Incentives and the Sorting of Altruistic Agents into Street-Level Bureaucracies

Margaretha Buurman* and Robert Dur†

July 14, 2008
Preliminary version

Abstract

Many street-level bureaucrats (such as caseworkers) have the dual task of helping some clients, while sanctioning others. We develop a model of such a street-level bureaucracy and study the implications of its personnel policy on the self-selection and allocation decisions of agents who differ in altruism towards clients. When bureaucrats are paid flat wages, they do not sanction, and the most altruistic types sort into bureaucracy. Pay-for-performance induces some bureaucrats to sanction, but necessitates an increase in expected wage compensation, which can result in sorting from both the top and bottom of the altruism distribution. We also show how client composition affects sorting and why street-level bureaucrats often experience an overload of clients.

Keywords: street-level bureaucracy, sorting, altruism, personnel policy, pay-for-performance.

JEL-codes: J3, J4, L3, M5

*Department of Economics, Erasmus University Rotterdam. Address: Erasmus University, FEW/H 8-34, P.O. Box 1738, 3000 DR Rotterdam, The Netherlands. E-mail: buurman@few.eur.nl

†Department of Economics, Erasmus University Rotterdam, Tinbergen Institute, and CESifo. Address: Erasmus University, FEW/H 8-15, P.O. Box 1738, 3000 DR Rotterdam, The Netherlands. E-mail: dur@few.eur.nl

1 Introduction

Street-level bureaucrats often have the dual task of helping some clients while disciplining others. Caseworkers are a case in point. On the one hand, their job is to allocate employment services and give job search assistance to clients who are willing but unable to find a job. On the other hand, they are supposed to sanction clients who rather live on a benefit than work from 9 to 5.[1]

The dual nature of the job implies that it is not straightforward to say what kind of people should optimally be hired by such street-level bureaucracies. While the helping aspect of the job makes altruistic or client-oriented people the ideal candidates, these people are likely to take clients' interests too much into account when encountering clients who should be sanctioned. In addition to this normative issue of what would be optimal candidates, the positive issue of what kind of people find a career in a street-level bureaucracy actually worthwhile is perhaps even more important. While assessment centers and talented HR managers may give agencies a glimpse of job applicants' motivations, their true motivations often remain hidden, implying that agencies should use other, more implicit instruments to promote self-selection of the most desired types of workers. These may include paying low base salaries and offering bonuses for good performance.

This paper studies these issues by developing a model of a street-level bureaucracy that serves different types of clients, some of which are in need of help (willing but unable clients) and others who should be sanctioned (non-willing clients). In addition, there exists a group of clients who should neither be helped nor sanctioned (willing and able clients). The agency hires bureaucrats whose task is to meet clients, assess their type, and allocate either help, no help, or a sanction. Bureaucrats are hired from a pool of potential job applicants who differ in their altruism towards clients they meet, ranging from complete indifference to highly altruistic. The agency cannot observe job applicants' types. However, it can affect the sorting of job applicants by its personnel policy. We study two different settings which are often observed in practice: 1) the bureaucracy pays a base salary only; 2) on top of the base salary, the bureaucracy offers bonus pay or nonpecuniary rewards for good performance. We obtain the following results.

[1] Other examples easily come to mind. For instance, teachers' main task is to help students learn, but from time to time their job also involves expelling disruptive students from the classroom. Soldiers taking part in peacekeeping missions often engage in both humanitarian activities and combat. And police officers both help and sanction people.

First, when bureaucrats' compensation consists of a base salary only, the bureaucrats' behavior is in line with the agency's preferences, except for the allocations to non-willing clients. As bureaucrats are (weakly) altruistic towards clients, bureaucrats do not sanction non-willing clients, but allocate no help instead. The most altruistic types among the potential job applicants self-select into the bureaucracy. Besides the base salary, the attractiveness of the job depends on the composition of the client population. In particular, the job becomes more attractive, and hence the base salary can be lower, when there are more clients in need of help. If the agency has monopsony power, it may be optimal to hire fewer agents than necessary to serve all clients, so as to reduce salary costs. Our model can thus explain why street-level bureaucracies are often plagued by limited resources and an overload of clients, as observed by e.g. Lipsky (1980).

Second, bonus pay (or nonpecuniary rewards) for good performance induces some bureaucrats (the least altruistic among the hired bureaucrats) to sanction non-willing clients. Generally, it is optimal for the agency to set bonus pay such that it induces only part of the bureaucrats to sanction: some bureaucrats care so much about the feelings of non-willing clients that it is too costly to induce them to sanction. Besides affecting bureaucrats' decisions, we show that bonus pay can be used by the agency to extract rents from the most altruistic bureaucrats. Since these bureaucrats do not sanction, a rise in bonus pay increases their income by less than the income of bureaucrats whose participation constraint binds. Optimal bonus pay is therefore higher than the value of sanctioning for the agency.

Third, the effect of pay-for-performance on the sorting of agents into bureaucracy depends crucially on whether the expected joys of helping the willing and unable clients exceed the expected sorrows of sanctioning non-willing clients. If the client population consists mainly of people in need of help, and the willing clients' benefit from help is high relative to the unwilling clients' pain of sanctions, there is still self-selection of the most altruistic types into the job. If this condition does not hold, the only way through which the agency can make sure that at least some of its agents sanction non-willing clients is by offering a combination of base salary and bonus pay that brings higher pecuniary payoffs than the outside option of agents willing to sanction. As a result, there is sorting from the top and the bottom of the altruism distribution, with highly altruistic agents choosing no sanction for the non-willing clients and earning low income and low-altruistic agents imposing sanctions and earning high bonus pay. When the bureaucracy values sanctions for non-willing clients a lot, the bureaucracy may optimally set personnel policy such that it only hires agents from the bottom of the altruism distribution. We

thus show that bonus pay can have a profound impact on the type of workers self-selecting into street-level bureaucracies.

We proceed as follows. The next section describes how our paper relates to the literature and discusses some stylized facts about the motivation of caseworkers, which we take as the leading example in our paper. Section 3 describes the model. Section 4 analyzes the simple case with flat wages and section 5 studies pay-for-performance. Finally, in section 6 we make some concluding remarks.

2 Related literature and some stylized facts

Our paper contributes to a recent literature in economics on incentives and workers' motivation in the public sector (Francois, 2000 and 2007, Dixit, 2002, Glazer, 2004, Besley and Ghatak, 2005, Prendergast, 2007, Delfgaauw and Dur, 2008, and Delfgaauw, 2008). Closest to our paper is Prendergast (2007) who studies sorting of purely altruistic agents into a street-level bureaucracy. There are four key differences between his paper and ours. First, we focus on jobs which involve a *dual* task of helping some clients and sanctioning others. Second, while in Prendergast (2007) bureaucrats earn flat wages, we allow bureaucracies to use incentives. Third, while Prendergast (2007) focuses on effort provision of the agents (assuming that agents cannot lie about the client's type), we assume that agent's information about client's type does not involve effort cost and is soft, giving discretion to the agent in his allocation decision. Last, we assume that agents are impurely altruistic in the sense that they only care about clients they meet and we abstract from hostile agents. We discuss the implications of these latter two assumptions along the way.

There is abundant evidence that a substantial part of people working in street-level bureaucracies are concerned about clients. Lipsky (1980: 72) observes that 'Those who recruit themselves for public service work are attracted to some degree by the prospect that their lives will gain meaning through helping others.' More recently, Le Grand (2003: 38) concludes that a part of public service employees, the 'act relevant knights' in his terms, are 'motivated by the need to perform the helping acts themselves'. Other studies on public sector workers' motivations include Edmonds et al. (2002) and Frank and Lewis (2004).

