THE ROLE OF STIGMA IN THE DESIGN OF WELFARE PROGRAMS

TOMER BLUMKIN, YORAM MARGALIOTH AND EFRAIM SADKA

10-08

OCTOBER, 2008

THE FOERDER INSTITUTE FOR ECONOMIC RESEARCH
AND
THE SACKLER INSTITUTE OF ECONOMIC STUDIES

The Role of Stigma in the Design of Welfare Programs

Tomer Blumkin* **Yoram Margalioth**** **Efraim Sadka*****

May 5, 2008

Abstract

We consider the notion of welfare stigma à la Besley and Coate (1992b). This stigma is attributed to welfare claimants by society when they are perceived as undeserving in the sense that they falsely claim to be eligible for welfare benefits. However, due to imperfect information, this stigma may be extended, with some probability, to all welfare claimants. We examine the implications of this kind of stigma for the design of welfare programs.

JEL Classification: H2, D6

Key Words: Welfare, Take Up, Targeting, Inequality, Stigma.

* Department of Economics, Ben-Gurion University, Beer-Sheba 84105, Israel; CesIfo; and IZA. E-mail address: tomerblu@bgumail.bgu.ac.il.

** The Buchman Faculty of Law, Tel-Aviv University, Tel-Aviv, Israel. E-mail address: margalio@post.tau.ac.il.

*** The Eitan Berglas School of Economics, Tel-Aviv University, Tel-Aviv 69978, Israel; CesIfo; and IZA. E-mail address: sadka@post.tau.ac.il.

1. Introduction

The goal of most welfare programs is to provide benefits to the least well-off individuals. These individuals are usually characterized by low earning ability or ill health – information that cannot be directly observed by the government, nor can it be easily verified.[1] To overcome these problems, the government uses various screening devices. Direct devices include means-testing by reviewing documentation, conducting interviews, and testing by specialists. Indirect screening includes: *targeting groups* (*tagging* according to Akerlof, 1978): basing welfare eligibility on observable characteristics such as old age, education or observable disability, correlated with ability; *targeting benefits*: offering in-kind benefits (e.g., wheelchairs) that deserving individuals; namely, the intended beneficiaries of the program, would find relatively more attractive (Nichols and Zeckhauser, 1982); and *welfare ordeals*: adding requirements, such as work or training requirements to the program (even if they entail pure deadweight costs) that undeserving individuals would find relatively costly and hence, self-select out of the program (Nichols and Zeckhauser, 1982; Besley and Coate, 1992a, 1995).

In designing such screening devices, the government faces a fundamental trade-off between benefits take-up; that is, ensuring that eligible individuals actually receive the benefits;[2] and targeting efficiency; namely, excluding non-eligible claimants.[3]

---

[1] Mental problems and back pains are the two leading causes of inability to work, stated in claims for government income support. Unfortunately, these causes are especially difficult to verify (King, 2004; see also Campolieti, 2002).

[2] Ample empirical research has shown that complex and time-consuming application processes (e.g., Moffitt, 2003; Currie, 2006), as well as stigma (Moffitt, (1983), significantly reduce welfare take-up rates. Lack of information about the existence of some programs and their terms or eligibility rules is another often-mentioned explanation for low take-up rates.

[3] A well-known example is the 'Dutch disease' of the 90s. In 1990, the proportion of Dutch workers aged 15 to 44 living on disability insurance was about three times higher than in other European countries. In that year, Prime Minister Lubbers called the Netherlands a "sick" country, and vowed to

Stigma is a major cause for low take-up (see, *inter alia*, Moffitt, 1983; Hancock et al., 2004; Pudney et al., 2006), and reducing welfare stigma is a commonly stated policy objective (e.g., Besley and Coate, 1992b, 1995). A recent example of a policy change in this direction was the introduction of Electronic Benefit Transfer (EBT) systems to provide food stamp program participants with a magnetic debit card that looks more like a regular debit or credit card.

In this paper, we draw a distinction between two different explanations for the phenomenon of welfare stigma. Stigma is a Greek word that refers to a kind of tattoo that was cut or burned into the skin of criminals, slaves, or traitors, in order to identify them visibly as blemished or morally polluted persons (Frey, 2003). Sociological theory defines stigma as a phenomenon whereby an individual with an attribute that is strongly discredited by his/her society is rejected because of that attribute (Goffman, 1963). The first well-known explanation for welfare stigma is that welfare recipients are publicly exposed as being unable to support themselves. The relevant attribute is therefore inability, which is an accurate description of the targeted welfare recipient from the government's perspective. If this indeed is the cause of stigma, it means that an individual who is unable to provide for him/herself is reduced in our minds from a whole and normal person to a tainted, discounted one. Explained in this way, welfare stigma is an unfortunate cost that should be minimized by making the process of application and collection of welfare benefits as discreet as possible or by providing the benefits universally.

