Welfare Stigma Re-examined[♠]

Tomer Blumkin[*] **Yoram Margalioth**[**] **Efraim Sadka**[***]

May, 2013

Abstract

We dissect welfare stigma into two types: traditional and statistical, and show that the latter can be employed as a desirable form of a welfare ordeal, as its costs are positively correlated with ability.

JEL Classification: H2, D6, J1

Key Words: Stigma, Welfare Ordeal, Social Norms.

[♠] The authors wish to thank an associate editor and three anonymous referees for their comments and constructive suggestions. The usual disclaimer applies.

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

[**] The Buchmann 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

In a second best setting, an egalitarian government seeking to target benefits to the least well-off members of society is faced with a fundamental screening problem of identifying the deserving individuals. The government uses various devices to overcome this problem. Direct devices include means testing by reviewing documentation, conducting interviews, and testing by specialists.[1] Indirect screening includes *targeting groups* [*tagging*, according to Akerlof (1978)]: basing welfare eligibility on observable characteristics such as old age, level of education or observable disability, correlated with ability;[2] *targeting benefits*: offering in-kind benefits (e.g., wheelchairs, daycare services) that deserving individuals; namely, the intended beneficiaries of the program, would find relatively more attractive [Nichols and Zeckhauser (1982); Cremer and Gahvari (1997)]; and *welfare ordeals*: adding requirements, such as work or training requirements to the program (even if they entail pure deadweight costs, i.e., take the form of "digging holes") that undeserving individuals would find relatively costly and hence, self-select out of the program [Nichols and Zeckhauser (1982); Besley and Coate (1992a) and (1995)].

In this paper we argue that stigma could be used as an efficient form of a welfare ordeal. In a seminal paper, Moffitt (1983), welfare stigma was defined as a feeling of lack of self-respect from participation in welfare due to an inability to support oneself. We refer to this type of stigma as "traditional stigma".

There is, however, a second possible explanation for welfare stigma, by and large overlooked by the literature, related to fraudulent behavior. People who pretend to have no ability to work, but who in fact possess such ability and live off welfare by

[1] See Besley and Coate (1995) for a general characterization of means-tested income maintenance programs.

[2] For a more recent treatment of the role of tagging in enhancing target efficiency see Boadway and Pestieau (2006) and Cremer, Gahvari and Lozachmeur (2010).

misreporting their income, are sanctioned through stigma. According to Besley and Coate (1992), because there is no way of fully distinguishing between deserving and undeserving claimants, society attributes to *all* welfare claimants some stigma costs. We refer to this type of stigma as "statistical stigma".

In this paper we study the role of statistical stigma as an ordeal mechanism to enhance the targeting/screening efficiency of the welfare system. Besley and Coate (1992) mentioned the possibility that welfare stigma may serve a useful policy role in reducing the number of undeserving claimants. But they did so briefly, at the end of a paper that was mainly dedicated to defining welfare stigma and suggesting ways to reduce it. Later literature on welfare stigma ignored the possibility of using stigma as a policy tool, emphasizing its cost to deserving claimants. Yaniv (1997) focuses on social stigma generated through public exposure in welfare programs, and shows that such stigma may constitute a strong deterrent to participation in the welfare program; even stronger than the expected punishment for dishonest claiming. Jacquet and Van der Linden (2006) analyze the case for tagging, viewing stigma as limiting its role.[3] Kleven and Kopchuk (2011) view stigma to be particularly bad at screening, analogizing it to a pure ordeal, which they suggest should be set to zero. Indeed, a major weakness of welfare ordeals is that they are costly to the deserving claimants themselves, hence, are not necessarily welfare-enhancing. Nichols and Zeckhauser (1982) argue that a pure ordeal may serve a screening function if it imposed a higher utility cost on high-ability individuals than on low-ability individuals, conditional on income.

[3] The kind of stigma they refer to is the traditional type which they define as the shame that a welfare recipient experiences being unable to fend for oneself and one's family. Contrary to the main assumption we make in this paper, they assume that social workers "costlessly observe the ability of workers and can prevent high-ability workers from benefiting from the assistance scheme"."