Perhaps the clearest example of street-level bureaucrats with dual tasks (both helping and sanctioning) are caseworkers. There is a rich empirical literature studying the motives and client-orientation of caseworkers. Blau

(1960: 347) studies the attitude towards clients of personnel in a public welfare agency and concludes that 'Most persons who took a job in the welfare agency were partly motivated by an interest in working with and helping poor people. They tended to look forward to establishing a warm, although not intimate, relationship with deserving and grateful clients, and considered the case worker as the agent of society who extended a helping and trusting hand to its unfortunate members.' Marston et al. (2005: 146) provide strong evidence for client-advocacy in a Danish employment project. They cite a bureaucrat as saying that: 'How am I supposed to activate people who are running around in the streets without a home– I can't (...) but I need to give them a temporary place to stay- or do something for them.' Heckman et al. (1996: 2) find that caseworkers in a US job-training program have 'a strong desire to aid the least well off.' Lastly, Considine (2000: 290) finds that Australian caseworkers do not like to sanction clients: 'They found it off-putting to subject job-seekers to the framing of highly legalistic agreements in their first weeks.' They also 'saw sanctioning as a last resort which implied a breakdown in their service and thus a loss of face for them and their agency.'

However, not all street-level bureaucrats and caseworkers have such warm feelings towards clients. Hernandez et al. (2003: 15-16) interviewed participants who used vocational rehabilitation (VR) services: '21% of the participants reported that they had VR counselors who were committed', but '29% of those participants found that VR counselors were unresponsive, particularly when they failed to return telephone calls and follow through with specific tasks that were discussed during appointments (for example, offering but never providing job placement services).' Using Swiss survey data, Behncke et al. (2007: 8-9) also find striking differences between caseworkers' attitudes towards clients. In their sample, 52% of caseworkers state that 'the wishes of the unemployed should be satisfied', while 9% of caseworkers 'assign placements in jobs and active labour market programmes independent of the wishes of the unemployed'. Lastly, Blau (1960: 347) notices that a few of the caseworkers in his study were motivated to seek their job by considerations such as a desire to dominate people.

3 The model

The model revolves around a principal (the benefit administration or public employment service) who hires an endogenous number of agents (casemanagers or caseworkers). The task of an agent is to meet clients (unemployed

workers or people on social benefits) and to allocate to each of them either an employment service, a sanction, or no help at all. Employment services can consist of schooling, job search assistance, assessments etc. A sanction can be a pecuniary penalty, but also a non-pecuniary penalty, for instance workfare where the client is obliged to do production work. For convenience, we normalize the number of clients each agent meets to one.

In what follows we describe the possible allocations and payoffs to the principal, agents, and clients, which are summarized in table 1.

Principal The principal's preferred allocation depends on the client's type. Clients differ in two respects: their motivation and their ability to find a job. For convenience we assume that clients belong to either one of the following four types.[2] The first type of clients, denoted by l, is willing to work, but not able to find a job without help. They need assistance in the form of employment services to improve their labor skills and job search behavior. When a client of type l receives employment services, the principal's payoff increases by $b - c > 0$, where b represents the gains from clients finding a job and leaving welfare with a higher probability, and c stands for the costs of the employment services. Without help the willing but unable clients would not be likely to find a job, leaving the principal a payoff we normalize to zero. Giving them a sanction is considered to be unfair by the principal. The principal receives some kind of payment z: the money collected from the penalty, the production value of the client under workfare, and possible positive effects of sanctions on the probability of leaving welfare and finding a job (see for instance Van den Berg et al., 2004). However, the principal loses x (well-being) from the wrongful treatment of the willing client. The principal's payoff from this allocation is assumed to be negative, $z - x < 0$. The second type of clients, denoted by m, is willing and able to find a job. The best decision for the principal would be *not* to help those clients resulting in a payoff which we normalize to zero. Giving them employment services would entail costs, but does not help them to find a job faster, resulting in a payoff of $-c < 0$ for the principal. A sanction would be considered unfair, implying a payoff of $z - x < 0$. The third type of clients is able to find a job, but not willing to do so. The fourth type of clients is neither willing, nor able to find a job. We label these last two types by their common denominator: the non-willing, n. According to the principal they should all receive a punitive

[2] Our labelling of clients resembles the ones mentioned in e.g. Marston et al. (2005: 149), Sol et al. (2007: 21), and Bunt et al. (2008: 37).

sanction for misbehavior, resulting in a payoff $z > 0$.[3] Allocating them employment services is a waste of money, $-c < 0$.[4] Allocating no help to them leaves the principal a payoff normalized to zero.

The principal knows the distribution of clients' types, but does not know the type of each individual client. He needs agents to sort this out for him and allocate the right service to a given number of clients. The principal has a monopoly in supplying allocations to these clients: clients cannot choose who monitors their job search behavior. The principal pays agents a base salary w, which lowers the principal's payoff. Further, the principal may use incentives which are discussed below. Hiring agents to make allocations is only optimal when the principal's utility U_p from doing this is equal to or higher than the principal's reservation utility, that is allocating all clients the same treatment. We assume throughout that the principal always finds it worthwhile to hire a strictly positive number of agents. The principal hires agents from a pool of heterogenous individuals which is sufficiently large so that the principal is never supply-constrained. Further, the principal is a monopsonist in the labor market for agents. This assumption only plays a role in the subsections where we derive the number of agents the principal wants to hire. We shall also discuss what happens when the principal has no monopsony power (that is, competes with other bureaucrats for workers).

Clients Clients are fully informed about their willingness and ability to find a job. The utility of a client U_c depends on his type and on the allocation made to him. As for the principal, we normalize clients' payoff to zero in case they receive no help. All clients dislike sanctions. These give them a negative payoff $(-v < 0)$. Because willing and unable clients like to have a job and need help to find it, they appreciate employment services. This gives them a positive payoff, $k > 0$. Willing and able clients are indifferent between receiving employment services and no help.[5] Non-willing clients prefer receiving no help and enjoy their leisure time to participating in employment services,

[3] In addition to the monetary payoff of imposing a sanction, there could be some feelings of satisfaction or justification that a non-willing client gets punished. To save on notation, we ignore these potential benefits.

[4] Although this might seem a strong assumption, relaxing it does not change our results much as long as the non-willing clients dislike employment services or as long as a sanction should be allocated as well.

[5] If the willing and able clients strictly prefer no help to participating in employment services, our results do not change. If they strictly prefer participating in employment services to receiving no help, there is an additional incentive problem, but our main arguments remain unaffected. In this case, the clients' and principal's preferences differ in two respects instead of one.

Table 1: Payoff to principal, client and agent of different allocations

Client's type	**Allocation**	**Payoff**		
		Principal	**Client**	**Agent**
Willing, unable l	Sanction	$z-x-w<0$	$-v<0$	$w-\theta_j v$
	No help	$-w<0$	0	w
	Employment services	$b-c-w>0$	$k>0$	$w+\theta_j k$
Willing, able m	Sanction	$z-x-w<0$	$-v<0$	$w-\theta_j v$
	No help	$-w<0$	0	w
	Employment services	$-c-w<0$	0	w
Non-willing n	Sanction	$z-w>0$	$-v<0$	$w-\theta_j v$
	No help	$-w<0$	0	w
	Employment services	$-c-w<0$	$-g<0$	$w-\theta_j g$

which gives them a negative payoff $-g$. Receiving employment services is, however, preferred to getting a sanction, $-v<-g<0$.

There are $L>0$ willing and unable clients, $M>0$ willing and able clients, and $N>0$ non-willing clients. The total number of allocations, denoted by Q, is endogenously determined by the principal (through his decision on the number of agents he wishes to hire), but cannot exceed the total number of clients, $Q \leq L+M+N$.

