

Does Welfare Ruin the Poor?

Edward J. Bird

Working Paper No. 5

June 1995

June 1995

This is work in progress, and comments are welcome.

Does Welfare Ruin the Poor?

Edward J. Bird, Assistant Professor of Political Science and Public Policy
W. Allen Wallis Institute of Political Economy
University of Rochester, Rochester, NY 14627-0158
Phone: (716) 275-7840. Fax: (716) 271-3900. email: ejbd@troi.cc.rochester.edu

Abstract: *Stigma*, the imposition of specific punishments for violating norms, has been said to explain the reluctance of eligible people to claim their welfare benefits. However, previous research has only shown that welfare has some unobserved cost; it has not shown that this cost is stigma, nor that anyone would find it optimal to impose stigma if expected to do so. This paper explores several equilibrium mechanisms that create unobserved costs of welfare receipt, including two that generate stigma. It proves that stigma can emerge in equilibrium as a result of cooperation and coordination failures. Aggregated across time and over thousands of interactions, even slight failures can impose considerable costs on 'undeserving' recipients. The paper also explores the interaction between stigma and public policy. Increases in welfare benefits erode the power of work norms, in the sense that fewer people pay attention to them. Many U.S. states have adopted the strategy of reducing poverty by building and sustaining work norms. The paper shows this to be an effective but costly approach.

Thank you for contributing important ideas: Randall Calvert, David Austen-Smith, Charles Michalopoulos, James Johnson.

I. Does Welfare Ruin the Poor?

In the 19th century, many held the view that poverty was a mark of bad character. 'Indiscriminate' charity and the giving of alms to the 'undeserving' poor was thought to be socially dangerous, because it rewarded immorality and hence encouraged more poverty. Without a morals test, public aid would degrade the character of the poor and create a sea of decadent vagrants (Himmelfarb, 1991).

Such was the theory, which still has proponents and opponents today. The goal of this paper is to see whether the theory makes sense given what we now know about social interaction among rational, self-interested individuals. Can we explain the existence of a social notion of the 'undeserving welfare recipient' as equilibrium behavior, and if we can, is it true that being more generous increases the number of such recipients?¹

The literature on welfare stigma seems the natural place to look for equilibrium notions of deserved and undeserved welfare receipt. Section II will argue, however, that the literature has yet to produce convincing evidence that stigma exists. Part of the problem is definitional. Here, I will define *stigma* as any cost of claiming welfare that is a) generated by the specific actions of some person in society, and b) occurs if and only if the welfare claimant is violating a social norm of self-sufficiency by not working when she or he 'ought' to. In the existing literature, however, 'stigma' is used to refer to anything that makes a dollar of public assistance worth less than a dollar of other income. Moffitt (1983) shows that some such cost exists empirically, but he does not show that the cost is related to stigma as defined here. Besley and Coate (1992) provide theories that assign the cost to two things: employer signals about the welfare recipient's ability, and taxpayers' resentment at having to

pay taxes to support the recipients. Neither of these causes involve social norms; neither invokes any notion of who 'ought' to receive benefits; neither uses terms like 'deserving' and 'undeserving.' Thus while these two leading papers in the welfare stigma literature make sound and convincing arguments about the costs of welfare receipt in general, they have nothing specifically to say about welfare stigma as defined here. They do not tell us why (or whether) society forms moral notions about the deserving and undeserving poor, nor why anyone would *act* on those notions in a way that produces real material consequences for others.

Any such activity is surely important from a practical standpoint. The Victorians were concerned lest social assistance policies erode the basic willingness to supply labor. In our day, welfare reforms by US states have focused again on behavioral norms, as recipients have been required to attend school, avoid out-of-wedlock childbirth, and avoid drug use or other petty crimes. From an academic standpoint, there has been a surge in interest in cultural phenomena (Lindbeck, 1994), values (Aaron, Taylor and Mann, 1993), and social capital (Putnam, 1995). It is an appropriate time to re-examine the relationship between public policies and social norms in general, and welfare stigma in particular.

This paper focuses on two specific questions regarding welfare stigma: existence and interaction with public policy. The first aspect is the more challenging of the two. Stigma, if it exists as defined here, is an action by one person to impose a cost on another. That action happens because and only because the stigmatized person has violated a work norm. It is a puzzle worth looking into why anyone would find it in their self-interest to perform such an action. Why would anyone want to lower the utility of a poor person merely

because that person seems to be 'undeserving?' Yet the paper provides two good reasons, both based on the simple idea that imposing stigma can make sense when individual behavior is affected by expectations about the behavior of others. Consider the repeated prisoner's dilemma. Two people, A and B, will be better off if they cooperate than if they do not. A's decision to cooperate depends on whether she expects B to cooperate. If there is some common-knowledge event that leads A not to expect cooperation, then A will not cooperate. Furthermore, B can work through A's logic and will therefore not expect A to cooperate. This in turn induces B not to cooperate, which confirms A's expectation. The basic insight of the paper is to treat welfare receipt, if in violation of a work norm, as a common knowledge event that induces failures in cooperation and coordination between the unfortunate recipient and those whom she encounters. These failures lower her utility and thus generate a cost of applying for welfare. In this story, stigma is not the result of an angry outburst, nor is it the result of internalized shame. It is a stream of thousands of awkward moments that alienate the 'unworthy' recipient.

To make this mechanism convincing as an explanation of stigma, it must be distinguished from other equilibrium mechanisms (such as those produced by Besley and Coate, 1992) that generate a cost of welfare claims. The paper does this in sections III, IV, and V. Section III sets up a basic model of welfare receipt that owes much to Besley and Coate's (1992) model. Section IV presents Moffitt's (1983) theory in the context of the Section III model. Section V does the same for one of the Besley and Coate (1992) models. The purpose of these sections IV and V is twofold: to allow comparison of existing theories to the new theories about to be presented, and to improve the existing theories by giving

their non-stigma explanations of welfare claim costs a better foundation as social equilibria. As it stands now, these foundations are weak. Neither Moffitt (1983) nor Besley and Coate (1992) show that the imposition of the *non-stigma* costs they analyze is incentive compatible for the actors presumed to impose them. Sections IV and V do this.

Section VI then discusses the general problem of making the imposition of stigma-related costs incentive-compatible. Section VII achieves this through a repeated prisoner's dilemma, while Section VIII uses a single-shot coordination game instead. Section IX presents a comparative static analysis of the stigma models. Section X provides general conclusions.

The main result of the paper is to demonstrate that stigma, defined as a cost of welfare claims experienced if and only if the claimant violates a work norm, can exist in equilibrium. It need not exist, however. When stigma exists, it functions in surprising ways. Welfare generosity does seem to 'ruin the poor' as the Victorians feared, because increases in welfare benefits cause more people to ignore work norms. Increasing the generosity of welfare has a positive feedback effect on the caseload: the increase in participation tends to encourage further participation. Making work norms more stringent does not, however, encourage more people to pay attention to them. It does not reduce the population of people who break the norms; on the contrary, harsher work requirements command less respect and increase the size of the defecting population.

II. The surprising absence of stigma as equilibrium behavior or empirical fact.

The stylized fact behind this paper is the finding that take-up rates for public

assistance are usually below 100 percent. Food Stamp take-up is between 50 and 60 percent (Ranney and Kushman, 1987); full participation would increase expenditures by 60 percent (Blank and Ruggles, 1993). AFDC take-up has risen and fallen over the last thirty years, with a low of 45 percent in 1967 and a high of 97 percent in 1979 (Moffitt, 1987). Most studies of AFDC participation find unexplained time effects, as though the underlying propensity to claim welfare fluctuates significantly across time (e.g. Clarke and Strauss, 1994; Hoynes and MaCurdy, 1994).

The unobserved forces that make take-up rates less than 100 percent are often referred to in general as 'stigma.' Moffitt (1983) follows Weisbrod (1970) in this, modeling stigma as anything that makes a dollar of welfare income worth less in utility terms than a dollar from other sources (see also Fraker and Moffitt, 1988). He does find that some such force, which I will call the 'welfare claim cost,' exists as an empirical matter, but of course that does not mean that stigma exists in the way it is defined here.²

What could the welfare claim cost be? It could be stigma in the sense we are using it here. Alternatively, Moffitt suggests that the hassles one faces in making a welfare claim generate a significant time and comfort cost. Besley and Coate (1992) trace the welfare claim cost to two forces: an employer's reluctance to hire welfare recipients, and a taxpayer's resentment of the tax costs those recipients generate. Neither Moffitt's nor Besley and Coate's explanations involve norm violations, however. They do not really explore the stigma as a socially-constructed moral idea. Rather, all three explain how the welfare claim cost can emerge for reasons that have nothing to do with the notions of 'the deserving poor.'