Our novel contribution to the literature is in arguing that statistical stigma is positively correlated with claimants' abilities, because people who know the claimant are assumed to have some information regarding her earning ability; hence, unlike the government that is unable to observe ability, they can partially distinguish between deserving and undeserving claimants.[4]

Our paper contributes to a large strand in the law-and-economics literature that views social norms as a form of self-regulation, substituting for or complementing traditional enforcement tools used by the government (detection, monetary sanctions etc.) When conduct is regulated by social norms, stigma functions as an enforcement mechanism [see McAdams and Rasmusen (2007)]. In our context, limiting benefits to deserving claimants is enforced by a social norm, with statistical stigma (controlled by exposure) serving as the enforcement mechanism.

Employing stigma as a policy tool requires careful analysis because stigma may apply not only to the regulated individual behavior but to individual attributes as well. We examine this insight, using the case of welfare stigma, dissecting it into two types: traditional and statistical. Traditional stigma applies to being unable to provide for oneself and is an attribute. Statistical stigma has to do with fraudulent behavior. Welfare claimants, who pretend to have no ability to work, or work but misreport their income in order to be eligible for welfare, are sanctioned through stigma. We show that statistical stigma can be employed as a non-cruel type of welfare ordeal, as its costs are positively correlated with ability.

[4] Even when the government can observe individuals' abilities but doing so is costly, the government may be better off relying on the informed public's assistance in enforcing its policy, through self regulation (social norms). For example, able-bodied people would feel shame parking in highly visible spots reserved for the handicapped.

The optimal level of stigma would depend on the extent to which the welfare stigma in the specific context is mostly traditional or statistical. The current policy of trying to minimize welfare stigma in each and every case may thus not be optimal.

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

2. Model

The population (normalized to 2) is equally divided between high- and low-ability individuals.[5] We let $\overline{w}$ and $\underline{w}$ denote their respective earned income levels; where $\overline{w} > \underline{w} \geq 0$. To render our analysis meaningful, we assume that the government is faced with a screening problem [in the spirit of Mirrlees (1971)] by letting the individual's income be private information, neither observable nor verifiable by the government. This assumption captures in a tractable way a realistic feature of welfare systems where benefits are means-tested but due to imperfect monitoring by the government misreporting takes place.

We assume that all individuals share the same preferences, given by:

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

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

The government aims to attain some minimal level of well-being, denoted by $\hat{w} > 0$, at minimal cost, where $\underline{w} < \hat{w} < \overline{w}$. Ideally the government would support the low-ability (deserving claimants) only, but given our informational assumptions is faced with a screening problem (sorting out the undeserving).

[5] 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.

2.1 Types of Stigma Costs

We assume that the stigma cost, *s*, takes the following form:

(2) $s(\gamma, w, x) = \gamma \cdot [\, p(w, x) \cdot z(x) + \alpha \,]$,

where $\gamma \in \{0,1\}$ denotes the degree of public exposure in welfare programs, and assumes the value of 0 when the welfare program is discreet and the value of 1, otherwise; $x \in [0, 2]$ denotes the number of individuals claiming welfare benefits; $z(x)$ measures the disutility associated with being an undeserving claimant, $p(w, x)$ denotes the probability that an individual of type *w* is perceived to be of high-ability, conditional on having *x* individuals claiming for welfare and on being observed as a claimant, and $\alpha > 0$ denotes the disutility associated with being a welfare claimant.

The stigma cost in (2) captures the two types of stigma referred to in the introduction. The term $\gamma \cdot p(w, x) \cdot z(x)$ captures the notion of statistical stigma. The term $\gamma \cdot \alpha$ captures the notion of traditional stigma (assumed fixed across claimants, for simplicity). Notice that in the absence of public exposure claiming welfare benefits entails no stigma costs (of either type).