Agents As soon as an agent meets a client, he knows the client's type. Thus, investigating a client does not involve cost of effort. Hence, an agent is always fully informed about the client's willingness and ability to work when he allocates a service or sanction.[6] Agent's utility depends first of all on his base salary w (see table 1). Second, the agent may be altruistic towards the client the agent meets. This is represented by $\theta_j U_c$, where θ_j measures agent j's altruism towards his client, and U_c is the utility of the client the agent meets. Since agents only have altruistic feelings towards clients they meet, they are impurely altruistic or have 'warm glow' preferences

[6] In this paper we are primarily interested in finding out what happens when the agent knows the correct type and has a certain discretion to allocate a service or sanction. Although the question how much effort an agent exerts to determine the correct allocation is also an interesting one (and studied in depth by Prendergast, 2007), it is beyond the scope of this paper.

in the sense of Andreoni (1989, 1995). We assume that for any j, $0 \leq \theta_j \leq \overline{\theta} < 1$,[7] and that θ_j is distributed according to the cumulative distribution function $\int_0^{\overline{\theta}} f(\theta)\, d\theta$ with mass R. Hence, the total number of potential agents is $R \int_0^{\overline{\theta}} f(\theta)\, d\theta$. Importantly, an agent's altruism θ_j is private knowledge. Altruistic agents take into account how their allocation decisions affect clients' welfare. Without significant loss of generality, we assume that whenever the agent is indifferent between allocations, the agent gives priority to what the client prefers. When the client is also indifferent, the agent decides to allocate what the principal prefers. The agent will only accept the job as a caseworker when his utility is equal to or above his reservation utility, $U_a \geq \bar{A}$.

Incentives As we shall see, the principal's and agents' preferences are not always in line. The principal can reward the agents with a bonus, denoted by π, for making correct allocations, without overruling them when wrong. For instance, we can think of a bonus for job placements when correct decisions lead to maximum job placements. Overruling is not possible, because before observing the outcome, time has past and the allocation has already been put into effect. The bonus for the agents can take a pecuniary or non-pecuniary form. We also show that bonuses for correct decisions and penalties for wrong decisions yield equivalent results.

Timing The principal offers a contract, describing the base salary and bonus. Each agent decides whether or not to take the job. Then each agent meets a client and takes a decision about the allocation. Lastly, the clients', principal's and agents' payoffs are realized.

4 Flat wages

We start by analyzing the case where the principal gives no incentives ($\pi = 0$). He just pays a base salary w.[8] We solve the game by backward induction and start by agents' decisions on allocations.

[7] We also discuss along the way what would happen when agents are allowed to have negative, hostile feelings toward their clients ($\theta_j < 0$).

[8] A practical example of this is discussed by Riccuci and Lurie (2001: 34), who conclude that "there are neither 'carrots' nor 'sticks' to motivate the workers" in the social welfare offices in Texas, Michigan and Georgia. Even though these offices use performance measures, "workers said that all front-line welfare workers are likely to receive the same performance ratings" (Riccuci and Lurie, 2001: 35).

4.1 Which allocations do agents make?

The principal's and agents' preferences align when meeting clients who are willing to work. When meeting a client who is willing but unable, agents allocate employment services, because this gives them a payoff of $w + \theta_j k$, which is higher than the payoff of allocating a sanction, $w - \theta_j v$, or allocating no help, w. When meeting a client who is willing and able, the agents' payoff of allocating no help is w, equal to the payoff of giving employment services and higher than the payoff of sanctioning, $w - \theta_j v$. Whenever the agents are indifferent between allocations, they make the allocations the clients prefer. And if the clients are indifferent too, the agents take the decision the principal prefers. In this case: no help. However, when meeting a client who is non-willing, the principal would prefer the agents allocating sanctions, but the agents allocate no help instead, resulting in a payoff of w. This is higher than when they impose a sanction, $w - \theta_j v$, or give employment services, $w - \theta_j g$, because these allocations hurt non-willing clients. So, the agents' decisions are not fully in line with those desired by the principal: Agents are not willing to sanction the non-willing clients, as they want to avoid the negative feelings they get from imposing sanctions. Instead, they allocate no help to these clients.

4.2 Which agents take the job?

An agent takes the job when his expected utility from taking the job is higher than or equal to his outside option utility, $EU_a \geq \bar{A}$. Using our previous results on agents' allocation decisions, the expected utility for agent j from taking the job is

$$EU_a = w + \frac{L\theta_j k}{L + M + N}. \tag{1}$$

That is, the agent enjoys his base salary and, with probability $\frac{L}{L+M+N}$, helps a willing and able client, which raises his utility by $\theta_j k$. The agent derives no such additional utility when encountering a willing and able client or when encountering a non-willing client, because, as we have seen, the agent will allocate these clients no help. Since $Lk > 0$, the participation constraint can be written as:

$$\theta_j \geq \widetilde{\theta} = \left(\bar{A} - w\right) \left[\frac{Lk}{L + M + N}\right]^{-1}. \tag{2}$$

We can distinguish three cases for the threshold level of agent's altruism $\widetilde{\theta}$. First, $\widetilde{\theta} > \overline{\theta}$. In this case, nobody is willing to take the job. Second, if $\widetilde{\theta} \leq 0$, then the whole labor force is willing to apply. In the third and most

interesting case where $0 < \widetilde{\theta} \leq \overline{\theta}$, only agents with a sufficiently high level of altruism are willing to take the job. We focus on this third case.[9] Notice that this implies that $\bar{A} - w > 0$: the base salary does not make up for foregoing the outside option. The reason is that the job gives agents an opportunity to help willing but unable clients, which increases altruistic agents' utility.

The self-selection process of the agents is affected by the composition of the client population, the employment services and sanction policy, and the principal's personnel policy. We discuss the influence of these aspects in turn.

Client population The higher the number of willing and unable clients, the more attractive the job is for altruistic agents, because there are a lot of clients needing employment services and thus a big chance of getting the warm feelings of helping them. So even for agents with a relatively low level of altruism θ_j, the job becomes interesting:

$$\frac{\partial \widetilde{\theta}}{\partial L} = \frac{-k\left(\bar{A} - w\right)(M+N)}{(Lk)^{2}} < 0.$$

The higher the number of willing and able clients, the less interesting the job becomes. These clients need no help. Thus there are no benefits from helping to compensate for the low-paying job:

$$\frac{\partial \widetilde{\theta}}{\partial M} = \frac{\bar{A} - w}{Lk} > 0.$$

The same holds for the number of non-willing clients. Because the agents do not give them sanctions, they avoid the negative feelings this would evoke for the clients and thus themselves. But they do not get positive feelings either. When there are more of these clients, the principal needs to raise the base salary to attract enough agents:

$$\frac{\partial \widetilde{\theta}}{\partial N} = \frac{\bar{A} - w}{Lk} > 0.$$

Employment services and sanction policy The employment services and sanctions can be more or less attractive to clients. Clients, for example, often like employment services where they themselves can have a say. The

[9] When we would allow for agents with $\theta_j < 0$ (hostile agents), we would get a result similar to Prendergast (2007). That is, agents from both ends of the spectrum, with very positive and very negative attitudes to clients, take the job.

more attractive the employment services for clients, the more interesting the job for agents:

$$\frac{\partial \widetilde{\theta}}{\partial k} = \frac{-L\left(\bar{A} - w\right)(L + M + N)}{(Lk)^{2}} < 0.$$

Making the sanction policy more or less fierce has no effect in this simple case, because agents do not impose sanctions anyway:

$$\frac{\partial \widetilde{\theta}}{\partial v} = 0.$$

Personnel policy The principal in this case has a simple personnel policy: He only offers a base salary, w. Raising this salary makes the job attractive to a larger number of agents:

$$\frac{\partial \widetilde{\theta}}{\partial w} = \frac{-(L + M + N)}{Lk} < 0. \tag{3}$$