Moreover, though the non-stigma explanations for the welfare claim cost are

interesting and intuitive, they are not compelling explanations as currently presented. In simplest terms, these theories fail to consider the incentive compatibility of their welfare claim costs. In this they share a flaw evident in almost all of the new applied literature on social effects, i.e. peer pressure (Kandel and Lazear, 1992), custom (Akerlof, 1980), conformity (Jones, 1984), tax evasion (Cowell, 1990; Gordon, 1989), and so on. In all of this work, social effects matter because they are assumed to influence individual utility. Thus, Moffitt (1983) has the welfare claim cost as a set of parameters that directly lower utility as the welfare claim increases. Or, Besley and Coate's (1992) taxpayer resentment model has the recipient's utility simply falling as the total cost of the welfare system rises. No one explains where the utility effects come from. Who is imposing the cost? How are they doing it, and why do they find it worthwhile to do so?

Thus in exploring welfare stigma, it is first necessary to provide some equilibrium justification for both the Moffitt (1983) and the Besley and Coate (1992) models. Fortunately, a growing theoretical literature in incentive-compatible social effects provides just the right resources for this. In general, individual agents can find it rational to enforce norms, impose peer pressure, indulge in fads and fashions, and conform to custom (see Bikhchandani, Hirshleifer, and Welch (1992) for an information-based fad model and a survey; in this light see also Kandori, 1992; and Matsuyama, 1992; on the role of leadership and communication, see Calvert, 1992; Calvert and Banks, 1993; Farrell, 1987 and 1993; Farrell and Gibbons, 1989; Crawford and Haller, 1990; and DeMarzo, 1992; for more general ideas see Hardin, 1990; Schotter, 1981; Kreps, 1990; Binmore, 1994, p. 139ff; Coleman, 1990; and Johnson, 1993). This theoretical literature has only rarely been applied

to particular real-world problems, e.g. Greif (1992).

In sum, the literature on social effects has not provided convincing theoretical arguments or convincing empirical evidence that welfare stigma exists. Thus we cannot yet explore important practical questions, such as whether the generosity of the welfare system has some influence on welfare stigma.

III. A Basic Model of Welfare Receipt

We can explore the influence of social forces in the context of welfare stigma through a simple model of welfare receipt. A good deal of the model's structure comes from Besley and Coate (1992).

In a population of N individuals, each person is endowed with a unit of labor which can be supplied to the market at the wage rate w_i for individual i . Work imposes a disutility of λ_i . Both w and λ are distributed uniformly, wages on $[w_l, w_u]$ and labor disutility on $[\lambda_l, \lambda_u]$. Individuals have two choices: whether or not to offer their labor, and whether or not to apply for welfare. The variable c (omitting the i subscript) equals one if the individual claims, and zero otherwise. The authorities, on the other hand, decide whether to grant the claim ($d = 1$) or not ($d=0$). I assume that the welfare bureaucracy functions under simple income eligibility rules and cannot observe wages or labor supply. Therefore, claims are granted to any applicant whose income $w \cdot l$ is less than the exogenous legal poverty line p . The grant, g , is assumed to take a simple form, replacing all income below the poverty line: $g = p - wl$. Total income is $y = cd p + (1-cd)wl$. Income provides utility according to a continuous, twice-differentiable function $u(y)$ with $u' > 0$, $u'' < 0$, and $u(0) = -\infty$. Total

utility is $u(y) - \lambda l$.

With these assumptions, individual choices boil down to either working and receiving income w , or not working and receiving a grant $g = p$. To see this, consider the case of $w < p$. Any claim will be accepted by the authorities, so that $d = 1$ if $c = 1$. In the case of a claim, however, the individual receives the income p whether or not she works. Given any labor disutility, she will choose not to work. On the other hand, should she choose not to claim benefits ($c = 0$) she will receive infinitely negative utility if she also refrains from working. Thus, for this individual, the choice is to work or to claim. Similar reasoning shows that individuals with $w > p$ also choose between working and claiming welfare.

This model allows us to see clearly the puzzle addressed by the welfare stigma literature. Those with wages above the poverty line will choose to work if $u(w) - \lambda > u(p)$. Some will do so, but others will not, responding instead to the work disincentive provided by the grant structure. Those who do not work receive welfare because the government observes only income, not wages or labor supply. The situation is less complex among those entitled to welfare, with wages below the poverty line. If $w < p$, then $u(w) - \lambda$ exceeds $u(p)$ for any positive λ . The model therefore predicts that all eligible people claim the public assistance to which they are entitled. In empirical fact, of course, they do not.

The next four models give reasons why. They share the same strategy of finding some factor X , the welfare claim cost, that causes a decrease in utility when welfare is claimed, so that utility becomes $u(y) - \lambda l - Xc$.

IV. The Hassle/GUF Model

One approach is to treat X as an exogenous and constant parameter of the model. This is Moffitt's (1983) interpretation; X is guilt, an internalized unpleasant emotion associated with doing something the recipient thinks is wrong. Of course, one can explain many things with a guilt-in-the-utility function (GUF) model, from tax compliance (Gordon, 1989) to work effort (Akerlof, 1980). Alternatively, X may instead be thought of as an unobserved material cost of making a welfare claim. Applying for welfare is known to be unpleasant and time-consuming, and this probably has an effect on the frequency of claims (Kane and Bane, 1994; Strauss, 1977; Corbett, 1994; Rank, 1994).

To put these ideas in an equilibrium framework, suppose utility is $u(y) - \lambda l - Xc$, y defined as above, so that individuals with $w < p$ will work if $u(w) - \lambda > u(p) - X$. By continuity of $u(\cdot)$, there exists for every wage w a critical value $\lambda_0 = u(w) - u(p) + X$, such that all individuals with wage w will claim welfare if their labor disutility exceeds λ_0 . Figure 1 illustrates the function $\lambda_0(w)$, which will play an important general role in providing intuition for the results of all the models in the paper. The dotted box in the figure represents the population, which is distributed uniformly and independently by wage and by labor disutility. The two curves in the figure show the critical labor disutility required to induce a welfare claim (and non-work), when there is no cost to obtaining welfare, and when this cost is X . When the cost is zero, all those with wages below the legal eligibility line p claim welfare (regions A and D). Moreover, some of those with higher wages cheat and claim welfare anyway (regions B and E). Again, these claims cannot be denied because it is assumed that the government observes only income, not wages or labor supply. When the

welfare claim cost is X , cheating is reduced from $B+E$ to just B , but the take-up among eligibles falls from $A+D$ to just A . Individuals in D are eligible for benefits but work anyway. The welfare caseload, C , is the area AB :

$$C = \int_{w_l}^{w'} \lambda_u - \lambda_0(w) dw \quad (1)$$

where w' will depend on the position of the λ_0 curve; if $\lambda_0(w_u) < \lambda_u$, $w' = w_u$, otherwise w' solves $\lambda_u = u(w') - u(p) + X$. Note that, all else constant, increases in the welfare claim cost will shift the λ_0 curve upward, reducing the caseload. Similarly, an increase in the generosity of welfare will shift λ_0 downward, increasing the caseload. Lastly, note that there is no reason to expect that the λ_0 curve will fall anywhere in the dotted box. If the curve falls entirely above the box, no one claims welfare. If it falls entirely below it, everyone claims welfare.³

This simple model of internalized or simplistic costs of welfare claims is sufficient to explain the fact that take-up rates are often below 100 percent. Such behavior can be incentive compatible for all involved, even in the complete absence of any kind of stigma as defined here. Thus, the mere fact that take-up rates are low does not indicate that welfare stigma exists.