One can provide a simple micro-foundation for the function $p(w, x)$. Suppose that while an individual's earning ability is unobserved by the government, other individuals may observe some noisy signal correlated with this ability, denoted by *n*. Suppose that the signal may assume two values, high (*H*) and low (*L*) and further assume that the signal is independently distributed across individuals. The signal is assumed to be informative, in the sense that a realization of a high (low) signal is more likely to occur when the individual is of high (low)-ability. Formally we assume that $\Pr[n = H \,|\, w = \overline{w}] = \Pr[n = L \,|\, w = \underline{w}] = q > 1/2$. Notice that the parameter *q* measures the accuracy of the signal: the higher *q* is the more informative the signal

turns out to be (in the limit where $q \to 1$ individuals are perfectly informed). As will be shown below, in equilibrium, all low-skill individuals (whose number is given by unity by assumption) will claim welfare benefits. As the total number of claimants is given by *x*, the number of undeserving (that is, high-skill) claimants is given by *x*-1 (provided that *x*>1, which will hold in equilibrium). Denoting by p^H and p^L, the probability that an individual claiming for welfare is perceived to be an undeserving claimant, conditional on observing the signals *H* and *L*, respectively, employing *Bayes' Rule*, it follows that $p^H = (x-1)\cdot q/[(x-1)\cdot q+(1-q)]$ and $p^L = (x-1)\cdot(1-q)/[(x-1)\cdot(1-q)+q]$.[6] Assuming, for simplicity, that individuals decide whether to claim for welfare benefits before the realization of the signal, the (unconditional) probability that a high (low)-skill individual is perceived to be an undeserving claimant is given, respectively, by:

$$\text{(3a)} \qquad p(\overline{w}, x, q) = q \cdot p^H + (1-q)\cdot p^L = \left[\frac{(x-1)\cdot q^2}{(x-1)\cdot q+(1-q)} + \frac{(x-1)\cdot(1-q)^2}{(x-1)\cdot(1-q)+q}\right],$$

and

$$\text{(3b)} \qquad p(\underline{w}, x, q) = (1-q)\cdot p^H + q \cdot p^L = \left[\frac{(x-1)\cdot q\cdot(1-q)}{(x-1)\cdot q+(1-q)} + \frac{(x-1)\cdot(1-q)\cdot q}{(x-1)\cdot(1-q)+q}\right].$$

The stigma costs entailed by high (low)-skill claimants are given, respectively, by:

$$\text{(4a)} \qquad s(\gamma, \overline{w}, x, q) = \gamma\cdot[p(\overline{w}, x, q)\cdot z(x) + \alpha],$$

and

$$\text{(4b)} \qquad s(\gamma, \underline{w}, x, q) = \gamma\cdot[p(\underline{w}, x, q)\cdot z(x) + \alpha].$$

[6] For example, p^H, the probability that a claimant is perceived to be undeserving conditional of observing a high signal, is measured by the fraction of undeserving claimants with a high signal [captured by the term in the numerator $(x\text{-}1)q$] out of the total number of claimants (both undeserving and deserving) with a high signal [captured by the term in the denominator $(x\text{-}1)q+(1\text{-}q)$].

It is straightforward to verify that $p(\overline{w}, x, q) > p(\underline{w}, x, q)$ and that $p(w, x, q)$ is increasing with respect to x (for all w). In words, the more productive an individual is, the more likely she will be perceived to be of higher ability (hence undeserving). Thus, introducing stigma (by exposure) serves as a screening mechanism to sort out the undeserving (pretending) individuals due to the differential costs positively correlated with ability (deservedness).[7] As the stigma cost rises with w, deserving claimants will be the first to join the welfare system as they bear the lowest stigma costs. Thus, provided that some high-skill individuals will apply for welfare (which will be the case in equilibrium), the larger the number of welfare claimants, the higher the number of undeserving claimants will be, hence the higher the stigma cost associated with being a welfare claimant.[8]

We make two additional assumptions. We assume that $z(x)$ decreases with respect to x and that $z(2)=0$. In words, as the number of welfare claimants increases, disutility associated with being on welfare decreases. When all individuals are claimants ($x=2$), there are only traditional stigma costs.[9] Finally, we make the technical assumption that $z''(x) \leq 0$.

2.2 The Fundamental Policy Dilemma

A welfare program is given by the pair $< t, \gamma >$, where t denotes the transfer and γ measures the degree of public exposure. There are two possibilities to consider.

---

[7] Notice 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.

[8] This form of 'negative externality' exerted on all welfare claimants is due to the imperfect information about the identity of the welfare claimant. It will disappear in the limiting case where q converges to unity (in this case, no statistical stigma will be entailed by low-skill claimants).

