mechanisms is sensitive to minor changes of the customer's utility function. Additionally, we illustrate how our results extend to more than two preference types and discuss possible policy implications.

ISSN 1993-4378 (Print)
ISSN 1993-6885 (Online)