A second possible explanation for welfare stigma is that the relevant attribute is not genuine inability but rather laziness or dishonesty. According to this explanation, society in general is sympathetic towards the unfortunate disabled and

---

"cure" it. The Netherlands was indeed "cured" in 1993 when greater scrutiny was introduced (Aarts and De Jong, 1996).

wishes to help them through welfare programs and otherwise (Wolf, 2004). However, in practice, there may also be people who pretend to be deserving and thereby manage to receive unwarranted welfare benefits. Because of this phenomenon, welfare recipients are viewed with some suspicion. According to this explanation, the cause of welfare stigma is that the welfare recipients "are held responsible for their condition" (Pettigrew, 1980). This is especially true in the US. Most Americans – elites, middle class, and even the poor – believe that effort is what principally accounts for how people do in life, and that those who are poor simply have not tried hard enough (Rainwater, 1974). A systematic investigation of American conceptions of the poor (Feagin, 1972) demonstrates that persons are believed to be poor primarily as a result of their own failings. Studies comparing American- and European-style welfare states suggest that one of the major reasons that Americans redistribute less than Europeans is that Americans believe that they live in an open and fair society and that if someone is poor, it is his own fault (Alesina, Glaesar and Sacerdote, 2001; Alesina and Angeletos, 2005). Similar views, however, exist in most of Europe as well, although to a lesser extent (e.g., Commission of the European Communities, 1977). The United Kingdom and Ireland are the most similar to the US; indeed, British and Irish respondents chose "laziness and lack of will power" most often as a chief cause of poverty. Even in Sweden, beneficiaries of unemployment insurance benefits were vulnerable to the same accusations of laziness applied to Americans in similar circumstances (Rainwater, 1982). Because of Sweden's preoccupation with problems of male alcoholism, the popular image of welfare chiselers applied more to men than to women (Rainwater, 1979).

This second explanation for welfare stigma was developed by Besley and Coate (1992b) as follows: Society cares about the welfare of the poor and is willing to

transfer resources to them, but only if the claimants are "deserving." Undeserving claimants are individuals who can support themselves but, nevertheless, choose to claim benefits at the expense of society at large. There is no way of distinguishing between deserving and undeserving claimants. Therefore, society attributes to *all* welfare claimants some stigma. This is a form of statistical discrimination (profiling) and results in adverse selection: the existence of unobserved "lemons" inflicts a cost on each and every member of the group.

On the individual level, the cost of stigma is determined by a combination of two elements: (i) being known as a welfare claimant; (ii) being perceived as having the ability to support oneself. The former depends on the level of exposure, which is a policy instrument controlled by the government, e.g., asking claimants to undergo procedures such as reporting in person to an agency every once in a while, standing in line at welfare agency offices or in specially designated lines in the supermarket.[4] The latter, being perceived as being able to support oneself, depends on the intimate knowledge of the people around the claimant regarding her innate earning ability.

People care about what others think about them if those others are sufficiently close to them. They care much less about what strangers think of them. Hence, stigma is mostly the result of being exposed as a welfare claimant to people who know the claimant, but are not close enough to her to actually share her interests as do family members or close friends. Exposure to people like neighbors, employers, repeated service providers or acquaintances inflicts the highest social stigma.

Focusing on Besley and Coate's (1992b) definition of welfare stigma, we re-examine the implications of stigma for the design of welfare programs. In particular,

[4] Like Besley and Coate (1992b), we believe that it is unlikely that negative self-image can arise independently of social attitudes; stigma depends on how people view the applicant. If the whole process were absolutely discreet; for example, through the internet, no stigma would arise. Hence, exposure can affect the level of stigma and thus serve as a policy instrument.

we study the role of stigma as an ordeal mechanism in enhancing the target efficiency of the welfare system.[5] We argue that stigma imposes differential costs on deserving and undeserving claimants. The reason for the differential effect is that people who know the claimant are assumed to have some information regarding her earning ability. Hence, high ability claimants are more likely to be perceived as undeserving and therefore incur higher stigma costs compared to low ability individuals. This distinction could help to sort out the needy individuals and enhance the target efficiency of the welfare system.[6]

The organization of the paper is as follows. In Section 2, we present the model. In Section 3, we analyze the government problem. Section 4 concludes.

2. The Model

We present a simple model with only the key ingredients necessary to illustrate our point. Consider an economy where individuals differ in their innate productive ability, denoted by *w*. The population is equally divided between high-ability and low-ability individuals.[7] We let $\overline{w}$ and $\underline{w}$ denote the productive ability (hence the hourly wage rate) of the high-ability type and low-ability type, respectively; where $\overline{w} > \underline{w} > 0$. We normalize the population of each ability-type to unity, with no loss in generality. For simplicity, we assume that each individual

[5] The literature on the design of welfare programs, by and large, overlooks this second type of stigma. A recent exception is Jacquet (2006), who incorporates this notion of stigma into the analysis of optimal income taxation.

[6] Kleven and Kopczuk (2005) suggest employing program complexity (such as elaborate and costly screening of claimants) as an ordeal mechanism to sort out the "pretenders." However, as indicated in the literature (see e.g., Lipsky, 1984; Blank and Ruggles, 1996; Daponte et al., 1999; Haider et al., 2003 and Orbach, 2006), this may also sort out the neediest individuals. In contrast, employing our notion of stigma does not suffer from this drawback, due to the differential effect on deserving and undeserving claimants.