4.3 Optimal personnel policy

We have seen that agents are willing to take the job when they are sufficiently altruistic, more precisely: when condition (2) holds. But how many agents does the principal want to hire? Recall that each agent makes one allocation. Using equation (2), the total number of allocations can be written as $Q = R\int_{\widetilde{\theta}}^{\overline{\theta}} f(\theta)\, d\theta$. Further, using our results on agents' allocation choices in section 4.1, the principal's expected payoff of hiring an agent is $\frac{L}{L+M+N}(b-c) - w$. Hiring an agent increases the number of employment services allocated to willing and unable clients, but comes at the cost of paying the base salary. The principal's optimization problem can thus be written as

$$\max_{w}\left[\frac{(b-c)L}{L+M+N} - w\right] R\int_{\widetilde{\theta}}^{\overline{\theta}} f(\theta)\, d\theta,$$

where $\widetilde{\theta}$ is described by equation (2). The first-order condition describing the optimal base salary is:

$$\frac{\partial EU_p}{\partial w} = \left[\frac{(b-c)L}{L+M+N} - w\right] Rf(\widetilde{\theta})\left[\frac{Lk}{L+M+N}\right]^{-1} - R\int_{\widetilde{\theta}}^{\overline{\theta}} f(\theta)d\theta = 0. \tag{4}$$

The condition shows us that by raising the base salary the principal attracts a number of additional agents, $Rf(\widetilde{\theta})\left[\frac{Lk}{L+M+N}\right]^{-1}$. This raises the principal's

expected payoff as these agents allocate employment services to the willing and unable clients, and lowers his payoff by the salary he has to pay them $\left[\frac{(b-c)L}{L+M+N} - w\right]$. Raising the base salary also implies that the principal has to pay higher salaries to all agents he hires, $-R\int_{\widetilde{\theta}}^{\overline{\theta}} f(\theta)\, d\theta$. The reason is that if the principal wants to hire more agents, he must offer a base salary that is attractive to less altruistic agents. In the optimum, the principal equates these marginal benefits and costs or, if the resulting number of potential allocations exceeds the number of clients, hires $L + M + N$ agents. Note that the principal may optimally choose to hire too few agents to serve all clients. This stems from the principal's monopsony power: increasing all agents' base salary may not make up for the payoff resulting from an increase in allocations. Hence, insufficient staffing and an overload of clients may to some extent be an optimal choice. If the principal lacks monopsony power, either all or none of the clients would be served, depending on whether the market wage is lower or higher than the expected payoff of an allocation.

5 Pay-for-performance

As we have seen in the previous section the principal and agents' preferences are not always in line: Agents do not sanction non-willing clients. Can the principal change the behavior of the agents and at what cost? In this section we investigate what happens when the principal rewards agents with a bonus $\pi > 0$ for every correct decision, without overruling them when wrong.[10] An example of such an incentive scheme is making agent's pay dependent of his clients' labor market performances.[11]

5.1 Which allocations do agents make?

The allocations to the willing and unable clients and to the willing and able clients were already in line with the preferences of the principal in the simple

[10] Alternatively, the principal may give a bonus only when an agent imposes a sanction. This incentive scheme, however, induces agents not only to sanction non-willing clients, but also the willing and able clients, which is costly to the principal.

[11] Klerman et al. (2005: 129) observe individual rewards for workers in California, as does Weissman (1997: 37) for Los Angeles County. For example, top performers or workers having more than 10 placements in the preceding month got rewards such as free movie tickets or banners with their name on it. Furthermore, Courty and Marschke (1997) and Heckman et al. (2002) study the use of performance standards for local training centers in the US, where centers were rewarded based on their participants' labor market outcomes.

case studied above. Pay-for-performance does not change the agents' allocations to these clients; it only increases the payoff to the agents of making this allocation by π. However, when the agent meets a non-willing client, bonus pay may induce an agent to sanction rather than allocate no help. His payoff of sanctioning becomes $w + \pi - \theta_j v$, while the payoff of allocating no help remains w and the payoff of allocating employment services remains $w - \theta_j g$. This means that, if the bonus π is high enough, the payoff of sanctioning is higher than the payoff of allocating no help. More specifically, an agent will give non-willing clients a sanction when his level of altruism is lower than the threshold level:

$$\theta_j \leq \widehat{\theta} = \pi / v. \tag{5}$$

When $\theta_j > \widehat{\theta} = \pi/v$, the agent's decision is not affected by the bonus: He allocates no help to the non-willing clients, because the bonus does not compensate for the negative feelings the agent experiences when sanctioning a non-willing client.

5.2 Which agents take the job?

An agent applies for the job when his expected utility from the job is higher than his outside option utility, $EU_a \geq \bar{A}$. We need to distinguish two groups of agents: those that sanction non-willing clients and those that do not.

The agent does not sanction $\left(\theta_j > \widehat{\theta} = \pi/v\right)$ The agent's expected utility, equation (1), changes into:

$$EU_a = w + \frac{(L+M)\,\pi}{L+M+N} + \frac{L\theta_j k}{L+M+N} \geq \bar{A},$$

implying the following participation constraint:

$$\theta_j \geq \widetilde{\theta} = \left(\bar{A} - w - \frac{(L+M)\,\pi}{(L+M+N)}\right)\left[\frac{Lk}{L+M+N}\right]^{-1}. \tag{6}$$

Compared to the case of flat wages, the participation constraint has changed in only one way: The agent earns a bonus π for the allocations to willing clients $(L+M)$. All comparative statics have the same sign as in the absence of bonuses, except for the effect of a higher number of willing and able clients:

$$\frac{\partial \widetilde{\theta}}{\partial M} = \frac{\bar{A} - w - \pi}{Lk} \gtreqless 0.$$

When the bonus π is sufficiently high, a higher number of willing and able clients increases the number of agents willing to take the job. Furthermore, note that an increase in the bonus makes the job more attractive:

$$\frac{\partial \widetilde{\theta}}{\partial \pi} = \frac{-(L+M)}{Lk} < 0.$$

This effect is smaller than that of raising the base salary (which is again given by (3)), because agents who do not sanction only receive the bonus when encountering willing clients. As we shall see, this has important implications for the optimal level of the bonus.

The agent sanctions $\left(\theta_j \leq \widehat{\theta} = \pi/v\right)$ Expected utility of agents who optimally decide to sanction non-willing clients reads:

$$EU_a = w + \pi + \frac{\theta_j (Lk - Nv)}{L+M+N} \geq \bar{A},$$

implying the following participation constraint:

$$\text{if } Lk - Nv > 0, \text{ then } \theta_j \geq \widetilde{\widetilde{\theta}} = \left(\bar{A} - w - \pi\right) \left[\frac{Lk - Nv}{(L+M+N)}\right]^{-1}; \quad (7)$$

$$\text{if } Lk - Nv < 0, \text{ then } \theta_j \leq \widetilde{\widetilde{\theta}} = \left(\bar{A} - w - \pi\right) \left[\frac{Lk - Nv}{(L+M+N)}\right]^{-1}. \quad (8)$$

First, consider the case where $Lk - Nv > 0$. That is, given that an agent sanctions non-willing clients, the job brings higher expected joys of helping the willing and unable clients than expected sorrows of sanctioning the non-willing clients. Then, as before, only agents with a sufficiently high level of altruism are willing to apply for the job, $\widetilde{\widetilde{\theta}} \leq \theta_j \leq \overline{\theta}$. Compared to agents who do not sanction, pecuniary payoffs are higher, because the agent gets a bonus for every allocation he makes, while non-pecuniary payoffs are lower, because the agent suffers a loss when sanctioning non-willing clients.