Before we examine more sophisticated models, it is useful to note that the assumptions of the GUF/Hassle model can be built into all of them. Because the GUF parameter X is fixed, it can be included as part of the labor disutility parameter. Hence in what follows, λ represents the *net* disutility of work. It is equal to the basic disutility of work, *minus* the disutility of welfare receipt.

V. Welfare Receipt as an Ability Signal

Besley and Coate (1992) propose that the welfare claim cost is the result of a more complicated social process involving the information content of the status 'former welfare recipient.' In the most general terms, suppose that economic agents are more inclined to engage in mutually-beneficial transactions with individuals who are endowed with high levels of some ability. If welfare receipt is informative in equilibrium, that is, if people can conclude that ability is systematically lower in the welfare population than the population in general, then the choice of work vs. welfare receipt will involve a material cost related to the lower returns that low-ability individuals can expect to receive. The Besley and Coate model applies this logic to the relationship between welfare reciprocity and potential future employment.

This is an intuitive explanation for low take-up, but, as argued above, the model as presented in Besley and Coate (1992) relies heavily on utility assumptions to achieve its results. In particular, they do not show that an equilibrium exists in which employers would find it in their self-interest to respond to the signal of welfare receipt. This section provides an existence proof. For all the results that follow, the solution concept is Bayesian equilibrium.

Employers are assumed to be uncertain about the labor disutility, and hence the productivity, of the prospective employees. A free market in labor forces all employers to offer each worker a wage equal to the expected value of the worker's marginal product. Further, the value of marginal product of a worker is $Q_i - k\lambda_i$, where Q and k are parameters. The Q_i are distributed uniformly and independently of λ on $[Q_1, Q_2]$, and let Q_1

$> k\lambda_u$ so that all workers produce positive output value. Employers observe Q at the individual level, but they only observe labor disutility in the aggregate. Thus they offer the wage $w_i = Q_i - kE(\lambda_i)$, where $E(\cdot)$ is an expectation operator over the aggregate observed by the employer. An employer who observes Q and some worker characteristic Z will pay wages $w_i = Q_i - kE(\lambda_i | Q_i, Z_i)$.

Let there be two periods. In period one everyone has the option of working or claiming welfare, while in period two everyone works. Let $\bar{\lambda}$, $\bar{\lambda}_{iw}$, and $\bar{\lambda}_{in}$ be the mean labor disutility in the whole population, among period-1 workers (as expected by worker i), and among period-1 non-workers (as expected by worker i), respectively. In period 1, $w_i = Q_i - k\bar{\lambda}$, all i . Let v_i be the expected period-two wage; $v_{iw} = Q_i - k\bar{\lambda}_{iw}$ for those who worked in period 1, and $v_{in} = Q_i - k\bar{\lambda}_{in}$ for those who did not. Two-period utility is $[u(y) - \lambda_i]$ + $[u(v_i) - \lambda_i]$, and an individual will work in period 1 if $u(w_i) - \lambda_i + u(v_{iw}) - \lambda_i$ exceeds $u(p) + u(v_{in}) - \lambda_i$. Thus even if there is a net utility advantage from welfare receipt in the first period ($u(p) - [u(w) - \lambda] > 0$), the second-period wage cost of welfare receipt ($u(v_{in}) - u(v_{iw})$) may be positive and large enough to dissuade a welfare claim. The wage differential is thus the welfare claim cost.

Theorem 1: Under these assumptions, there exists an equilibrium in a) welfare receipt is a credible signal of low productivity, and b) some people who are eligible for welfare choose not to claim it.

Proof. To simplify the proof initially, assume all individuals have the same base

productivity Q . Under the information conditions given, there will be a λ_0 such that the individual with this labor disutility would be indifferent between working and not working in period 1:

$$\lambda_0 = u(w) - u(p) + \Delta \quad (2)$$

where Δ is the second-period wage differential $u(v_w) - u(v_n)$. Individuals with labor disutility higher than λ_0 will not work in period 1. Let $\bar{\lambda}_w = (\lambda_0(\Delta) - \lambda_1)/2$ and $\bar{\lambda}_n = (\lambda_u - \lambda_0(\Delta))/2$ be the average labor disutilities of workers and non-workers (assuming $\lambda_1 < \lambda_0 < \lambda_u$; see below). Evidently $\bar{\lambda}_w - \bar{\lambda}_n = (\lambda_u - \lambda_1)/2$, hence $\Delta > 0$; employers will in general penalize welfare reciprocity by making lower wage offers. With $v_w = Q - k\bar{\lambda}_w$ and $v_n = Q - k\bar{\lambda}_n$, the second-period wage penalty becomes

$$\Delta = u(Q - k\bar{\lambda}_w(\Delta)) - u(Q - k\bar{\lambda}_n(\Delta)) \quad (3)$$

Existence of equilibrium thus involves the question of whether there are fixed points in (3). Reviewing the formulas for $\bar{\lambda}_w$, $\bar{\lambda}_n$, and λ_0 , we can see that the utility function arguments in (3) are linear functions of Δ , and we can write

$$\Delta = u(A + B\Delta) - u(C + D\Delta) \quad (4)$$

where A , B , C , and D are parameters with $B < 0$ and $D > 0$. From the properties of $u(\cdot)$, $\Delta'(\Delta) < 0$ and $\Delta''(\Delta) > 0$; if $\Delta(0) > 0$, there will be a unique fixed point. $\Delta = 0$ implies the absence of any wage penalty from welfare receipt. Under such a circumstance everyone ignores the second-period effect, so that those with labor disutilities below $u(w) - u(p)$ will choose to work. The average labor disutility in this group will remain below that of the

welfare recipients, a fact that generates second-period wages such that $v_w > v_n$. This then implies $u(v_w) > u(v_n)$, which means $u(A) - u(C) > 0$. $\Delta(0)$ therefore exceeds 0; a unique fixed point exists.

This fixed point will provide an employer-employee equilibrium that makes welfare a credible signal whenever $\lambda_1 < \lambda_0 < \lambda_u$. In that case, the population sharing the productivity value Q will split between welfare and work, and the split will imply that the labor disutility of welfare recipients exceeds that of workers. As Q rises, however, the incentive to first-period work rises, and we cannot rule out the possibility that at some $Q' < Q_u$, the entire population of Q' individuals chooses to work. Here, welfare receipt cannot provide information about labor disutility because, obviously, there are no welfare recipients. Similarly, as Q falls, the incentive for first-period work falls, and we cannot rule out the possibility that for some $Q'' > Q_l$, the entire population of Q'' individuals chooses to claim welfare. Again, welfare provides no information. We thus have three equilibria, in one of which (when base productivity falls between Q' and Q'') welfare receipt provides information about the productivity of workers. In that equilibrium, welfare take-up will be less than 100 percent. ■

Given that there exists some base productivity Q_i with $Q'' < Q_i < Q'$, employer discrimination will exist as an equilibrium response to the signal provided by welfare recipiency. In such an equilibrium, some workers whose wages make them eligible for welfare choose instead to work, so as to avoid the penalty of lower second-period wages. This makes take-up rates less than 100 percent. Employer discrimination can have this effect whether or not the employers or employees care anything about the 'deservedness' of welfare

claims. Thus, contrary to Besley and Coate's (1992) claim, the fact that employer discrimination can make take-up rates low is no evidence of stigma, as defined here.

VI. Can Stigma Exist?

Now that we are aware of at least three sensible reasons why welfare dollars generate less utility than other dollars, stigma becomes an interesting puzzle. Stigma is a cost of welfare that is not related to individual guilt, hassles, or employer's use of information. As defined here, stigma happens only as a reaction to the violation of a social norm regarding who ought, and who ought not, to claim welfare. Though there are many ways of expressing this norm, the most natural way is to express it as a market wage. Thus, anyone whose market wage is very low (such as the disabled or the elderly), or whose wage is average but associated with high work-related expenses (such as single mothers with many young children), ought legitimately to claim welfare. Those with high wages ought not to claim welfare but rather work instead. As a reaction to norm violation, then, stigma can only happen when the welfare recipient A interacts with another person B under conditions of common knowledge regarding A's wage and A's status as a welfare recipient.