[7] There are in fact more than just two types in the economy. The other types are clearly observable as non-deserving and none of these types apply for welfare benefits. Taxing the other types would finance the benefits claimed by the two types we consider explicitly. The assumption that the two types are of equal size does not affect the qualitative nature of the results.

supplies one unit of labor, hence *w* also denotes the gross labor income of an individual of type *w*. Finally, following Mirrlees (1971), we assume that an individual's innate ability is private information and unobserved by the government.

Individuals may be eligible for welfare benefits. To be eligible, an individual may be asked to report to the welfare agency and be subject to some requirements. Such procedures entail public exposure and hence give rise to stigma costs.[8]

All individuals share the same preferences represented by the following utility function:

(1) $u(c, s, \delta) = c - \delta \cdot s$,

where *c* denotes consumption, *s* denotes stigma cost and δ is an indicator function that assumes the value of 1 if the individual claims for welfare benefits, and zero otherwise.

We assume an egalitarian government whose re-distributive preferences are represented by a *Rawlsian* welfare function. That is, the government aims to attain some minimal level of well being for all individuals, denoted by $\hat{u} > 0$, at minimal cost. We assume that $\hat{u} = \hat{w}$, where $\underline{w} < \hat{w} < \overline{w}$. That is, the typical high-ability individual, whose innate ability exceeds the threshold $\hat{w}$, can attain the minimal level of well being without the need for supplementary government support, and will be henceforth referred to as an undeserving claimant. On the other hand, the low-ability individual cannot attain this minimal level of well-being without government support, and will be henceforth referred to as a deserving claimant.

Invoking the notion of social stigma used by Besley and Coate (1992b), we assume that the stigma cost, *s*, suffered by an individual of type *w* who claims for welfare benefits, takes the following form:

[8] See the discussion in footnote 4 above.

(2) $s(\gamma, w, x) = \gamma \cdot p(w, x) \cdot z(x)$,

where $0 \le \gamma \le \bar{\gamma} < 1$ is a measure of public exposure in welfare programs (determined by the government); say, the fraction of the population that knows a certain individual is claiming for benefits; x is the number of individuals who are claiming for welfare benefits; $z(x)$ measures the disutility associated with being an undeserving welfare claimant and $p(w, x)$ denotes the probability that an individual of type w is perceived to be of high-ability, conditional on the fact that x individuals are claiming for welfare. The type of stigma we focus on here derives from the fact that people cannot fully distinguish between deserving and undeserving individuals and hence assign some "cost of doubt" to all welfare claimants. Note that the literature (e.g., Moffitt, 1983) identifies another type of stigma cost, unrelated to being perceived as potentially undeserving, driven by the mere fact that welfare recipients are unable to support themselves. To put our argument regarding the potential desirable feature of stigma in sharpest relief, we set aside this other type of stigma cost, without discounting its importance.

We plausibly assume that $p(w, x)$ strictly increases with respect to w, strictly increases with respect to x, for $x \ge 1$, and satisfies $p(w, x') = 0$ for $x' \le 1$. In other words, the more productive an individual is, the more likely she will be perceived to be of higher ability, hence as an undeserving welfare claimant. Thus, introducing stigma serves as a screening mechanism to sort out the undeserving (pretending) individuals.[9] Moreover, as the stigma cost rises with respect to innate ability, deserving claimants will be the first to join the welfare system as they bear the lowest

[9] Note that imposing any sort of participation costs on welfare claimants may serve as a self-selection screening device when marginal utility diminishes with respect to consumption. However, we chose to focus on the enhancement of this self-selection effect through the imposition of differential costs on claimants. In order to capture the latter effect in sharpest relief, we simplify by assuming that utility is linear with respect to consumption.

stigma costs. Thus, the larger the number of welfare claimants, the higher the number of undeserving claimants, hence the higher the stigma cost associated with being a welfare claimant. Finally, when all individuals who claim for welfare benefits are truly unable to support themselves and to attain the threshold level of well-being (that is, $x \leq 1$), they are rationally perceived to be deserving claimants. Hence, stigma cost is zero for all individuals. We further assume that $z(x)$ strictly decreases with respect to x and that $z(2)=0$. In other words, as the number of welfare claimants increases, the disutility associated with being on welfare decreases (being a welfare recipient becomes a more prevalent social norm). When all individuals are welfare claimants (x=2) there are naturally no stigma costs. Finally, we make the technical assumption that $z''(x) < 0$.

We naturally assume that individuals are rational in their beliefs in the sense that the probability assigned to all individuals as being undeserving claimants is indeed given by the actual number of undeserving claimants. Formally (for $x \geq 1$),

(3) $$p(\overline{w}, x) + p(\underline{w}, x) = x - 1.$$

[Note that $x - 1$ is the number of undeserving claimants ("pretenders")].