Next, consider the case where $Lk - Nv < 0$. It is easy to see that in this case, if the expected pecuniary payoffs are smaller than the outside option utility, $\bar{A} - w - \pi > 0$, only agents with hostile feelings $\theta_j < \widetilde{\widetilde{\theta}} < 0$, would be willing to apply. However, we have assumed $\theta_j \geq 0$ for any j, and so if the principal wants to attract agents who choose to sanction non-willing clients, he must offer $w + \pi \geq \bar{A}$. The first agents interested in a job like this are the ones with low levels of altruism, agents who do not care too much about the clients' feelings, $0 \leq \theta_j \leq \widetilde{\widetilde{\theta}}$.

The comparative static results are similar to those derived above with one exception: Making the sanction policy more fierce (raising v) makes agents less willing to apply for the job, because they do impose a sanction when encountering non-willing clients:

$$\frac{\partial \widetilde{\widetilde{\theta}}}{\partial v} = \frac{N\left(\bar{A} - w - \pi\right)(L + M + N)}{(Lk - Nv)^2} > 0.$$

Furthermore, raising the bonus has the same effect as raising the base salary, because these agents receive bonuses for all allocations, implying that base salary and bonus pay are perfect substitutes:

$$\begin{aligned}
\text{if } Lk - Nv &> 0, \text{ then } \frac{\partial \widetilde{\widetilde{\theta}}}{\partial w} = \frac{\partial \widetilde{\widetilde{\theta}}}{\partial \pi} = \frac{-(L + M + N)}{Lk - Nv} < 0; \\
\text{if } Lk - Nv &< 0, \text{ then } \frac{\partial \widetilde{\widetilde{\theta}}}{\partial w} = \frac{\partial \widetilde{\widetilde{\theta}}}{\partial \pi} = \frac{-(L + M + N)}{Lk - Nv} > 0.
\end{aligned}$$

Although the signs are opposite in the two cases, the interpretation is still the same: Raising the bonus or base salary attracts more agents, in the first case from the top and in the second case from the bottom of the distribution of altruistic agents.

5.3 Optimal personnel policy

By setting the base salary w and the bonus π, the principal determines the number of agents that will be hired as well as affects their allocation decisions. There are two cases that need to be distinguished: The case where the expected nonpecuniary payoffs of the job for agents willing to sanction are positive, $Lk - Nv > 0$, and the case where these are negative.

Nonpecuniary payoffs positive when agents sanction ($Lk - Nv > 0$) In this case, the job is mainly a job of helping needy people getting a better chance on the labor market, even for agents who sanction non-willing clients. Hence, as we have seen, the job is particularly attractive to altruistic agents. Using our results about the allocations agents make (section 5.1) and which agents take the job (section 5.2), we know that, if the principal decides to induce at least part of the agents to sanction non-willing clients, $R\int_{\widehat{\theta}}^{\overline{\theta}} f(\theta)d\theta$ allocations will be made by agents willing to take the job but not willing to sanction, and $R\int_{\widetilde{\widetilde{\theta}}}^{\widehat{\theta}} f(\theta)d\theta$ allocations will be made by agents

willing to take the job and sanction non-willing clients. The resulting expected payoffs to the principal of these two groups of agents are respectively $\left[\frac{(b-c)L}{L+M+N} - w - \frac{(L+M)\pi}{L+M+N}\right]$ and $\left[\frac{(b-c)L+zN}{L+M+N} - w - \pi\right]$ for each allocation made by them. The principal's optimization problem can thus be written as:

$$\max_{w,\pi} \left[\frac{(b-c)\,L}{L+M+N} - w - \frac{(L+M)\,\pi}{L+M+N}\right] R \int_{\widehat{\theta}}^{\overline{\theta}} f(\theta)\, d\theta + \left[\frac{(b-c)\,L+zN}{L+M+N} - w - \pi\right] R \int_{\widetilde{\widetilde{\theta}}}^{\widehat{\theta}} f(\theta)\, d\theta \tag{9}$$

where $\widehat{\theta}$ is described by (5) and $\widetilde{\widetilde{\theta}}$ by (7). It is easy to verify that the participation constraint of agents who do not sanction, described by (6), is not binding, unless $\widehat{\theta} < \widetilde{\theta}$. In the latter case, the bonus is too low to induce any agents to sanction non-willing clients and, hence, the optimization problem is the same as in the case of flat wages. If $\widehat{\theta} \geq \widetilde{\theta}$, the first-order conditions for the optimal base salary and bonus are:

$$\frac{\partial EU_p}{\partial w} = \left[\frac{(b-c)\,L+zN}{L+M+N} - w - \pi\right] Rf(\widetilde{\widetilde{\theta}}) \left[\frac{Lk-Nv}{L+M+N}\right]^{-1} - R\int_{\widetilde{\widetilde{\theta}}}^{\overline{\theta}} f(\theta)d\theta = 0 \tag{10}$$

$$\frac{\partial EU_p}{\partial \pi} = \frac{N(z-\pi)}{L+M+N} Rf(\widehat{\theta})\frac{1}{v} + \left[\frac{(b-c)L+zN}{L+M+N} - w - \pi\right] Rf(\widetilde{\widetilde{\theta}}) \left[\frac{Lk-Nv}{L+M+N}\right]^{-1} - \frac{L+M}{L+M+N} R\int_{\widehat{\theta}}^{\overline{\theta}} f(\theta)d\theta - R\int_{\widetilde{\widetilde{\theta}}}^{\widehat{\theta}} f(\theta)d\theta = 0 \tag{11}$$

Raising the base salary w has similar effects as in the case of flat wages, except that now the additional agents sanction non-willing clients. The principal's payoff increases by the difference between the expected payoffs of these agents helping willing and unable clients and sanctioning non-willing clients and the costs of their salary and bonus, as described by the first term in (10). Furthermore, as before, the principal has to pay a higher salary to all agents, which is reflected by the last term of (10). Due to this monopsony effect, the principal may again optimally choose to hire too few agents to serve all clients.

Raising the bonus π has three effects. Firstly, it induces $\frac{1}{v}Rf(\widehat{\theta})$ additional agents to sanction and receive a bonus rather than allocate no help to

non-willing clients. This results in a total increase of the principal's payoff described by the first term in (11): The principal gains z from each additional sanction to non-willing clients at the cost of paying an additional bonus π. Secondly, as for the base salary, by raising the bonus the principal attracts additional agents willing to sanction, which is reflected by the second term in (11). Notice that this term is identical to the first term in (10), which reiterates our result above that raising the base salary or the bonus have the same effect on recruitment of agents willing to sanction. Lastly, the marginal costs of raising the bonus are described by the last two terms of (11): Agents are compensated better for correct decisions.

In the optimum the principal equates the marginal benefits and costs of raising the bonus and of raising the base salary. Combining the first-order conditions gives:

$$(\pi - z)\,\frac{f(\widehat{\theta})}{v} = \int_{\widehat{\theta}}^{\overline{\theta}} f(\theta)d\theta,$$

which implies that the optimal bonus π exceeds the value to the principal of sanctioning a non-willing client (z). The intuition follows. By raising the bonus, some additional agents are induced to sanction, which raises the principal's payoff by $(\pi - z)$, as discussed above. Hence, it is optimal for the principal to raise the bonus at least to the point where the bonus equals the value to the principal of sanctioning non-willing clients. This echoes the familiar results that, with risk-neutral agents, optimal bonus pay equals the full marginal product. However, there is an additional benefit of raising the bonus. Recall that an increase in the bonus enables the principal to reduce the base salary by the same amount without losing any agents, because the bonus and base salary are perfect substitutes for the marginal agents. Expected wage compensation for agents who do not sanction decreases, however. They bear the full loss of the reduction in the base salary, but gain only partly from the increase in the bonus as they do not sanction. Raising the bonus thus enables the principal to extract rents from the agents who do not sanction non-willing clients. In the optimum, the bonus therefore exceeds the principal's value of sanctioning non-willing clients. Nevertheless, if $\overline{\theta}$ is sufficiently high, the optimal bonus does not induce all agents to sanction non-willing clients. The principal simply finds it too costly to induce highly altruistic agents to impose sanctions. When we have many identical principals competing for the same workers instead of a monopsonist, competition leads to an optimal bonus π equal to the value to the principal of sanctioning a non-willing client (z).