A. The puzzle of stigma. Let us imagine two people interacting under this common-knowledge condition. B observes that A's wage is higher than the normative wage, so B concludes that A is violating the social understanding of who ought to be working. The puzzle of stigma is, why would B act any differently toward A *merely* as a result of this conclusion? If we rule out as motivations sadism or any other personal satisfaction from hurting those who violate society's norms,⁴ B should not ordinarily care whether A is

working or not. If A is the customer and B is the grocer, it would make no sense for B to refuse to sell A an apple *merely* because A has violated a work norm.

The answer to the puzzle lies in the nature of the interaction. Simple market exchanges, for example, should never result in punishment for norm violations. This is because the piece of information that norm violation represents has no influence on the two parties' expectations about behavior. Whether or not there is a known norm violation, the grocer believes that the welfare recipient desires the apple at the given price, and the recipient believes that the grocer is willing to sacrifice the apple at that price. Norm violation by either party has no influence on these expectations and valuations, and hence it will not change prices, actions, or utilities.

B. A prisoner's dilemma justification. Norm violation can alter expectations in two more complex types of interactions, however. In prisoner's dilemma (PD) games, incentives to cooperate depend on expectations of future cooperation. Anything that changes these expectations can eliminate cooperation, lowering the payoffs to both parties. A norm violation, though completely unrelated to the interaction, can have this effect.

Suppose A and B are neighbors who interact by loaning house and garden tools to one another. The single-shot game involves A and B standing at their fence, A holding a hedge trimmer, B holding power saw, and both considering whether to loan the respective tool to the other. Taking A's perspective for a moment, there is always a risk that B will not return the tool, and in the single-shot game this risk is not affected by whether or not B loans his tool to A. For A, then, the strategy 'Loan tool' is strictly dominated by 'Do not loan tool.' The same holds for B. The only equilibrium involves no loans, and both parties are worse

off than if they had made the loans. In general, all neighborhood capital pool arrangements have this prisoner's dilemma character.

Now suppose that the game repeats indefinitely and A and B find themselves in a dynamic cooperative equilibrium. Suppose A and B are playing grim trigger strategies (loan tools until some breakdown condition occurs, e.g. the other fails to return a tool; then never loan another tool forever). Consider the following trigger condition: "The agreement breaks down if either party violates a work norm." If this trigger strategy is in equilibrium (something that has to be shown), then B will find it incentive-compatible to punish A for violating a work norm. If A violates the norm, A will expect B to pull the grim trigger and refuse to loan tools forevermore. With this expectation, A's utility-maximizing response is also to refuse to loan tools forevermore. B, fully understanding A's reasoning, now expects A not to loan tools anymore. And with this expectation, B's utility-maximizing response is to refuse loans. Both A and B fall into the no-cooperation equilibrium, which involves lower payoffs for both. Norm violation by can thus cause B to change his behavior in such a way that A's utility falls. Therefore, B can find it sensible to impose a stigma cost on a socially improper welfare claim.

C. A coordination justification. A second mechanism of rational stigma is the single-shot coordination (CD) game, in which two players have the incentive to match one another's actions but do not know what those actions will be. Schelling (1960) gives the example of two people in separate rooms who will be allowed to take home a share of \$1 million if they can independently write down the same division of spoils between them. Suppose neighbors A and B both want to drive to Flint, and have complete control over when they leave. If A

and B choose to leave at the same time, they can share the travel expenses and save themselves hundreds of dollars. In general, all scheduling and bargaining interactions can be analyzed as coordination problems like this one.

Schelling points out that there are many equilibria to this game; the two neighbors are equally happy leaving at any hour of the day, so long as both leave at the same time. The difficulty is in communicating to one another the time at which each should be ready to go. The start time must be common knowledge, and any common-knowledge start time is self-confirming as an equilibrium. Suppose both parties are aware that normally, in their culture, people begin a travel day at 9 AM. Then both will expect the other at 9 AM, and this expectation will be correct. The time "9 AM" is a focal point, outside the game, that induces a particular choice of coordination equilibrium.

A norm violation by one of the parties, that itself has nothing to do with the game, can induce a breakdown in the coordination equilibrium. Suppose A and B understand their strategies as being contingent on whether the other player has violated a work norm. That is, given the focal point at 9 AM, B's strategy is to prepare for travel at that time so long as he knows that his fellow-traveler has not violated a work norm. Now suppose that A has violated the work norm, and that the violation is common knowledge. If B's strategy is in equilibrium (something that has to be shown), A does not expect B to be ready for travel at 9 AM. A, having no idea when B will travel, but still hoping to travel with him, chooses some other time to leave. Fully understanding A's reasoning, and still hoping to travel with him, B now knows that A is going to choose some random time *other* than 9 AM. Therefore B certainly will not choose to leave at 9 AM. Moreover, B will also have to choose a

random time other than 9 AM, as no one time is any more likely than another given what A is doing. With both parties choosing random times, in all likelihood they will not meet for travel, and therefore the norm violation induces behavior by B that lowers utility for A. B voluntarily imposes stigma on A.

In both these stories, stigma has a very specific flavor. It does not involve personal shame; no one wags a finger at the undeserving welfare recipient; no one scolds them in public. There are no incidents whereby one party chooses to execute punishment. Rather, stigma comes from a long series of awkward moments. Meetings are missed; bargaining takes a bit more time; neighbors hesitate to talk; and shop-owners are reluctant to give credit. Aggregated over enough people and time periods, these moments can have a serious impact on the cost of claiming welfare. Stigma becomes a choking cloud of alienation. Understood this way, this paper's theory of stigma comes closer to the casual empirical facts about stigma (e.g. Goffman, 1963) than any other.

D. Common knowledge. All of the above assumes that A's status as a welfare recipient, as well as A's wage, are known to B. Indeed this information is common knowledge between them. The information conditions have the following influence on our reasoning: both the prisoner's dilemma and the coordination explanations for welfare stigma have greater power as the number of people who are aware of a norm violation rises. If norm violations cannot be detected, there can be no stigma as defined here. Stigma can only occur when both parties to a transaction become aware, somehow, that one of them has claimed welfare but has a high wage. Casual observation suggests, however, that such awareness can emerge in countless circumstances. In the prisoner's dilemma approach,

single-shot games are repeated indefinitely and this gives ample opportunity for the players to learn about one another. Similarly, one can imagine in coordination games that norm violations are revealed at random. Perhaps a Food Stamp coupon slips out of an expensive purse. In either case, one can easily imagine circumstances in which norm violations become common knowledge. I have chosen not to model the processes by which learning occurs, however. In the models that follow, it should be understood that the population of individuals who share common knowledge about one another's wages and welfare receipt status can in principle be a very small one.⁵

E. The norm. Furthermore, I do not consider the emergence or the existence of the norm itself. The norm is assumed to exist at the level ω . Anyone whose wage is above this level, yet who claims welfare instead of working, will be considered a *defector* from the norm. How such a norm gets established or changed is a coordination problem of a higher sort (Kreps, 1990).

The following sections prove that the prisoner's dilemma and coordination explanations of stigma are coherent.

VII. A Prisoner's Dilemma Model of Welfare Stigma.

Let the population consist of $N+1$ individuals who jointly observe one another's wages and work choices. Suppose individual utility is as follows:

$$u_i = u(y_i) - \lambda_i l_i + \gamma \sum_{j \neq i} \sum_{t=1}^{\infty} \delta^{t-1} R_{ijt} \quad (5)$$

where δ is a rate of time preference ($0 < \delta < 1$), γ is the utility weight placed on prisoner's dilemma payoffs, and R_{ijt} is the expected payoff (in utils) in period t of an

infinitely repeated game between individual i and some other individual j . Each stage of the game has the following payoff structure:

Entries: (Row payoff, Column payoff)		Column Player	
		Cooperate	Defect
Row Player	Cooperate	B, B	-1, 2B
	Defect	2B, -1	0, 0

The payoffs are symmetric; the roles played by i and j have no effect on outcomes.

Individual decisions follow a specific sequence. First, nature sets the norm, ω , which becomes common knowledge. Second, everyone chooses whether or not to work. Third, each pair of players repeats the above stage game indefinitely. (For simplicity, I will assume that players in one repeated PD do not observe actions in any other; once the game reaches the repeated PD stage, each pairing acts independently of the others.)