[9] These assumptions are in line with the literature on social norms [see e.g., Lindbeck, Nyberg and Weibull (1999)].

When $\gamma = 0$ there is no public exposure, hence no stigma costs incurred. In such a case all individuals claim for benefits for any level of transfer, *t*, set by the government. Notice that we simplify by assuming full take-up when no public exposure is entailed. In light of the empirical evidence low take-up rates are also attributed to factors unrelated to stigma, such as lack of information and complexity, which we abstract from modeling explicitly. These two additional factors narrow down the set of welfare claimants. Our study thus examines the potential screening role of stigma within this restricted pool of welfare claimants.

When $\gamma = 1$, individuals who claim for benefits suffer from stigma, and unless the level of transfer set by the government is high enough, some (and potentially all) individuals may refrain from applying (creating a take-up problem). To ensure full take-up by deserving claimants the government has to offer a sufficiently generous transfer. In such a case, the system will always attract some undeserving claimants, as well ($x>1$). This follows from the fact that when only deserving claimants are on welfare, the statistical stigma cost entailed by claiming for welfare is zero for both deserving and undeserving claimants [follows by substituting $x=1$ into (3a) and (3b)].

The fundamental trade-off faced by the government is between take-up and targeting. Setting exposure to zero implies no participation costs thereby ensuring full take-up. However, it attains poor targeting as benefits are provided to low- and high-ability claimants likewise. When public exposure is introduced, for example, by providing benefits in-kind, as in the case of food stamps[10] or public housing, the government has to raise the transfers to ensure participation and compensate for the stigma costs entailed. Overall cost may be reduced and redistribution enhanced, thanks to the differential stigma costs entailed by claimants (due to the statistical

[10] Supplemental Nutrition Assistance Program (SNAP).

stigma component) that improves the sorting out of undeserving (high-ability) claimants, thereby attaining better targeting of benefits.

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

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

Assuming the transfer is set high enough to ensure all low ability individuals apply, an equilibrium is given by the number of claimants in the program, $1 < x^* \leq 2$, which satisfies:

(6) $t \geq s(\gamma, \overline{w}, x^*, q)$, with an equality holding when $x^* < 2$.[11]

3. Characterizing the Optimal Policy

Formally, the optimal policy is a solution to the following optimization problem:

$$
\begin{aligned}
&\min_{\gamma \in \{0,1\}, t} [t \cdot x] \\
&s.t. \\
&(i) \quad \underline{w} + t - s(\gamma, \underline{w}, x, q) \geq \hat{w} \ , \\
&(ii) \quad t \geq s(\gamma, \overline{w}, x, q), \text{ with an equality holding when } x < 2,
\end{aligned}
\tag{7}
$$

Constraint 7(*i*) requires that the utility derived by deserving claimants would exceed the minimal threshold, $\hat{w}$ (this condition implies that all low-ability individual will apply for welfare benefits as $\hat{w} - \underline{w} > 0$). Condition 7(*ii*) serves to determine the number of welfare claimants in equilibrium. In an interior equilibrium (i.e., $x < 2$), high-ability individuals are indifferent between claiming or not. The number of high-

[11] We stick to an informal exposition to abbreviate notation, but strictly speaking we define a *Rational-Expectations Equilibrium*, where beliefs (perceptions) given by the probability expressions in (3a) and (3b) are consistent with the actual choices of the individuals (whether or not to apply for welfare benefits) and derived by *Bayesian* updating.

ability claimants just suffices to raise the stigma costs to a level that renders this indifference.