One can provide a simple micro-foundation for the function $p(w, x)$. We will stick to this micro-foundation in what follows. Suppose that while an individual's productive ability is private information, other individuals may observe some signal correlated with this ability (say, years of schooling, family background, the car she drives, etc.). Specifically, suppose that the signal may assume two values, high (H) and low (L), and let $n(w)$ denote the probability that the signal's value is high when an individual is of ability w. Suppose further that $\overline{n} \equiv n(\overline{w}) > n(\underline{w}) \equiv \underline{n}$. Thus, the signal is informative, in the sense that a high-ability individual is more likely to obtain

a high signal. Let $\bar{q}$ and $\underline{q}$ denote the probability of being perceived as an undeserving claimant, conditional on observing signals *H* and L, respectively. Employing *Bayes' Rule*, it follows that $\bar{q} = (x-1)\cdot\bar{n}/(\underline{n}+\bar{n})$ and $\underline{q} = (x-1)\cdot(1-\bar{n})/(2-\bar{n}-\underline{n})$. It thus follows that $p(w,x) = n(w)\cdot(\bar{q}-\underline{q})+\underline{q}$ for $x \geq 1$ and zero otherwise. The accuracy of the signal is measured by the difference $\bar{n}-\underline{n} > 0$; as the difference increases, the signal becomes more informative, and when $\bar{n}-\underline{n} = 1$ (that is, $\bar{n}$ = 1 and $\underline{n}$ = 0) the signal is perfectly informative. To simplify the exposition, without affecting the qualitative nature of our results, we let $\bar{n} = 1-\underline{n} = q > 1/2$, where the parameter *q* measures the accuracy of the signal. It is straightforward to verify that all the properties of the model are satisfied. Substituting the above expression for $p(w,x)$ into equation (2) yields:[10]

(4a) $s(\gamma,\bar{w},x) = \gamma\cdot(2q^2-2q+1)\cdot(x-1)\cdot z(x)$,

(4b) $s(\gamma,\underline{w},x) = \gamma\cdot(-2q^2+2q)\cdot(x-1)\cdot z(x)$.

A welfare program is given by the pair $<t,\gamma>$ where *t* denotes the transfer (a uniform benefit to which an individual is entitled) and γ measures the degree of public exposure (affecting the stigma entailed by claiming for benefits). A universal system is captured in our framework by the case where $\gamma = 0$; that is, no eligibility requirements are set (or, alternatively, requirements are discreet and do not involve any public exposure). In this case, all individuals claim for benefits.[11] A selective system is captured by the case where $\gamma > 0$. In such a case, only a fraction of the

[10] We assume that individuals decide whether to apply for welfare benefits before they observe the realization of the signal (high or low).

[11] Even universal transfer programs, that is, non-excludible programs, do not have a 100% take-up rate. However, making the programs selective; that is, making the benefits available only to a selected few as in the case of means-tested programs or programs for the disabled, significantly exacerbates the take-up problem (Currie, 2006).

individuals claims for benefits. Note, however, that even when $\gamma > 0$, the case where all individuals claim for benefits forms an equilibrium. This follows from the fact that $z(2)=0$, so that no stigma costs are entailed when being a welfare claimant becomes the social norm. Note further, that a selective system will always attract some undeserving claimants. This follows from the fact that when only deserving claimants are on welfare, the stigma cost entailed by claiming for welfare is zero; this cannot form an equilibrium.

Note that an individual of type w will claim for benefits when faced with the welfare system $< t, \gamma >$, if and only if:

(5) $t - s(\gamma, w, x) \geq 0$.

Given a welfare program $< t, \gamma >$, an equilibrium is given by the number of claimants in the program, $1 < x^* \leq 2$, which satisfies:

(6) $t \geq s(\gamma, \overline{w}, x^*)$, with a strict inequality only when $x^* = 2$, in which case, all claim for welfare benefits.

3. The Optimal Policy

We next turn to formulating the optimal policy. The government seeks to attain some minimal level of utility (for all individuals) at minimal cost, by choosing the transfer (t) and the degree of exposure (γ).[12] We simplify by assuming that exposure is costless. Formally, the optimal policy is a solution to the following optimization problem:

[12] We compare a selective welfare system with a universal one. We consider the case of a hybrid system that is partially selective later on.

$$
(7) \quad \begin{aligned}
&\min_{\gamma \in [0,\bar{\gamma}],t,x} [t \cdot x] \\
&s.t. \\
&(i) \quad \underline{w} + t - s(\gamma, \underline{w}, x) \geq \hat{w} \ , \\
&(ii) \quad t \geq s(\gamma, \overline{w}, x), \text{ with a strict inequality only when } x = 2,
\end{aligned}
$$

where $s(\gamma, w, x)$ is given by equation (4).

Constraint (*i*) requires that the utility derived by all deserving claimants (the low-ability individuals) would exceed the minimal threshold, $\hat{w}$. Condition (*ii*) requires that individuals' beliefs about the deservedness of welfare claimants are self-fulfilling. It is straightforward to observe that constraint (*i*) is binding.[13]

3.1 Comparing the Selective and the Universal Systems

To obtain the optimal policy, we compare the optimal universal system with the optimal selective one. By the above considerations, the optimal universal system is simply given by $t = \hat{w} - \underline{w}, \gamma = 0$ and $x = 2$. That is, the transfer, *t*, is set at a level, which is just sufficient to ensure that the low-ability individual will attain the threshold level of utility. The cost entailed by implementing the optimal universal system is, hence, given by:

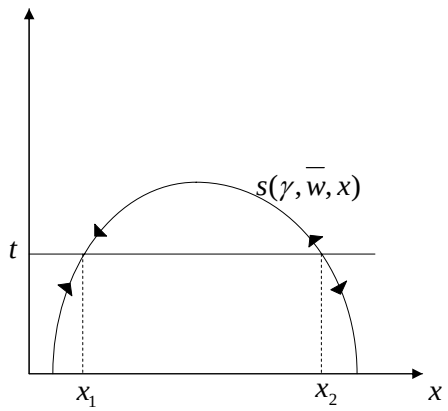
$$
(8) \qquad \cos t^{universal} = 2 \cdot (\hat{w} - \underline{w}).
$$

We next turn to the more interesting case of a selective welfare system. In such a case, $\gamma > 0$, $x < 2$ and both constraints (*i*) and (*ii*) are binding. First, it is straightforward to verify, by virtue of the assumption that $z''(x) < 0$, that

[13] To see this, suppose, by way of contradiction, that condition (*i*) is not binding. There are two cases to consider. Consider first the case where the optimum is obtained by a universal system ($\gamma = 0$), which implies that *x*=2. In such a case, constraint (*ii*) is trivially non-binding and one can slightly reduce *t* without violating either of the two constraints and reduce the cost of the welfare system, thus obtaining the desired contradiction. Next, consider the case where the optimum is obtained by implementing a selective system (setting $\gamma > 0$), which implies that *x*<2. In such a case, condition (*ii*) is binding, but then, by fixing *x* – the number of welfare claimants, one can slightly reduce both γ and *t* while maintaining constraint (*ii*) as binding and not violating condition (*i*) by continuity considerations. By reducing the cost of the welfare system, we obtain the desired contradiction.

$\partial^2 s(\gamma, \overline{w}, x) / \partial x^2 < 0$. Thus, the expression on the right-hand side of constraint (*ii*) is strictly concave; hence, for any degree of exposure, γ, and provided that the level of transfer, *t*, is sufficiently small, there exist two values of *x* for which constraint (*ii*) holds as an equality (see the figure below). Namely, there are in general two candidate equilibria for a selective welfare system. It is straightforward to verify that only one of the two equilibria is locally stable (the one for which $\partial s / \partial x > 0$, given by x_1 in the figure). We will henceforth confine attention to the stable equilibrium.[14]

Figure: Equilibria in a Selective Welfare System



Fixing the degree of public exposure, γ, a (locally stable) equilibrium for a selective welfare system is given by the (unique) pair $< t, x >$ that solves the following system of two equations given in (7):

[14] While the figure indicates the possibility of having no intersection, we will henceforth assume that a well-defined solution of the system of two constraints, (i) and (ii), for the two unknowns (*t* and *x*) exists for any degree of exposure. Note that there (always) exists a third equilibrium when the degree of exposure is positive, namely an equilibrium where all individuals claim for welfare (which is also locally stable). This equilibrium is essentially a universal program. To avoid trivial conclusions, we assume, however, that when a selective system is implemented, the relevant equilibrium will be one where only a fraction of the individuals claim for benefits.

(9) $t = \hat{w} - \underline{w} + \gamma \cdot (-2q^2 + 2q) \cdot (x-1) \cdot z(x)$,

(10) $t = \gamma \cdot (2q^2 - 2q + 1) \cdot (x-1) \cdot z(x)$.

Substituting for *t* from equation (10) into equation (9) yields:

(11) $\hat{w} - \underline{w} = \gamma \cdot (2q-1)^2 \cdot (x-1) \cdot z(x)$.

Substituting for γ from equation (11) into equation (10) yields:

(12) $t \cdot x = (\hat{w} - \underline{w}) \cdot A(q) \cdot x$,

where $A(q) = (2q^2 - 2q + 1)/(2q-1)^2$.

The expression on the right-hand-side of equation (12) measures the cost entailed by implementing a selective welfare system as a function of *x*. It follows from equation (12) that the cost increases with respect to the number of claimants. To minimize this cost, the optimal degree of exposure should be set at $\overline{\gamma}$, the highest degree that is feasible.[15] We thus conclude that the optimal selective system is given by the triplet $\overline{\gamma}$, $t(\overline{\gamma}, \hat{w}, q)$ and $x(\overline{\gamma}, \hat{w}, q)$, where *t* and *x* are given by the solution to equations (9) and (10) for $\gamma = \overline{\gamma}$. The cost entailed by implementing a selective system is given by:

(13) $\cos t^{selective} = (\hat{w} - \underline{w}) \cdot A(q) \cdot x(\overline{\gamma}, \hat{w}, q)$.

Comparing the cost entailed by the optimal universal system [equation (8)] and that entailed by the optimal selective system [equation (13)] implies:

(14) $\cos t^{selective} < \cos t^{universal} \Leftrightarrow 2 / A(q) > x(\overline{\gamma}, \hat{w}, q)$.

As $1 < x(\overline{\gamma}, \hat{w}, q) < 2$, it is straightforward to verify that for a sufficiently small *q* [any *q* for which $A(q)>2$, noting that $A'(q)<0$], the universal system dominates; whereas, for high enough values of *q* [*q* sufficiently close to 1, that is $A(q)$ sufficiently close to

[15] To see this, note that as we focus on the stable equilibrium, the expression on the right-hand side of equation (11) strictly increases with respect to *x*. Thus, as we increase γ, *x* decreases. As, by assumption, increasing γ is costless, one attains the optimum by setting $\gamma = \overline{\gamma}$. Note that allowing for costly exposure would not change the qualitative nature of our results.