Let Φ be the set of stage-game pure strategies $\{C, D\}$. (I will ignore equilibria involving mixed strategies). Let $\Psi = \{0, 1\}$ be the set of possible labor strategies. Let $W = [w_1, p]$ be the set of possible values for ω . This assumes that no norm of welfare receipt can require that no one should apply for welfare, and that the norm of welfare eligibility is less than the legal eligibility line (otherwise there would be no stigma). Finally, at any stage in a single repeated PD game, the history h^t contains a) the norm, b) the labor choices of the two players, and c) all of their previous plays in the stage game. The set of all possible histories at time t is $H^t = W \times \Psi^2 \times \Phi^{2(t-1)}$. A strategy for that individual game is, first, a

map from W into Ψ , and second, a sequence of maps from H^t into Φ .

There are many equilibria in the game, including one in which labor choices are independent of both the announced norm and the subsequent repeated game, which in turn is also independent of labor supply choices. That is, the existence of a common-knowledge social norm is no guarantee of its enforcement; that norms are stated does not guarantee that they have social impact.

There is an equilibrium in which they do have social impact, however. The strategies in this equilibrium involve two kinds of status. *Defectors* are players with wages above ω who claim welfare benefits. All others are called *conformists*.

Theorem 2: Under the assumptions behind the repeated prisoner's dilemma game, there exists an equilibrium with a unique number of defectors, in which defection is costly, i.e. in which stigma exists.

Proof: The following strategies, if performed by all players, constitute an equilibrium:

1. a) If $w_i \leq \omega$, choose $l_i = 0$.

1. b) If $w_i > \omega$, choose $l_i = 1$ if and only if $\lambda_i < u(w_i) - u(p) + (N - M_i)\gamma B'$, where M_i is the expected number of defectors and B' is the discounted present value of a stream of payoffs B in the repeated PD.

2) If i and j are conformists, Player i plays {Tit for Tat} in the prisoner's dilemma. If either is a defector, Player i plays {Always Defect}. (The Tit for Tat strategy is to Cooperate in period 1, and thereafter play whatever the opponent played in the preceding

period.)

After being matched, two players observe their respective status as conformists and defectors. {Tit for Tat} and {Always Defect} being best replies to themselves (for low enough discount rates), matched conformists will optimize by following the former strategy, and all other pairs will optimize by following the latter. The discounted repeated game payoff for paired conformists is $B' = B/(1-\delta)$, for all other pairs, 0.

When choosing labor supply, players must consider the likely frequency of defection in the prisoner's dilemma. Suppose player i expects that M_i members of the population of N players choose to defect. The expected repeated-game payoff for a conformist is $(N - M_i)\gamma B'$.

Conformists include all low-wage individuals (i.e. $w < \omega$), and high-wage individuals who work. Low-wage people cannot be defectors, and therefore will receive conformist payoffs in the repeated game whether or not they claim welfare. Work will be optimal for a low-wage player i if

$$u(w_i) - \lambda_i + (N - M_i)\gamma B' > u(p) + (N - M_i)\gamma B' \quad (6)$$

Since $w_i < p$ (by $\omega < p$), and $\lambda_i > 0$, the choice $l_i = 0$ is indeed best. If $w_i > \omega$, failure to work results in defector status, with the consequence that neither conformists nor defectors will cooperate in the repeated game; payoffs there will be zero. Work will be best if

$$u(w_i) - \lambda_i + (N - M_i)\gamma B' > u(p) \quad (7)$$

which implies the condition (b) above.

If we require rational expectations on the part of all players, then all should share the

same expectation about the number of defectors, i.e. $M_i = M$, all i , and in equilibrium the actual number of defectors should equal the expectation. Thus it is necessary to show that there exists an M that makes defection optimal for exactly M high-wage workers.

The argument for this can be motivated through Figure 2. The norm ω has been placed low in the wage distribution; the poverty line ($p > \omega$) is slightly higher. The line λ_0 shows the work disutility required to induce welfare claims. The lower curve assumes no repeated PD effect, while the upper shows the critical λ for high-wage workers when expected PD payoffs are $(N - M)B'$, i.e. $\lambda_0 = \{\lambda_i: \lambda_i = u(w_i) - u(p) + (N - M)\gamma B'\}$. Given this ω and this M , individuals in region AE will claim welfare. Those in F are eligible for welfare but will not claim it; note that the take-up rate in the figure is realistic, about 70 percent. Those in B will claim welfare as eligibles, though they will be in violation of the social norm. Those in D not only violate the social norm, they also break the law because their wages exceed the legal poverty line. Individuals in region G avoid cheating the welfare system only because of the existence of stigma. Finally, those in region H work regardless of stigma and legal eligibility. The model allows a rich variety of welfare receipt behavior.

The number of defectors M is given by the area of region BD, which is determined by the position of λ_0 ; λ_0 is in turn a function of M , with $\lambda_0'(M) < 0$. Its area is

$$Z(M) = \int_{\omega}^{w'} \lambda_u - [u(x) - u(p) + (N-M)\gamma B'] dx \quad (8)$$

where w' is as defined in Section IV above. Thus equilibrium requires a value of M such that $M = Z(M)$; the defection population must have a size that elicits defection by exactly that many people.

Under mild conditions on the integral $Z(M)$, such an M exists and is unique. To see this, note that

$$Z'(M) = \int_{\omega}^{w'} B' dx = (w' - \omega) \gamma B' \quad (9)$$

by Leibniz's rule. Because $w' > \omega$, $Z' > 0$. Further, Z'' is approximately zero, though not identically zero because the limit of integration w' depends on M . Assume for the moment that Z is linear. Then if it can be shown that $Z(M)$ intersects the 45-degree line somewhere on the domain of M , it will be true that the intersection point is a unique solution. See Figure 3, which illustrates the requirements. K is the maximum number of defectors. If $Z(K) < K$ but $Z(0) > 0$, there will be exactly one point M^* such that $Z(M^*) = M^*$.

$Z(0)$ will exceed 0 if at least one person would find it optimal to defect when no one else is defecting. When no one defects, $M = 0$ and $\lambda_0 = \{\lambda_i; \lambda_i = u(w_i) - u(p) + N\gamma B'\}$. The person with highest labor disutility and lowest wage (above ω) will defect if

$$CI : \lambda_u > u(\omega) - u(p) + N\gamma B' \quad (10)$$

which means only that λ_0 must be below λ_u for some $\omega < w < w_u$. Facing the maximum sanction (punishment by all), the highest-disutility worker, if offered the minimum wage still subject to sanction, chooses to accept the sanction and avoid work. If the condition is not met, then the expectation $M = 0$ is self-confirming: when no one defects, no one wants to defect.

$Z(K)$ will be less than K if at least one person would find it optimal to work even if

no one else is working. When no one works, $M = K$ and $\lambda_0 = \{\lambda_i: \lambda_i = u(w_i) - u(p) + (N - K)\gamma B'\}$. The person with lowest labor disutility and highest wage will work if

$$C2 : \lambda_l < u(w_u) - u(p) + (N - K)\gamma B' \quad (11)$$

or, λ_0 must be above λ_l for some $\omega < w < w_u$. Facing the lowest possible sanction, the lowest-disutility worker, if offered the highest wage, will choose to avoid that sanction and work instead. If this condition is not met, the expectation $M = K$ is self-confirming: when everyone defects, everyone wants to defect.

C1 and C2 cannot both fail at the same time, as this would imply a λ_0 curve that is both above λ_u and below λ_l for all $\omega < w < w_u$. If one or the other fails, the unique equilibrium M is a polar value, 0 or K . With $Z(M)$ linear, the conditions C1 and C2 ensure that the slope of Z is positive but less than one, as depicted in Figure 3. The slope of Z indicates the impact of adding one person to the ranks of the defectors on the incentive to defect. The slope is positive because shifting one person from conformity to defection reduces the size of the conforming population, thus lowering the cost of defection. This in turn encourages further defection. If adding ten defectors induces further defections by fewer than ten other people, the adjustment process between M and $Z(M)$ will converge on an M^* between 0 and K . That M^* will be stable. If adding ten defectors induces further defections by more than 10 people, the adjustment process will force M^* to 0 or K . In Figure 3, $Z(M)$ would intersect the 45-degree line from below, and any intermediate value of M^* would be unstable. Thus the requirement that the right-hand side of (9) be positive but less than one ensures that the unique M^* is an intermediate value.