3.1. Stigma-Free vs. Stigma-Inducing Systems

We compare the optimal stigma-free system with the optimal stigma-inducing one. The optimal stigma-free system is simply given by $t = \hat{w} - \underline{w}, \gamma = 0$ and $x = 2$. That is, exposure is set to zero and 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 [constraint 7(*i*) is binding]. The cost entailed by implementing the optimal stigma-free system, hence, is given by:

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

We next turn to the more interesting case of a stigma-inducing welfare system. In such a case, $\gamma = 1$ and $x < 2$. By differentiating the expression in (4a) employing the assumption that $z''(x) \leq 0$, it is straightforward to verify that $\partial^2 s(1, \overline{w}, x, q) / \partial x^2 < 0$. Thus, the expression on the right-hand side of constraint 7(*ii*) is strictly concave; hence, provided that the level of transfer, *t*, is sufficiently small, there exist two values of *x* for which constraint 7(*ii*) holds as an equality (see figure below).[12] Namely, there are in general two candidate equilibria for a stigma-inducing 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 our attention to the stable equilibrium. It is also straightforward to

[12] This concave pattern of the stigma function is driven by two conflicting forces at play: (i) the probability of being perceived as an undeserving claimant, which rises with respect to the number of undeserving applicants; and (ii) the disutility associated with being an undeserving claimant which decreases as the norm of claiming benefits becomes more prevalent.

observe that constraint 7(*i*) is binding in the optimal solution.[13] 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, 7(i) and 7(ii), for the two unknowns (*t* and *x*) exists. Note that there also exists a third equilibrium where all individuals claim for welfare (which is also locally stable). This equilibrium is clearly dominated by the stigma-free system due to the traditional stigma costs entailed by claimants under the former configuration (the induced stigma does not confer any targeting benefits due to the absence of statistical stigma). To avoid trivial conclusions, we assume, however, that when a stigma-inducing system is implemented, the relevant equilibrium will be one where only a fraction of the individuals claim for benefits (an 'interior' equilibrium). In the illustrative numerical example we provide below there will be certain parameter values for which the 'interior' equilibrium will not exist. In such a case, the stigma inducing regime will result in the corner equilibrium in which all high-skill individuals will claim for welfare.

A (locally stable) equilibrium for a stigma-inducing welfare system is given by the (unique) pair $< t, x >$ that solves the following system of two equations given in (7):

(9) $$t = \hat{w} - \underline{w} + p(\underline{w}, x, q) \cdot z(x) + \alpha \,,$$

(10) $$t = p(\overline{w}, x, q) \cdot z(x) + \alpha \,.$$

[13] To see this, suppose, by way of contradiction, that condition 7(*i*) is not binding. Then one can slightly reduce *t*. By virtue of the stability property, the number of claimants will drop in response, to maintain the equality of 7(*ii*). This will occur without violating condition 7(*i*), by continuity considerations. By reducing the cost of the welfare system, we obtain the desired contradiction.

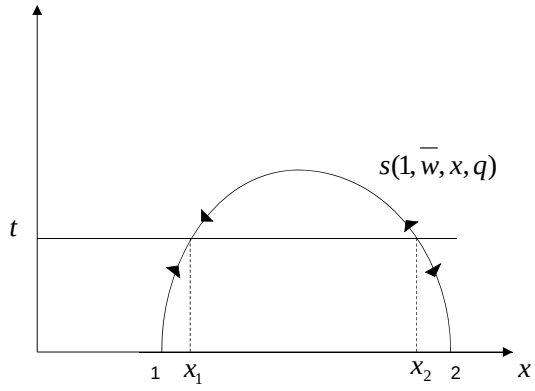

Figure 1: Equilibria in a Stigma-Inducing Welfare System

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

(11) $\hat{w} - \underline{w} = [\,p(\overline{w}, x, q) - p(\underline{w}, x, q)] \cdot z(x)$.

Substituting for *z*(*x*) from equation (11) into equation (10) yields the following expression for the cost entailed by implementing the stigma-inducing system:

(12) $\cos t^{stigma\_inducing} = x \cdot t = x \cdot [A(x,q) + \alpha]$,

where $A(x,q) \equiv \dfrac{p(\overline{w}, x, q) \cdot (\hat{w} - \underline{w})}{p(\overline{w}, x, q) - p(\underline{w}, x, q)}$, and $x(\hat{w}, q)$ is given by the implicit solution to the equation in (11).