1], the selective system prevails. Employing condition (11), it follows by virtue of the stability property of the equilibrium that $x(\bar{\gamma}, \hat{w}, q)$ strictly decreases with respect to *q*. By continuity, it follows, then, that there exists a unique value of *q*, above which the selective system would dominate, and below which the universal system would form the optimal program. Moreover, this threshold level of *q* strictly decreases with respect to $\bar{\gamma}$ and strictly increases with respect to $\hat{w}$. The latter implies that for a given $\bar{\gamma}$ and *q*, the selective system would dominate for sufficiently small values of $\hat{w}$, whereas the universal system would dominate for large enough values of $\hat{w}$.

The rationale for these results is straightforward. As *q* increases, the differences in stigma costs across individuals with different abilities become more pronounced. This follows from the fact that individuals can more efficiently distinguish between low- and high-ability individuals and, consequently, between more- and less- deserving welfare claimants. For a large enough *q*, the gains from introducing stigma costs, thereby deterring high-ability individuals from being welfare claimants, outweigh the costs associated with the need to increase the transfer (to ensure participation of low-ability individuals). As the degree of exposure ($\bar{\gamma}$) increases, this effect becomes more pronounced and the case for favoring a selective system becomes stronger. Finally, note that when the minimal level of well-being guaranteed by the government ($\hat{w}$) is small, the stigma cost required to support the selective system is moderate and hence the targeting advantage of that system relative to a universal one renders the former socially desirable. As $\hat{w}$ increases, the government has to raise the transfer; hence, more individuals choose to join the welfare program. This, in turn, results in higher stigma costs required to deter potential "pretenders," thereby forcing the government to further raise the transfers (in

order to ensure that deserving individuals will attain the minimal level of well-being). This amplification mechanism eventually renders the universal system (where transfers rise linearly with $\hat{w}$) socially desirable.

The result we obtain bears similarity to the result obtained by Besley and Coate (1992a). Besley and Coate justify the use of workfare as a self-selection screening mechanism to target benefits to low-ability individuals. The idea is that low-ability individuals have lower opportunity costs entailed by participating in workfare programs and hence a workfare requirement can reduce the costs of the welfare system and allow for enhanced poverty alleviation relative to a universal welfare system (that imposes no such requirements). Similarly, in our paper, shifting from a universal system to a selective one imposes stigma costs on participants (due to the inevitable existence of undeserving claimants). The effect of stigma is differential. This novel aspect of our paper derives from the fact that we plausibly assume that while the government cannot observe the individual's productive ability, it is partially observable by others. In particular, the more productive an individual is, the more likely she is to be perceived as such. Hence, the more likely she is to be tagged as an undeserving claimant. This differential effect implies that higher-ability individuals face higher stigma costs from participating in welfare programs. Thus, introducing a selective system can achieve enhanced targeting. When the differential effect is strong enough (depending on the degree of exposure feasible and the amount of information possessed by individuals) a selective system would dominate.

3.2 The Optimality of a Hybrid System

Most welfare systems are neither purely selective nor completely universal. It would be interesting to see whether a hybrid system, composed of a universal component

and a selective one, would be socially desirable. It turns out, indeed, that for a wide range of parameters, a hybrid (partially selective) system would form the optimal policy. Next, we illustrate this point.

As demonstrated in the previous section, for any minimal level of well-being set by the government, $\hat{w}$, there is a unique value of *q*, for which the cost entailed by a selective system is equal to that associated with a universal welfare system. This value of *q* is implicitly given by the solution to:

(15) $2 / A(q) = x(\bar{\gamma}, \hat{w}, q)$.

Moreover, given *q*, for any minimal level of well-being exceeding $\hat{w}$, the universal system would dominate, whereas, for any minimal level of well-being below $\hat{w}$, the selective system would prevail.

Now, fix q and let $\hat{u}$ denote the target welfare threshold set by the government. Suppose this threshold is set high enough, such that a universal system would be desirable (formally, $\hat{u} > \hat{w}$). We will show that rendering this system slightly selective would reduce the costs incurred by the government. Thus, a hybrid system would form the optimal solution.

To see this, recall that the optimal universal system is given by $t^* = \hat{u} - \underline{w}$, and, obviously, $\gamma = 0$. Now, consider an alternative system, which is partially selective. Formally, let the universal component be denoted by $0 < t' < t^*$. Employing our derivations in the previous section, the optimal selective program is given by the (unique) pair $< t, x >$ that solves the following system of two equations [similar to equations (9) and (10) above]:

(16) $t = \hat{u} - t' - \underline{w} + \bar{\gamma} \cdot (-2q^2 + 2q) \cdot (x - 1) \cdot z(x)$,

(17) $t = \bar{\gamma} \cdot (2q^2 - 2q + 1) \cdot (x - 1) \cdot z(x)$,

Note that equation (16) requires that the utility derived by a low-ability individual who participates in the selective system would be equal to $\hat{u}-t'$, so that combined with universal transfer, t', the minimal level of well-being, $\hat{u}$, will be attained. Following some algebraic manipulations (which replicate the derivations in the previous section and hence are omitted), one obtains the following expression for the cost entailed by the optimal selective program:

(18) $\cos t^{selective} = (\hat{u}-t'-\underline{w})\cdot A(q)\cdot x(\overline{\gamma},\hat{u}-t',q)$,

where $A(q)=(2q^2-2q+1)/(2q-1)^2$.