Nonpecuniary payoffs negative when agents sanction ($Lk - Nv < 0$) In this case, the job is mainly about disciplining instead of helping clients. As we have seen in section 5.2, in order to induce at least some of the agents to sanction non-willing clients, the total pecuniary payoffs must be at least equal to the outside option utility, $w + \pi \geq \bar{A}$, implying that some agents from the bottom of the altruism distribution sort into the agency. Using our previous results about which allocations agents make (section 5.1) and who will take the job (section 5.2), we know that $R\int_0^{\widetilde{\widetilde{\theta}}} f(\theta)\, d\theta$ allocations are made by agents willing to sanction and $R\int_{\widetilde{\theta}}^{\overline{\theta}} f(\theta)\, d\theta$ allocations are made by agents not willing to sanction.[12] The principal's optimization problem is:

$$\max_{w,\pi} \left[\frac{(b-c)L + zN}{L+M+N} - w - \pi\right] R\int_0^{\widetilde{\widetilde{\theta}}} f(\theta)\, d\theta + \left[\frac{(b-c)L}{L+M+N} - w - \frac{\pi(L+M)}{L+M+N}\right] R\int_{\widetilde{\theta}}^{\overline{\theta}} f(\theta)\, d\theta, \tag{12}$$

where $\widetilde{\theta}$ is described by (6) and $\widetilde{\widetilde{\theta}}$ by (8). The first-order conditions are:

$$\frac{\partial EU_p}{\partial w} = \left[\frac{(b-c)L + zN}{L+M+N} - w - \pi\right] Rf\left(\widetilde{\widetilde{\theta}}\right)\left[-\frac{Lk - Nv}{L+M+N}\right]^{-1} + \left[\frac{(b-c)L}{L+M+N} - w - \frac{\pi(L+M)}{L+M+N}\right] Rf\left(\widetilde{\theta}\right)\left[\frac{Lk}{L+M+N}\right]^{-1} - R\int_0^{\widetilde{\widetilde{\theta}}} f(\theta)\, d\theta - R\int_{\widetilde{\theta}}^{\overline{\theta}} f(\theta)\, d\theta = 0 \tag{13}$$

$$\frac{\partial EU_p}{\partial \pi} = \left[\frac{(b-c)L + zN}{L+M+N} - w - \pi\right] Rf(\widetilde{\widetilde{\theta}})\left[-\frac{Lk - Nv}{L+M+N}\right]^{-1} + \left[\frac{(b-c)L}{L+M+N} - w - \frac{\pi(L+M)}{L+M+N}\right] Rf(\widetilde{\theta})\left[\frac{L+M}{L+M+N}\right]\left[\frac{Lk}{L+M+N}\right]^{-1} - R\int_0^{\widetilde{\widetilde{\theta}}} f(\theta)d\theta - \frac{L+M}{L+M+N} R\int_{\widetilde{\theta}}^{\overline{\theta}} f(\theta)d\theta = 0 \tag{14}$$

[12] Using (5), (6) and (8), it is easy to verify that $\widetilde{\widetilde{\theta}} < \widehat{\theta} < \widetilde{\theta}$ in all cases where the principal hires neither all nor none of the potential agents. In other words, all agents hired from the top of the altruism distribution do not sanction, while all agents hired from the bottom do sanction non-willing clients. Hence, the incentive constraint (5) is redundant.

The first term in (13) shows that, as before, by raising the base salary w, the number of agents willing to sanction increases. However, in this case, it increases the number of agents not willing to sanction as well, resulting in additional payoffs described by the second term of (13). Furthermore, the principal has to pay a higher salary to all agents, denoted by the third and fourth term in (13), which can lead to the same monopsony result for the optimal number of allocations as before: too few agents to serve all clients. The first two terms in (14) show that by raising the bonus, the number of agents willing to sanction as well as the number of agents not willing to sanction increase. Finally, the last two terms of (14) describe the increase in bonus paid to all hired agents. Combining the first-order conditions and rewriting gives:

$$\left[\frac{(b-c)\,L}{L+M+N}-w-\frac{\pi\,(L+M)}{L+M+N}\right]f\left(\widetilde{\theta}\right)\left[\frac{Lk}{L+M+N}\right]^{-1} \tag{15}$$
$$=\int_{\widetilde{\theta}}^{\overline{\theta}} f\,(\theta)\,d\theta.$$

$$\left[\frac{(b-c)\,L+zN}{L+M+N}-w-\pi\right]f\left(\widetilde{\widetilde{\theta}}\right)\left[-\frac{Lk-Nv}{L+M+N}\right]^{-1} \tag{16}$$
$$=\int_{0}^{\widetilde{\widetilde{\theta}}} f\,(\theta)\,d\theta,$$

Hence, the principal hires agents not willing to sanction until the expected payoffs of hiring additional agents are equal to the increase in rents he has to leave to these agents. The principal can hire more of those agents by increasing their pecuniary payoffs through a raise in base salary, which is not fully compensated by a lower bonus. This increase in pecuniary payoffs to agents not willing to sanction does not affect the number of agents willing to sanction, because the principal can compensate the raise in their base salary fully by lowering the bonus, as these are perfect substitutes for them.

From (16) we see that the principal hires agents willing to sanction until the expected payoff of hiring another agent willing to sanction is equal to the rents the principal has to pay the rest of the sanctioning agents. The principal attracts more agents willing to sanction by raising the pecuniary payoffs to these sanctioning agents by increasing the bonus more than deceasing the base salary. Because agents not-willing to sanction receive bonuses over less

allocations, the principal can compensate their increase in bonuses by lowering the base salary somewhat, so that their willingness to take the job is not affected.

In the optimum the principal equates the marginal benefits and costs of raising the base salary and bonus and thus determines the optimal number of agents willing to sanction and optimal number of agents not willing to sanction. This may imply hiring agents from both ends of the distribution of altruistic types: highly altruistic agents who do not sanction non-willing clients and earn low income and low-altruistic agents who sanction and earn high bonus pay. When the value to the principal of sanctioning non-willing clients is sufficiently high, it may even be optimal to hire low-altruistic agents willing to sanction only. However, the opposite may also be the case. When the value of sanctioning is sufficiently low, the principal does not find it worthwhile to offer high pecuniary payoffs so as to attract agents who are willing to sanction. This results in sorting of the most altruistic types only, as in the case of flat wages.

5.4 Punishment instead of a bonus

What happens when the principal, instead of a bonus π for correct allocations, gives the agent a penalty $-\pi$ for every wrong allocation? The incentive constraint remains the same. As soon as $\theta_j \leq \widehat{\theta} = \pi/v$, the agent will sanction non-willing clients. The participation constraints do change somewhat. If the agent is not willing to sanction, the participation constraint in condition (6) becomes:

$$\theta \geq \widetilde{\theta} = \left[\bar{A} - w + \frac{N\pi}{L+M+N}\right]\left[\frac{Lk}{L+M+N}\right]^{-1}.$$

While in the case of a bonus for correct allocations the agent receives an expected bonus of $\frac{(L+M)\pi}{L+M+N}$, he now has to pay an expected punishment of $\frac{N\pi}{L+M+N}$. To attract the same number of agents the principal has to raise the base salary w by the difference between the expected bonus and penalty.