The trade-off between the stigma-free and the stigma-inducing systems is reflected in the fact that the transfer per recipient is higher under the stigma-inducing system [$A(x,q) + \alpha > \hat{w} - \underline{w}$, for q<1 and $\alpha > 0$] so as to compensate for both types of stigma costs; whereas, the number of claimants is higher under the stigma-free system [$x(\hat{w}, q)$<2], as the presence of statistical stigma sorts out some of the undeserving claimants. Comparing the cost entailed by the optimal stigma-free system [equation (8)] and that entailed by the optimal stigma-inducing system [equation (12)] implies:

(13) $\cos t^{stigma\_inducing} < \cos t^{stigma\_free} \Leftrightarrow 2 \cdot (\hat{w} - \underline{w}) > x(\hat{w}, q) \cdot \left[A[x(\hat{w}, q), q] + \alpha \right].$

A first simple analytical result can be derived by considering the limiting case, letting $q \to 1$ and $\alpha \to 0$, yielding:

(14) $\cos t^{stigma\_inducing} < \cos t^{stigma\_free} \Leftrightarrow 2 > x(\hat{w}, q)$.

By continuity considerations, as $1 < x(\hat{w}, q) < 2$, it follows that for values of q sufficiently high and values of α sufficiently low, the stigma-inducing system prevails.[14]

The rationale for the result is straightforward. As q increases, the difference in stigma costs across individuals with different abilities becomes 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 and small enough α, the targeting gains from introducing statistical stigma, by imposing differential stigma costs on low- and high-ability claimants, thereby sorting out some of the high-ability individuals from the welfare system, outweigh the costs associated with the need to increase the transfer to ensure the minimal guaranteed level of well-being for low-ability individuals, compensating them for the stigma costs incurred (due to both statistical stigma and traditional stigma). In the limiting case where $q \to 1$ and $\alpha \to 0$ low-ability claimants incur no stigma costs at all and there is no need for compensation relative to the stigma-free system.

To further explore the optimal policy we conducted some simulations.

[14] Recall that $q \to 1$ implies that $p(\overline{w}, \cdot) \to 1$, and $p(\underline{w}, \cdot) \to 0$.

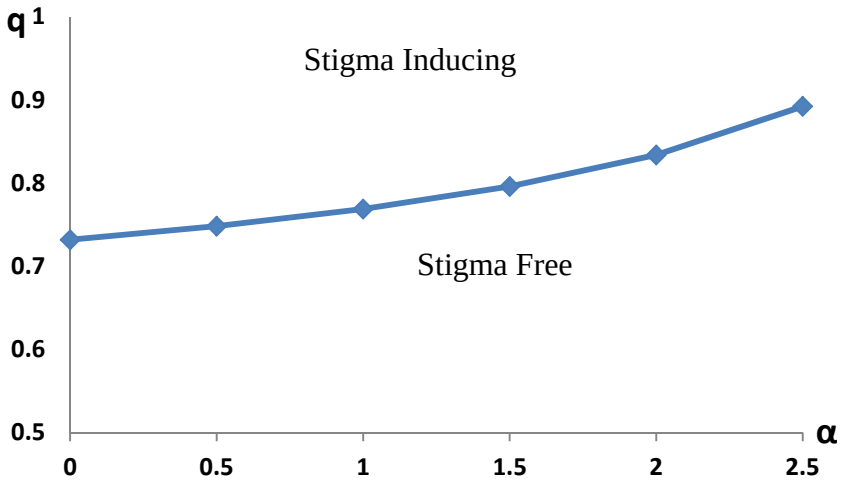

Figure 2: The Optimal Stigma Regime

Figure 2 depicts the optimal regime for different values of q and α. The parametric assumptions used for the calculation of the figure are: $z(x)=200-100x$, $\underline{w}=0$ and $\hat{w}=3$. Several remarks are in order. Our calculations indicate that there exists a stable interior equilibrium under the stigma-inducing regime for any $q>0.62155$. For values of q lower than this threshold, the only stable equilibrium is the corner one where all high-skill individuals claim benefits. This equilibrium is strictly dominated by the stigma-free regime (due to the traditional stigma costs entailed) for any positive value of α. Within the range of parameters for which an interior stable equilibrium exists, there exists a threshold level of q for each value of α above which the stigma-inducing regime prevails and below which the stigma-free regime is welfare superior. Notably, this threshold is monotonically increasing with respect to α. The observed patterns derive from the fact that as q increases, the individuals become more informed about the identity of the welfare claimants. This enhances the targeting advantages associated with statistical stigma which imposes differential costs on deserving- and undeserving claimants. In equilibrium, under the stigma-inducing regime, the number of welfare claimants is decreasing with respect to