The reduction in cost entailed by the universal component (due to the reduction in transfer) is given by:

(19) $\Delta\cos t^{universal} = 2\cdot(t^*-t') = 2\cdot(\hat{u}-t'-\underline{w})$.

The hybrid system would dominate the universal system if-and-only-if the following condition holds:

(20) $$\begin{aligned}&\cos t^{selective} < \Delta\cos t^{universal} \Leftrightarrow (\hat{u}-t'-\underline{w})\cdot A(q)\cdot x(\overline{\gamma},\hat{u}-t',q) < 2\cdot(\hat{u}-t'-\underline{w})\\ &\Leftrightarrow 2/A(q) > x(\overline{\gamma},\hat{u}-t',q)\end{aligned}$$

The last inequality is satisfied when t' is sufficiently large [formally, when $t'>\hat{u}-\hat{w}>0$, by virtue of equation (15)]. As the universal system dominates the selective one, by construction, it follows that the optimal solution must be a hybrid system.

The rationale for the result is straightforward. Even when shifting entirely from a universal system to a selective one is undesirable, due to the large stigma costs entailed, replacing a small part of the universal transfer with a selective component is called for, because then the targeting advantage of the selective system would save costs without imposing high stigma costs.

4. Conclusions

In this paper, we discuss the effect of welfare stigma on the design and magnitude of welfare programs. The type of stigma we focus on here derives from the fact that people cannot fully distinguish between deserving and undeserving individuals and hence assign some "cost of doubt" to all welfare claimants.

We demonstrate that this type of stigma can be employed as an efficient ordeal mechanism to differentiate between truly deserving welfare claimants and undeserving "pretenders," relying on the public's partial ability (in contrast to the government's lack of such ability) to distinguish between the two types of welfare claimants.[16] This tool is likely to be efficient in circumstances where it is relatively easy for the public to distinguish between truly deserving welfare claimants and undeserving ones; and the minimal desired standard of living is relatively moderate. We also establish a case for a hybrid welfare system that consists of both a universal (stigma-free) component and a selective (stigma-generating) part, as most welfare systems are, in practice.

The literature identifies another type of stigma cost, unrelated to being perceived as potentially undeserving, driven by the mere fact that welfare recipients are unable to support themselves. To put our argument regarding the potential desirable feature of stigma in sharpest relief, we set aside this type of stigma cost, without discounting its importance. Clearly, choosing the best policy would depend on the relative magnitude of the two types of stigma costs, which is likely to be culture-dependent and vary across countries. A challenge for the government would be to increase the stigma costs associated with being perceived as potentially

[16] For example, welfare eligibility is often limited to individuals who cannot work. The government may encounter substantial difficulties in tracing ineligible individuals who work informally but nonetheless claim benefits. This, however, need not apply to the public that, to some extent, can observe participation in the informal sector. The 'pretending' claimants will thus suffer from a higher stigma cost than that incurred by the deserving claimants who are truly unable to work.

undeserving without affecting (or better, decreasing) the costs associated with the other notion of welfare stigma. One possible way of achieving this goal may be to publicize the names of caught welfare-chiselers, thus raising the public awareness to welfare frauds.

References

Aarts, Leo and Philip de Jong, (1996) "The Dutch Disability Program and How it Grew", in Curing the Dutch Disease, eds. Leo Aarts, Richard Burkhauser and Philip de Jong, (Aldershot, U.K.: Avebury, 1996).

Akerlof, George A., (1978) "The Economics of Tagging as Applied to the Optimal Income Tax," American Economic Review 68, 8-19.

Alesina, A., E. Glaeser, and B. Sacerdote, (2001) “Why Doesn’t the United States Have a European-style Welfare State?” Brookings Papers on Economic Activity 2, 187-278.

Alesina, A. and G.-M. Angeletos, (2005) “Fairness and Redistribution: US Versus Europe,” American Economic Review 95, 913-935.

Besley, Timothy and Stephen Coate, (1992a) “Workfare versus Welfare Incentive Arguments for Work Requirements in Poverty-Alleviation Programs,” American Economic Review 82 (1), 249—61.

---- (1992b). "Understanding welfare stigma: Taxpayer resentment and statistical discrimination," Journal of Public Economics 48, 165-184.

---- (1995). “The Design of Income Maintenance Programs” Review of Economic Studies 62 (2), 187—221.

Blank, Rebecca and Patricia Ruggles (1996). "When Do Women Use Aid to Families with Dependent Children and Food Stamps? The Dynamics of Eligibility versus Participation," 31 J. Human Resources 57.

Campolieti, Michele (2002). “Moral Hazard and Disability Insurance: On the Incidence of Hard-to-Diagnose Medical Conditions in the Canada/Quebec Pension Plan Disability Program,” Canadian Public Policy/Analyse de Politiques 28, 419-441.