Whether that requirement is met or not, there will exist a unique M^* such that

defection is optimal for exactly M^* individuals, with $0 \leq M^* \leq K$. ■

Because the most interesting outcome is the one that allows for the richest and most realistic variety of behavior, further analysis of the model will assume conditions C1 and C2 hold. The slope term $(w' - \omega)\gamma B'$ will be assumed to fall between 0 and 1, and M^* will be assumed to fall between 0 and K .

Finally, note that if Z is non-linear, the above argument will still hold if the second derivative of Z is sufficiently close to zero. Referring to Figure 2, the first derivative is a one-unit slice along the entire right hand boundary of region BD, holding w' constant. The second derivative is the triangle added to this slice when w' shifts to the right. It seems plausible to assume the triangle is small relative to the slice, and that therefore Z'' is near zero. Section IX considers the effect of small changes in the vicinity of M^* , and for these purposes the $Z(M)$ function will be assumed to be linear.

VIII. A Coordination Model of Welfare Stigma.

Now consider stigma as the outcome of coordination failures. Let there be $N+1$ individuals who jointly observe one another's wage and work status, and let utility have this form:

$$u_i = u(y_i) - \lambda_i l_i + \gamma \sum_{j \neq i} BI(a_{ij} = a_{ji}) \quad (12)$$

where $I(a = b)$ is an indicator function, equaling 1 if $a = b$ and 0 otherwise. a_{ki} is an action of individual k with respect to individual i , and B is the benefit obtained by both individuals

if their actions match. Recalling the bargaining example, assume that the action space is a bounded finite set; actions are drawn from the set $A = [a_l, a_h]$ on the integers. The practical impact of this assumption is that the probability of a match between two randomly-chosen actions is nearly zero. (One could allow some benefit to be obtained from 'near-misses' but this would complicate the model without adding anything significant.)

The game proceeds as follows. First, nature sets the norm ω . Second, all players choose whether or not to work. Third, each player simultaneously chooses an action with respect to each of the other N players. Payoffs are then received.

$\Psi = \{0,1\}$ being the set of all labor-choice possibilities, and W being the set of possible norms, a strategy in this game is, first, a map from W into Ψ , and second, N maps from $W \times \Psi^{N+1}$ into A .

The second stage is an example from the class of n -player pure-coordination games, from which we can derive some general characteristics of any outcomes. First, any action strategy $a_{kl} = z$, all k and l , will be a best response to itself. Given that all others play z , no one player has the incentive to play anything else. Because any z in A can serve in this role, the game has many equilibria. Second, if the set of pure-strategy equilibria is finite, a mixed-strategy equilibrium exists in which all players play each action in the pure equilibrium set with equal probability (given that coordination payoffs are independent of the actions). Mixed-strategy payoffs will be considerably lower than coordinated payoffs. Finally, anything that may serve as a focal point in the game will induce an equilibrium (Schelling, 1960). Were a leader to make a common-knowledge suggestion "Play action z " then all players would expect the action z from others and would find it optimal to play z as

well.

I will assume that a focal point in the CD game does exist, denoted α . Player i 's strategies will involve either playing α against her opponent, hoping to receive the benefit B of coordination, or playing something other than α . If her choice is $a' \neq \alpha$, she receives a zero payoff if the opponent has played α , or any other action other than a' . Indeed, not knowing what the opponent's action is, player i can do no better than to choose a mixed strategy that assigns equal probability to all actions in A except α . The opponent's best response to such a strategy is also to randomize. The probability of any matched pair of actions from the set A (excluding α) is $1/(a_u - a_l)^2$, and there are $(a_u - a_l)$ possible matches; the expected payoff from the mixed strategies is therefore $\varepsilon = B/(a_u - a_l)$. For any sizeable strategy set A , this payoff will be considerably below B .

As above, the strategies use two kinds of status, *defectors* and *conformists*.

Theorem 3: Under the assumptions behind the coordination game, there exists an equilibrium with a unique number of defectors, in which defection is costly, i.e. in which stigma exists.

Proof: the following strategies (a) - (c) will constitute an equilibrium if followed by all players i :

a) If $w_i \leq \omega$, choose $l_i = 0$.

b) If $w_i > \omega$, choose $l_i = 1$ if and only if $\lambda_i < u(w_i) - u(p) + (N - M_i)\gamma(B - \varepsilon)$,

where M_i is the expected number of defectors.

c) If both i and j are conformists, play α in the coordination stage. Otherwise play an action $a' \neq \alpha$ with $\Pr(a') = 1/(a_u - a_l)$, all a' .

At the coordination stage, α is a best response to α , so that once two players are aware that both are conformists, α becomes the optimal play for both. Conformists meeting defectors will expect the mixed strategy, to which the mixed strategy is itself the best reply (as argued above). Thus given that expectations of behavior are formed with respect to defection status, the strategies in (c) are optimal. They imply payoffs of $(N - M_i)\gamma B + M_i\gamma\epsilon$ for conformists and $N\gamma\epsilon$ for defectors.

It remains to show that these payoffs induce welfare claims from all with wages below ω , work under condition (b) for all with wages above ω , and that expected defections M_i are rational. However, these requirements are identical to those of the prisoner's dilemma model (replacing $B - \epsilon$ with B'). Low-wage individuals are never defectors and receive the benefits of coordination regardless of their work choice. For them, welfare is indeed optimal because

$$u(w_i) - \lambda_i + (N - M_i)\gamma B + M_i\gamma\epsilon < u(p) + (N - M_i)\gamma B + M_i\gamma\epsilon \quad (13)$$

For high-wage individuals, coordination benefits are $(N - M_i)\gamma B + M_i\gamma\epsilon$ if they work but only $N\gamma\epsilon$ if they do not. Thus work is optimal only if

$$u(w_i) - \lambda_i + (N - M_i)\gamma B + M_i\gamma\epsilon > u(p) + N\gamma\epsilon \quad (14)$$

which is the same condition as given in part (b). Lastly, these strategies will involve rational expectations if they are based on a value of M that is shared by all players, and induces exactly M defections. Section VII demonstrated that a unique value exists. Analogous

conditions indicate when this value of M implies universal defection, universal conformity, or something in between. ■

IX. Comparative Statics.

The prisoner's dilemma and coordination models of stigma provide information on three issues of practical importance. First, how do caseloads and defection rates respond to changes in the generosity of the welfare system? Second, how do they respond to changes in norms? Third, what is the practical significance of potentially unstable equilibrium dynamics? The first issue is important because it addresses the Victorian concern that welfare generosity would erode the power of work norms. The second is important because policymakers today have grown more interested in the strategy of changing work norms, rather than program parameters, in the fight against poverty. The third is important because it provides intuition about historical episodes, such as the AFDC explosion in 1967-1972.

The first two issues will be considered in a limited context, under the assumptions that a) the equilibria under consideration have unique M^* values that lie between 0 and K , and b) $Z''(M) = 0$. The third issue requires that these conditions do not hold. The focus is on the prisoner's dilemma model, though results would be exactly the same under the coordination failure approach.

A. Program parameter effects. Because the welfare system here is a highly stylized one, its generosity can be summarized in the single parameter p . In a model without stigma (see Figure 1), the caseload is

$$C = (\lambda_u - \lambda_l)p + \int_p^{w'} \lambda_u - [u(x) - u(p)] dx \quad (15)$$

Differentiating and re-arranging, this implies

$$\frac{dC}{dp} = (\lambda_u - \lambda_l) + (w' - p)u'(p) \quad (16)$$

When the program becomes more generous, there are two effects. First, more people are eligible and a slice of the population along the right-hand boundary of AD in Figure 1 is added to the welfare rolls. This effect is given by $\lambda_u - \lambda_l$. Second, greater generosity increases the incentive of higher-wage individuals to avoid work. This appears as a downward shift in the λ_0 curve in Figure 1, and adds a slice along the lower right boundary of region B. The effect is given in (16) by $(w' - p)U'(p)$. Both effects are positive, so we can conclude that increasing welfare generosity increases caseloads.