The same holds for the participation constraint of agents who are willing to sanction. Condition (7) becomes:

$$\theta \geq \widetilde{\widetilde{\theta}} = \left[\bar{A} - w\right]\left[\frac{Lk - Nv}{L+M+N}\right].$$

So, in comparison to the situation where the agent receives a bonus for all his correct allocations, the agent now receives nothing. The nonpecuniary payoffs

are the same. This implies that the principal has to raise the base salary w by the forgone bonus π to attract the same number of agents as in the case of a bonus for correct allocations.

When nonpecuniary payoffs are positive, the participation constraint of agents willing to sanction is binding. This implies that the principal can extract rents from agents not willing to sanction by raising the penalty for wrong allocations, while leaving the base salary unchanged. The participation constraint of agents willing to sanction remains unaffected. Thus, the results are the same as with bonuses for correct allocations.

6 Concluding remarks

In this paper we studied the use of pay-for-performance to align the agent's and principal's preferences in street-level bureaucracies where agents have the dual task of helping some clients while disciplining others. Our theoretical work has some clear predictions which can be tested in future empirical research using datasets as for example those used by Behncke et al. (2007, 2008) combined with data on agent's altruism and incentives. This would allow us to test whether there is indeed a relation between agent's level of altruism and the incentives they receive (or the lack thereof) and whether the composition of the client population is indeed related to agents' willingness to take the job and the allocations they make. Also, it would be interesting to study whether huge differences in caseloads, as for example observed in Bloom et al. (2003), are related to differences in payoffs of hiring additional workers, as measured by the degree of competition between agencies and the composition of clients.

Of course, there are several other ways to align preferences than pay-for-performance. We discuss three of these alternatives here. A first alternative is to monitor (a part of) the agent's allocations before they are put into effect and punishing and overruling the agent when a wrong allocation is detected (see e.g. Van der Veen 1990). In a first draft we explored this case. If the agent does not care about the effects on the client of overruling by the principal, the results are similar to the case of pay-for-performance studied above.[13] But the results change when we assume that the agent, once he has met the client, cares to some extent about the effects on the client of overruling.[14] Like bonus pay, monitoring induces part of the bureaucrats with lower levels

[13] Thus, we follow Andreoni's (1989, 1995) definition of "warm glow" feelings: the agent only experiences these from his own deeds.

[14] The agent is to some extent purely altruistic, but only to clients he has met.

of altruism to sanction non-willing clients. But when the monitoring rate is sufficiently high, there can be further consequences for sorting: Sorting into the job from the bottom of the altruism distribution only, even from agents not willing to sanction. The reason is that, with a high monitoring rate, the agency is likely to overrule the bureaucrat when observing that the bureaucrat has not imposed a sanction on non-willing clients. When the bureaucrat's sorrows of the principal's imposition of sanctions on a part of his clients are larger than the joys of helping others, the job is no longer attractive to altruistic people and the agency needs to offer a relatively high wage to attract people who will all be little concerned about clients' welfare.

A second alternative to align the principal and agent's preferences is to use a statistical assignment program, profiling. This can for example help casemanagers asses clients' needs and make the correct allocations. Many countries use such a procedure to tailor services to those who are likely to need them most (OECD, 2007, Black et al., 2003, Rosholm et al., 2004, Ebert and O' Leary, 1996, Bell and Orr, 2002). Another reason to introduce profiling is to avoid casemanagers' bias. Or, as Bell and Orr (2002: 281) put it, to promote that: 'identical persons will get the same treatment, regardless of who their caseworker might be'. The use of profiling constrains the agent's room for discretion if there are consequences for the agent when he diverts from the advised allocation of the profiling system too often. This can be considered as ex-ante monitoring and is thus likely to imply the same results as described in the previous paragraph on monitoring.

A third alternative to align preferences is using clients' complaints to inform the principal that a mistake has been made. However, complaints would be uninformative in our model, because agents do not sanction willing clients. The only clients who would complain about sanctions are (a part of the) non-willing clients, who are sanctioned deservedly. The other allocations, employment services to willing and unable clients and no help to either willing and able and (the rest of the) non-willing clients are in line with those clients' preferences. No complaints are to be expected from those allocations. Complaints, however, are informative in models where agents need to exert effort to make a correct allocation, as studied by Prendergast (2007).

We refrained from any private costs to agents of sanctioning clients, like for example physical threats, lots of paperwork, or unwillingness to become engaged in the legal process of sanctions and appeal (see e.g. Considine 2000). This allowed us to focus on the consequences of giving incentives to altruistic agents. If there were private costs of sanctioning, agents would be even less likely to sanction clients. To make an agent sanction, the bonus needs to compensate for these private costs, thus decreasing the principal's payoffs of

agents sanctioning non-willing clients.

We assumed that agents have different levels of altruism towards clients, but each agent's level of altruism is the same towards all his clients. Thus, we assumed that the agent does not discriminate against some clients. But as Lipsky (1980: 108) observes: "some clients simply evoke workers' sympathy or hostility (...) workers may be inclined to 'give the underdog a break' or may favour clients with similar ethnic backgrounds, as when racial or ethnic favoritism prevails in discriminatory decision making." These feelings can of course influence agents' allocation decisions as well. Feelings of reciprocity (Fong et al., 2004; Fong, 2004; Fehr and Gächter, 2000) can also play a role. People tend to treat friendly and deserving people better than hostile and undeserving people. Reciprocity would imply that agents respond to some extent differently to willing-clients than to non-willing clients. If we allow agents in our model to be hostile towards non-willing clients and altruistic towards willing-clients, they would allocate more sanctions to non-willing clients. Thus, this would improve the alignment between the principal's preferences and the agent's allocations. However, these reciprocal feelings are in practice not likely to be strong enough to achieve perfect alignment, which is also clear from the empirical literature discussed in section 2.

A last note on agent's altruism is that it does not have to be stable over time. As Blau (1960: 347, 348) notices "the attitudes of most new case workers toward clients were strongly positive, if somewhat sentimental and idealistic (...) But as he encountered clients who blamed him personally for not helping them enough (...) and clients met his trusting attitude by cheating and lying, the newcomer tended to experience a 'reality shock' (...) This disillusioning experience might make a worker bitter and callous, or induce him to leave the job, and even those who did not have either of these extreme reactions tended to change their orientation to clients." This is an interesting subject for future work.

Finally, we have restricted the agents' tasks to making a decision about an allocation. However, agents sometimes perform several other tasks as well. For instance, make allocations as well as execute them. Agents may even have to decide on whether to execute these services themselves or to delegate them to other agents, public or private. Which decisions do agents make in this case? And who will take the job? These are also interesting questions for future research.

A Notation

- U_a =agent's utility function
- $\bar{A}$ =agent's outside option utility
- U_p =principal's utility function
- w =agent's base salary
- θ =agent's level of altruism towards clients
- z =principal's payoff when non-willing clients receive a sanction
- $z - x$ =principal's payoff when willing clients receive a sanction
- $-c$ =principal's costs when non-willing or willing and able clients receive employment services
- b =principal's payoff when willing and unable clients receive employment services
- L =number of clients of type l , willing and unable.
- M =number of clients of type m, willing and able.
- N =number of clients of type n, non-willing.
- Q =number of allocations
- U_c =client's utility function
- $-v$ =clients' costs of a sanction
- $-g$ = non-willing clients' costs of employment services
- k = willing and unable clients' benefits of employment services
- R =mass of the probability distribution function of agents
- π =bonus for correct allocations
- $-\pi$ =penalty for wrong allocations

References

[1] Andreoni, James (1989), Giving with Impure Altruism: Applications to Charity and Ricardian Equivalence, *Journal of Political Economy*, 97(6): 1447-1458.