q. Notably, the level of transfer offered to claimants [which is equal to the stigma costs entailed by the undeserving claimants, by virtue of condition (10)] is also decreasing with respect to *q*. There are two conflicting effects on the level of stigma costs incurred. For a fixed number of claimants, an increase in *q* results in a corresponding increase in the level of stigma costs. However, the resulting decrease in the number of claimants reduces the level of stigma costs. Our calculations indicate that the second effect dominates. Hence, all is all, both the number of claimants and the level of transfer decrease with *q*, hence, total government expenditure decreases with respect to *q*. For sufficiently high *q* total government expenditure under the stigma-inducing regime would be lower than that under the stigma-free regime (which is independent of *q*). The threshold level of *q* above which the stigma-inducing regime dominates is increasing with respect to the level of traditional stigma, as the screening gains associated with statistical stigma need to be large enough to compensate for the extra costs entailed due to the traditional stigma to warrant the implementation of a stigma-inducing regime.

4. Conclusion

A grave view of welfare stigma dominates the literature and guides policymakers. For example, the desire to reduce the traditional welfare stigma was the rationale for 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. Another example is the support for expanding the Earned Income Tax Credit (EITC) program, inter alia, as a means to eliminate welfare stigma.

In this paper, we re-examine the implications of stigma for the design of welfare programs, by offering a potential role for stigma as an ordeal mechanism

enhancing the targeting efficiency of the welfare system. The rationale underlying welfare ordeals is that the less needy derive lower gains from participation in the welfare program. Therefore, imposing participation costs on welfare claimants may sort out the less needy individuals. However, this could prove to be a double-edged sword as the ordeal inflicts costs on the deserving claimants as well.

We argue that employing statistical stigma as a welfare ordeal has the advantage of imposing differential costs positively correlated with ability (welfare deservedness). That said, we do not advocate the immediate enhancement of welfare stigma, but rather call into question the current policy of minimizing stigma costs in all circumstances.

References

Besley, T., and S. Coate. 1992. "Understanding Welfare Stigma: Taxpayer Resentment and Statistical Discrimination," *Journal of Public Economics* 48: 165-184.

_____. 1992a. “Workfare versus Welfare Incentive Arguments for Work Requirements in Poverty-Alleviation Programs,” *American Economic Review* 82: 249-61.

_____. 1995. “The Design of Income Maintenance Programs,” *Review of Economic Studies* 62: 187-221.

Boadway, R., and P. Pestieau. 2006. "Tagging and Redistributive Taxation," Annales d'Economie et de Statistique, ENSAE, issue 83-84, pages 05, Juillet-D.

Cremer, H., F. Gahvari. 1997. "In-Kind Transfers, Self-Selection and Optimal Tax Policy", *European Economic Review* 41:97-114.

Cremer, H., F. Gahvari and J.M. Lozachmeur. 2010. “Tagging and Income Taxation: Theory and an Application", *American Economic Journal: Economic Policy* 2:31-50.

Jacquet, L. and B. Van der Linden. 2006. “The Normative Analysis of Tagging Revisited: Dealing with Stigmatization,” *FinanzArchiv* 62: 168-198.

Kleven, H. and W. Kopczuk. 2011. "Transfer Program Complexity and the Take-Up of Social Benefits." *American Economic Journal: Economic Policy* 3: 54-90.

Lindbeck, A., S. Nyberg, and J. Weibull. 1999. "Social Norms and Economic Incentives in the Welfare State", *Quarterly Journal of Economics* 114:1-35.

McAdams, R., and E. Rasmusen. 2007. “Norms in Law and Economics,” in *Handbook of Law and Economics*, vol. 2, ed. A. Polinsky and S. Shavell. Amsterdam: North-Holland.

Moffitt, R. 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* 72:372-377.

Yaniv, G. 1997. “Welfare Fraud and Welfare Stigma,” *Journal of Economic Psychology* 18: 435-451.