How is this story affected by the presence of stigma? First, note that the caseload is now $C = (\lambda_u - \lambda_l)\omega + M$ (see Figure 2): everyone with wages below ω claims welfare, and to this we add the population of defectors, M . The first group is unaffected by any change in p , so that changes in the caseload can only be produced by changes in the defector population. Its size is given by Z :

$$Z(M) = \int_{\omega}^{w'} \lambda_u - [u(x) - u(p) + (N - M)\gamma B'] dx \quad (17)$$

Differentiation yields

$$\frac{dM}{dp} = \frac{u'(\omega)(w' - \omega)}{1 - (w' - \omega)\gamma B'} \quad (18)$$

This differs in important ways from the no-stigma case. First, there is no wholesale addition to the welfare rolls; raising p does not add to regions AE in Figure 2. Second, the increase in work disincentives is more pronounced. The downward shift effect on the λ_0 curve, given by $u'(\omega)(w' - \omega)$, is increased because the denominator term $1 - (w' - \omega)B'$ is positive but less than 1 (from equation 9). Stigma enhances the reaction of caseloads to program changes, because it amplifies any increase in the incentive to defect. Stigma creates a positive social feedback mechanism that amplifies the effect of any policy that increases caseloads.

Stigma also responds to program generosity in the way the Victorians feared. Because the right-hand term in Equation (18) is positive, increasing the generosity of the welfare system increases defection. Welfare apparently does 'ruin the poor' in the Victorian sense. It induces more people to ignore self-sufficiency norms.

B. Norm effects. From (17), differentiation with respect to ω yields

$$\frac{dM}{d\omega} = - \frac{\lambda_u - \lambda_0(\omega)}{1 - (w' - \omega)\gamma B'} \quad (19)$$

where $\lambda_0(\omega) = u(\omega) - u(p) + (N-M)\gamma B'$. The first element gives the level response of defection to a unit change in ω ; in Figure 2, it can be seen as the impact of slicing off the left-hand boundary of region BD. This reduces the defection population and increases the conformist population, which in turn raises the cost of defection and further lowers the number of defectors. This add-on effect is produced by an upward shift in λ_0 , which slices off the lower-right boundary of BD. Successive movement of ω rightward induces successive

upward shifts in λ_0 , until defection disappears. Thus, one could have advised the Victorians that they could have encouraged more people to respect their norms of self-sufficiency by relaxing those norms.

Note further that relaxing norms may or may not add to the caseload. Moving ω to the right adds a slice along the right-hand boundary of AE. The gain in A is just a transfer from B, however. The caseload rises only if the addition to E exceeds the lower-right boundary loss of BD.

C. The Unstable Equilibrium Case. Now consider the implications of an integral $Z(M)$ with slope greater than 1. This is not unrealistic by any means. It would be the case, for example, if the payoffs to cooperation and coordination B were suitably large. In the discussion of Figure 3, it was noted that any $M^* = Z(M^*)$ produced by a steeply sloped- Z curve would be unstable, in the sense that any deviation from that value would induce magnified responses in the defector population, until one or the other polar value was reached. Imagine that society happens to be at the all-defect equilibrium (the upper right-hand corner of the box in Figure 3). How could one induce society to go to the all-conform equilibrium? It would be sufficient to move the upward-sloping $Z(.)$ curve so that it intersects the 45-degree line to the *right* of K. In that case, the adjustment dynamics beginning at any point between 0 and K, including the all-defect point, point to the left. Society would instantly move to the all-conform equilibrium. In practical terms, this strategy would involve doing anything that move M^* to the right. Note, however, that upward-sloping $Z(.)$ implies that the signs of the comparative static effects in (18) and (19) are reversed. Now $dM/dp < 0$ and $dM/d\omega > 0$. Thus to shift M^* to the right it is, as before,

necessary to either lower benefits or raise the norm.

The Victorian dilemma was a legitimate one. They correctly feared that being kinder to the poor would cause more of them to defect from social norms of self-sufficiency. The only way to maintain adherence to norms was to be less kind, or to make the norms less harsh by expanding the notion of 'deserving' to include higher-wage individuals. In the end, the inheritors of the Victorian legacy chose the second option. The Commission on the Poor Laws, reporting in 1909, still laid great emphasis on the distinction between 'deserving' and 'undeserving' poor. By 1919, however, these notions were dead (Levine, 1988, p. 216). Instead, virtually all claimants were assumed to be truly needy. Evidently the norm of true need had shifted upward during the Great War.

X. Summary of Findings.

What is the relevance of this theoretical story for current social policy debates?

First, the models presented here show that the Victorian dilemma may well have been real. Social norms of work and self-sufficiency can indeed be eroded by the generosity of the welfare system. The effect is intuitive, and the prior literature on welfare stigma and social effects has long hinted that it exists, but this paper has proven its existence in equilibrium under very general assumptions.

Second, the paper has shown that the trade-off between work norms and welfare generosity is not necessary. There are several ways of explaining the fact that public assistance purchases less utility than other income. In many explanations, work norms play no role. Thus, empirical findings that take-up rates fall considerably below 100 percent do

not constitute compelling evidence that welfare stigma exists.

Third, the paper has explored the options facing decisionmakers hoping to help the poor. Increasing the generosity of the welfare system increases the caseload and erodes work norms. Making work norms less harsh also increases the caseload, but it strengthens work norms. At this writing, the strategy of many US states seems to be to make work norms more strict. The research here suggests that this will decrease the caseload but increase the number of 'undeserving' welfare recipients.

Fourth, the paper has shown that the power of norms depends critically on the degree to which individuals interact with one another in particular types of situations: coordination problems and repeated prisoner's dilemmas. Both interaction types are more commonly found in the life of the community than in the purely private sphere. Thus one way to counteract any feared erosion of norms, or to build norms where none exist, would be to facilitate greater reliance on the local community. One could delegate coordination and repeated prisoner's dilemma problems (e.g. crime-fighting, welfare, infrastructure, and amenities) down to neighborhood decision-making bodies. Doing so would increase the reliance of individuals on one another, and thereby make social norms more important in the behavior of all.

Endnotes

1. The Victorians' fears that welfare ruins the poor can be interpreted in three ways. First, it might be a simple recognition of the now well-understood work disincentive effects of means-tested transfers. Second, it might be a conjecture about preference formation: giving welfare to the parents increases the disutility of work in the children. Third, it might be a conjecture about work norms: giving too much welfare will lessen its stigma, diluting the effectiveness of social norms of self-sufficiency. The first interpretation does not capture the essence of the 19th century argument, however, which almost always stressed the effect of alms-giving on people other than the recipient. The second interpretation also is inaccurate in ignoring the immediacy of the commentators' alarm, in that the sea of vagrants would come right away, not in a few generations. This paper will consider the third interpretation, that increases in welfare benefits have an immediate effect on welfare stigma.
2. There are empirical studies that might be interpreted as evidence of stigma as understood here. The literature on neighborhood and peer effects at times seems to suggest that the society around us has some influence on what we do. Unfortunately, most of this research suffers from an identification problem (Manski, 1993), namely, how can we claim to estimate the average behavior of a group of individuals (which is the definition of mean regression) when the average behavior of those individuals is one of the right-hand-side variables? It is not surprising that the literature on these effects has produced only mixed results (Evans Oates and Schwab, 1992; Datcher, 1982; Corcoran et al., 1989; Case and Katz, 1991; see An, Haveman and Wolfe, 1993 for strong evidence of family effects; see Jencks and Mayer, 1990 for a review).
3. Obviously I am ignoring the government budget constraint. It would be trivial, though, to extend the model by arbitrarily adding to the distribution of wages at the upper end, creating a large population of individuals who under no circumstances would prefer welfare to work. These individuals could then be assumed to bear any tax burden for those who do choose welfare.
4. I am assuming B does not gain any direct pleasure or pain from attempting to change A's behavior. In Besley and Coate's taxpayer resentment model, it is implicit that taxpayers enjoy imposing some kind of utility cost on welfare recipients. For if the taxpayer wishes only to bring his own tax bill down, punishing one recipient cannot possibly be worth the cost. The model ignores the collective action problem involved in getting many taxpayers to impose costs on many recipients. My hope is to overcome this problem by finding incentive-compatible punishment schemes. Unfortunately, resentment over the tax bill cannot be one of them.
5. Another possibility, explored by Kandori (1992), would allow individuals to label one another. This seems too unrealistic.