[2] Andreoni, James (1995), Warm-Glow Versus Cold-Prickle: The Effects of Positive and Negative Framing on Cooperation in Experiments, *Quarterly Journal of Economics*, 110(1): 1-21.

[3] Behncke Stefanie, Markus Frölich, Michael Lechner (2007), Unemployed and Their Caseworkers: Should They Be Friends or Foes?, IZA Discussion Paper No. 3149, Institute for the Study of Labor.

[4] Behncke Stefanie, Markus Frölich, Michael Lechner (2008), A Caseworker Like Me: Does the Similarity between Unemployed and Caseworker Increase Job Placements?, IZA Discussion Paper No. 3437, Institute for the Study of Labor.

[5] Bell, Stephen H. and Larry L. Orr (2002), Screening (and Creaming?) Applicants to Job Training Programs: the AFDC Homemaker–Home Health Aide Demonstrations, *Labour Economics*, 9(2): 279-301.

[6] Berg, Gerard J. Van den, Bas Van der Klaauw and Jan C. Van Ours (2004), Punitive Sanctions and the Transition Rate from Welfare to Work, *Journal of Labor Economics*, 22(1): 211-241.

[7] Besley, Timothy and Maitreesh Ghatak (2005). Competition and Incentives with Motivated Agents, *American Economic Review*, 95(3): 616–636.

[8] Black, Dan A., Jeffrey A. Smith, Mark C. Berger and Brett J. Noel (2003), Is the Threat of Reemployment Services more Effective than the Services Themselves? Evidence from a Random Assignment in the UI System, *American Economic Review*, 93(4): 1313-1327.

[9] Blau, Peter (1960), Orientation toward Clients in a Public Welfare Agency, *Administrative Science Quaterly*, 5(3): 341-361.

[10] Bloom, Howard,Carolyn J. Hill and James A. Riccio (2003), Linking Program Implementation and Effectiveness: Lessons from a Pooled Sample of Welfare-to-Work Experiments, *Journal of Policy Analysis and Management*, 22(4): 551-575.

[11] Bunt S., M. Grootscholte, D.R. Kemper, C van der Werf (2008), *Work First and the Prospects on the Labour Market: Research into the Effects of Work First in the Netherlands*, The Hague: Council for Work and Income.

[12] Considine, Mark (2000), Selling the Unemployed: The Performance of Bureaucracies, Firms and Non-Profits in the New-Australian "Market" for Unemployment Assistance, *Social Policy & Administration*, 34(3): 274-295.

[13] Courty, Pascal and Gerald Marschke (1997), Measuring Government Performance: Lessons from a Federal Job-Training Program, *American Economic Review*, 87(2): 383-388.

[14] Delfgaauw, Josse (2008), Dedicated Doctors: Public and Private Provision of Health Care with Altruistic Physicians, Tinbergen Institute Discussion Paper TI 2007-010/1.

[15] Delfgaauw, Josse, Robert Dur (2008), Incentives and Workers' Motivation in the Public Sector, *Economic Journal*, 118 (525): 171–191.

[16] Dixit, Avinash K. (2002), Incentives and Organizations in the Public Sector: An Interpretative Review, *Journal of Human Resources*, 37(4): 696-727.

[17] Eberts, Randall W. and Christopher O' Leary (1996), Profiling Unemployment Insurance Beneficiaries, W.E. Upjohninstitute for Employment Research.

[18] Edmonds, S., Sharp, C., and Benefield, P. (2002). Recruitment to and Retention on Initial Teacher Training – A Systematic Review, Slough: National Foundation for Educational Research.

[19] Fehr, Ernst and Simon Gächter (2000), Fairness, and Retaliation: The Economics of Reciprocity, *Journal of Economic Perspectives*, 14(3): 159-181.

[20] Fong, Christina, Samuel Bowles and Herbert Gintis (2006), Strong Reciprocity and the Welfare State, In: Serge-Christophe Kolm and Jean Mercier Ythier, eds: *Handbook on the Economics of Giving, Reciprocity, and Altruism*, North-Holland: Elsevier.

[21] Fong, Christina (2004), Which beliefs matter for redistributive politics? Target-Specific versus General Beliefs about the Causes of Income, Mimeo, Carnegie Mellon University, Department of Social and Decision Sciences.

[22] Francois, Patrick (2000), Public Service Motivation as an Argument for Government Provision, *Journal of Public Economics*, 78(3): 275–299.

[23] Francois, Patrick (2007), Making a Difference, *RAND Journal of Economics*, 38(3): 714-732.

[24] Frank, S. A. and G. B.Lewis (2004), Government Employees: Working Hard or Hardly Working?, *American Review of Public Administration*, 34(1): 36–51.

[25] Glazer, Amihai. (2004), Motivating Devoted Workers, *International Journal of Industrial Organization*, 22(3), 427–440.

[26] Heckman James J. , Carolyn Heinrich and Jeffrey A. Smith (2002), The Performance of Performance Standards, *Journal of Human Resources*, 37(4): 778-811.

[27] Heckman, James, Jeffrey A. Smith, Christopher Taber (1996), What Do Bureaucrats Do? The Effects of Performance Standards and Bureaucratic Preferences on Acceptance into the JTPA Program, NBER working paper 5535.

[28] Hernandez B., J. Rosen, J. Cometa, K. Bailey and R.D. Luna (2003), Employment, Vocational Rehabilitation, and the Ticket to Work Program: Perspectives of Latinos with Disabilities, Submitted by DePaul University and Access Living of Chicago to the Disability Research Institute, Chicago.

[29] Klerman Jacob Alex , Gail L. Zellman, Tammi Chun, Nicole Humphrey, Elaine Reardon, Donna O. Farley, Patricia A. Ebener and Paul Steinberg (2005), Welfare Reform in California: State and County Implementation of CalWORKs in the Second Year, Rand Research.

[30] Le Grand, Julian (2003), *Motivation, Agency, and Public Policy; Of Knights and Knaves, Pawns and Queens*, Oxford: Oxford University Press.

[31] Lipsky, Michael (1980), Street-Level Bureaucracy: The Dilemmas of Individuals in the Public Service, New York: Sage.

[32] Marston Greg, Joergen Elm Larsen, Catherine McDonald (2005), The Active Subjects of Welfare Reform: A Street-Level Comparison of Employment Services in Australia and Denmark, *Social Work & Society*, 3(2): 141-154.

[33] OECD (2007), *OECD Employment Outlook 2007,* Paris: OECD.

[34] Prendergast, Canice (2007), The Motivation and Bias of Bureaucrats, *American Economic Review*, 97(1): 180–96.

[35] Riccuci, Norma M. and Irene Lurie (2001), Employee Performance Evaluation in Social Welfare Offices, *Review of Public Personnel Administration*, 21(4): 348-350.

[36] Rosholm, Michael, Michael Svarer and Bo Hammer (2004), A Danish profiling system, IZA Discussion Paper No. 1418, Institute for the Study of Labor.

[37] Sol, E., J. Castonguay, H. van Lindert, Y. van Amstel (2007), Work first werkt: Op weg naar evidence based-work first, Utrecht: Divosa.

[38] Veen, Romke Jan van der (1990), *De Sociale Grenzen van Beleid: Een Onderzoek naar de Uitvoering en Effecten van het Stelsel van Sociale Zekerheid* (Thesis), Leiden/Antwerpen: Stenfert Kroese.

[39] Weissman, Evan (1997), Changing to a Work First Strategy: Lessons from Los Angeles County's GAIN Program for Welfare Recipients, New York: Manpower Demonstration Research Company.