Commission of the European Communities. (1977). The Perception of Poverty in Europe. Brussels, March.

Currie, Janet (2006) "The Take Up of Social Benfits" in Public Policy and the Income Distribution, Alan J. Auerbach, David Card, and John M. Quigley (editors) (New York: Russell Sage, 2006).

Daponte, Beth O., Set Sanders and Lowell Taylor (1998). "Why do low-income households not use food stamps? Evidence from an experiment," J. Hum Resources 34:612–628.

Feagin, J.R. (1972). “America’s Welfare Stereotypes,” Social Science Quarterly 52, 921-33.

Goffman, Erwin (1963). Stigma: Notes on the Management of Spoiled Identity. Prentice-Hall, Englewood Cliffs, NJ.

Haider, Steven J., A. Jacknowitz, and Robert F. Schoeni. (2003) "Food Stamps and the Elderly: Why is Participation So Low?" Journal of Human Resources, 38 (supplement): 1080-1111.

Hancock, R., et al. (2004). "The Take-up of Multiple Means-Tested Benefits by British Pensioners: Evidence from the Family Resources Survey" Fiscal Studies 25, 279-303.

Jacquet, Laurence (2006). "Optimal Disability Assistance when Fraud and Stigma Matter" Queen's Economics Department Working Paper No. 1098

King, Judith (2004). "Impediments to Gainful Employment," Social Security Journal 66, 9-34.

Kleven, Henrik and Wojciech Kopczuk (2005) "Welfare Programs Complexity and the Take Up of Social Benefits" (research paper).

Lipsky, Michael (1984) "Bureaucratic Disentitlement in Social Welfare Programs" 58 Soc. Serv. Rev. 3.

London, Rebecca A. (2003) "Which TANF Applicants Are Diverted, and What Are Their Outcomes?," 77 Social Service Review 373.

Mirrlees, James A. (1971) "An Exploration in the Theory of Optimum Income Taxation", Review of Economic Studies, 38, 175-208

Moffitt, Robert (1983). "An Economic Model of Welfare Stigma," American Economic Review 75, 1023-1035.

Nichols, A., and R. Zeckhauser (1982). "Targeting Transfers Through Restrictions on Recipients," American Economic Review, 72(2), 372—377.

Orbach, Barak, Y., (2006). "Unwelcome Benefits: Why Welfare Beneficiaries Reject Government Aid?," Law & Inequality 24: 107-150.

Pettigrew, T., (1980). "Social Psychology's Potential Contributions to an Understanding of Poverty," in Covello, V.T. (Ed.), Poverty and Public Policy. Schenkman, Cambridge, MA.

Pudney, S. et al. (2006). "Simulating the Reform of Means-Tested Benefits with Endogenous Take-Up and Claim Costs" Oxford Bulletin of Economics, 68, 135-166.

Rainwater, L. (1974). What Money Buys. New York: Basic Books.

-------- (1982) “Stigma in Income-Tested Programs,” in Garfinkel, I. (Ed.), Income-Tested Transfer Programs: The Case For and Against. Academic Press, London, pp. 19-46.

Wolff, J. (2004), “Disability Among Equals”, in Philosophy and Disability, K. Brownle and A. Cureton (eds.), Oxford University Pres.

Yaniv, Gideon (1997). "Welfare Fraud and Welfare Stigma,” Journal of Economic Psychology 18, 435-451.

THE FOERDER INSTITUTE FOR ECONOMIC RESEARCH and
THE SACKLER INSTITUTE FOR ECONOMIC STUDIES
The Eitan Berglas School of Economics
Tel Aviv university

Recent list of Working Papers

1-2008	Limor Hatsor	The Allocation of Public Education Resources
2-2008	Anat Bracha Donald J. Brown	Affective Decision Making: A Behavioral Theory of Choice
3-2008	Saul Lach Gil Shiff Manuel Trajtenberg	Together but Apart: ICT and Productivity Growth in Israel
4-2008	Ariel Rubinstein	Comments on Neuroeconomics
5-2008	Tomer Blumkin Yoram Margalioth Efraim Sadka	Incorporating Affirmative Action into the Welfare State
6-2008	Chemi Gotlibovski Yoram Weiss	Retirement Decisions of Married Couples
7-2008	Ayala Arad	The Tennis Coach Problem: A Game-Theoretic and Experimental Study
8-2008	Victor Lavy M. Daniele Paserman Analia Schlosser	Inside the Black of Box of Ability Peer Effects: Evidence from Variation in Low Achievers in the Classroom
9-2008	Hosny Zoabi	Talent Utilization, a Source of Bias in Measuring TFP
10-2008	Tomer Blumkin Yoram Margalioth Efraim Sadka	The Role of Stigma in the Design of Welfare Programs

---

Foerder Institute for Economic Research, Tel-Aviv University, Tel-Aviv, 69978 Israel, Tel: 972-3-640-9255; fax: 972-3-640-5815; e-mail: foerder@post.tau.ac.il
Papers from 2000 onwards (and some from 1998 and1999) can be downloaded from our website as follows:
http://econ.tau.ac.il/research/search_workingPapers.asp