References

- Aaron, Henry J., Thomas E. Mann, and Timothy Taylor, eds. (1993), *Values and Public Policy*, Washington: The Brookings Institution.
- Akerlof, George A. (1980), "A Theory of Social Custom, of Which Unemployment May Be One Consequence," *Quarterly Journal of Economics* **94**, 749-95.
- An, Chong-Bum, Robert H. Haveman, and Barbara Wolfe (1993), "Teen Out-of-Wedlock Births and AFDC Receipt: The Role of Childhood Events and Economic Circumstances," *Review of Economics and Statistics* **75**(2), May, 195-208.
- Besley, Timothy and Stephen Coate (1992), "Understanding Welfare Stigma: Taxpayer Resentment and Statistical Discrimination," *Journal of Public Economics* **48**(2), July, 165-83.
- Bikhchandani, Sushil, David Hirshleifer, and Ivo Welch (1992), "A Theory of Fads, Fashion, Custom, and Cultural Change as Informational Cascades," *Journal of Political Economy* **100**(5), pp. 992-1026.
- Binmore, Ken (1994), *Game Theory and the Social Contract, Volume 1: Playing Fair*, Cambridge: MIT Press.
- Blank, Rebecca M. and Patricia Ruggles (1993), "When Do Women Use AFDC and Food Stamps? The Dynamics of Eligibility vs. Participation," NBER Working Paper No. 4429, August.
- Calvert, Randall L. (1992), "Leadership and Its Basis in Problems of Social Coordination," *International Political Science Review*, **13**(1), pp. 7-24.
- Calvert, Randall L. and Jeffrey S. Banks (1993), "Communication and Efficiency in Coordination Games," University of Rochester.
- Case, Anne C., and Lawrence F. Katz (1991), "The Company You Keep: The Effects of Family and Neighborhood on Disadvantaged Youths," NBER Working Paper No. 3705.
- Clarke, George R. G., and Robert P. Strauss (1994), "Children as Income-Producing Assets: The Case of Teen Illegitimacy and Government Transfers," Rochester Center for Economic Research Working Paper No. 389.
- Coleman, James S. (1990), *Foundations of Social Theory*, Cambridge: Harvard University Press.
- Corbett, Thomas (1994), "Changing the Culture of Welfare," *Focus* **16**(2), Summer-Fall, Madison: Institute For Research on Poverty.
- Corcoran, Mary, Roger Gordon, Deborah Laren, and Gary Solon (1989), "Effects of Family and Community Background on Men's Economic Status," NBER Working Paper No. 2896.
- Cowell, Frank A. (1990), *Cheating the Government: The Economics of Tax Evasion*, Cambridge: MIT Press.
- Crawford, Vincent P. and Hans Haller (1990), "Learning How to Cooperate: Optimal Play in Repeated Coordination Games," *Econometrica* **58**(3), May, 571-95.
- Datcher, Linda (1982), "Effects of Community and Family Background on Achievement," *Review of Economics and Statistics* **64**, pp 32-41.
- De Marzo, Peter M. (1992), "Coalitions, Leadership and Social Norms: The Power of Suggestion in Games," *Games and Economic Behavior* **4**(1), January, 72-100.

- Evans, William N., Wallace E. Oates, and Robert M. Schwab (1992), "Measuring Peer Group Effects: A Study of Teenage Behavior," *Journal of Political Economy* 100(5), October, pp. 966-91.
- Farrell, Joseph (1987), "Cheap Talk, Coordination, and Entry," *Rand Journal of Economics* 18, 34-39.
- _____ (1993), "Meaning and Credibility in Cheap-Talk Games," *Games and Economic Behavior* 5, 514-531.
- _____ and Gibbons, Robert (1989), "Cheap Talk Can Matter in Bargaining," *Journal of Economic Theory* 48, 221-37.
- Fraker, Thomas and Robert Moffitt (1988), "The Effects of Food Stamps on Labor Supply: A Bivariate Selection Model," *Journal of Public Economics* 35(1), February, 25-56.
- Goffman, Erving (1963), "Stigma: Notes on the Management of Spoiled Identity," Englewood Cliffs, New Jersey: Prentice-Hall.
- Gordon, James P.F. (1989), "Individual Morality and Reputation Costs as Deterrents to Tax Evasion," *European Economic Review* 33(4), April, 797-805.
- Greif, Avner (1992), "Cultural Beliefs and the Organization of Society," mimeo.
- Hardin, Russell (1990), "The Social Evolution of Cooperation," in K.S. Cook and M. Levi (eds.), *The Limits of Rationality*, Chicago: University of Chicago Press.
- Himmelfarb, Gertrude (1991), *Poverty and Compassion: The Moral Imagination of the Late Victorians*, New York: Vintage Books.
- Hoynes, Hilary and Thomas MaCurdy (1994), "Has the Decline in Benefits Shortened Welfare Spells?" *American Economic Review*, 84(2), May, 43-48.
- Jencks, Christopher and Susan E. Mayer (1990), "The Social Consequences of Growing Up in a Poor Neighborhood: A Review," in L. Lynn and M. McGeary (eds.), *Inner-City Poverty in the United States*, Washington, DC: National Academy Press, pp. 111-187.
- Johnson, James (1993), "Is Talk Really Cheap? Prompting Conversation Between Critical Theory and Rational Choice," *American Political Science Review* 87(1), 74-86.
- Jones, Stephen R. G. (1984), *The Economics of Conformism*, Oxford: Basil Blackwell, 1984.
- Kane, Thomas J. and Mary Jo Bane (1994), "The Context for Welfare Reform," in Mary Jo Bane and David T. Ellwood (eds.), *Welfare Realities*, Cambridge: Harvard University Press, pp. 1-27.
- Kandel, Eugene and Edward P. Lazear (1992), "Peer Pressure and Partnerships," *Journal of Political Economy* 100(4), August, 801-817.
- Kandori, Michihiro (1992), "Social Norms and Community Enforcement," *Review of Economic Studies* 59(1), January, 63-80.
- Kreps, David M. (1990), "Corporate Culture and Economic Theory," in James E. Alt and Kenneth Shepsle (eds.), *Perspectives on Positive Political Economy*, Cambridge: Cambridge University Press, 90-143.
- Levine, Daniel (1988), *Poverty and Society: The Growth of the American Welfare State in International Comparison*, New Brunswick: Rutgers University Press.
- Lindbeck, Assar (1994), "Welfare State Disincentives With Endogenous Habits and Norms," mimeo, University of Stockholm, December.
- Manski, Charles F. (1993), "Identification of Endogenous Social Effects: The Reflection Problem," *Review of Economic Studies* 60(3), 531-42.

- Matsuyama, Kiminori (1992), "Custom Versus Fashion: Path-Dependence and Limit Cycles in a Random Matching Game," Hoover Institution Domestic Studies Program, Stanford University, Working Papers in Economics E-92-11.
- Moffitt, Robert (1983), "An Economic Model of Stigma," *American Economic Review* 73(5), December, 1023-35.
- _____ (1987), "Historical Growth in Participation in Aid to Families With Dependent Children: Was There a Structural Shift?", *Journal of Post-Keynesian Economics* 9(3), Spring, pp. 347-63
- Putnam, Robert D. (1995), "Bowling Alone: Democracy In America at the End of the Twentieth Century," mimeo, Harvard University.
- Rank, Mark Robert (1994), *Living on the Edge: The Realities of Welfare in America*, New York: Columbia University Press.
- Ranney, Christine K. and John E. Kushman (1987), "Cash Equivalence, Welfare Stigma, and Food Stamps," *Southern Economic Journal* 53(4), April, 1011-27.
- Schelling, Thomas C. (1960), *The Strategy of Conflict*, London: Oxford University Press.
- Schotter, Andrew (1981), *The Economic Theory of Social Institutions*, Cambridge: Cambridge University Press.
- Strauss, Robert P. (1977), "Information and Participation in a Public Transfer Program," *Journal of Public Economics* 8, pp. 385-96.
- Weisbrod, Burton A. (1970), "On the Stigma Effect and the Demand for Welfare Programs: A Theoretical Note," *Institute for Research on Poverty Discussion Paper No. 82-70*, Madison, Wisconsin.

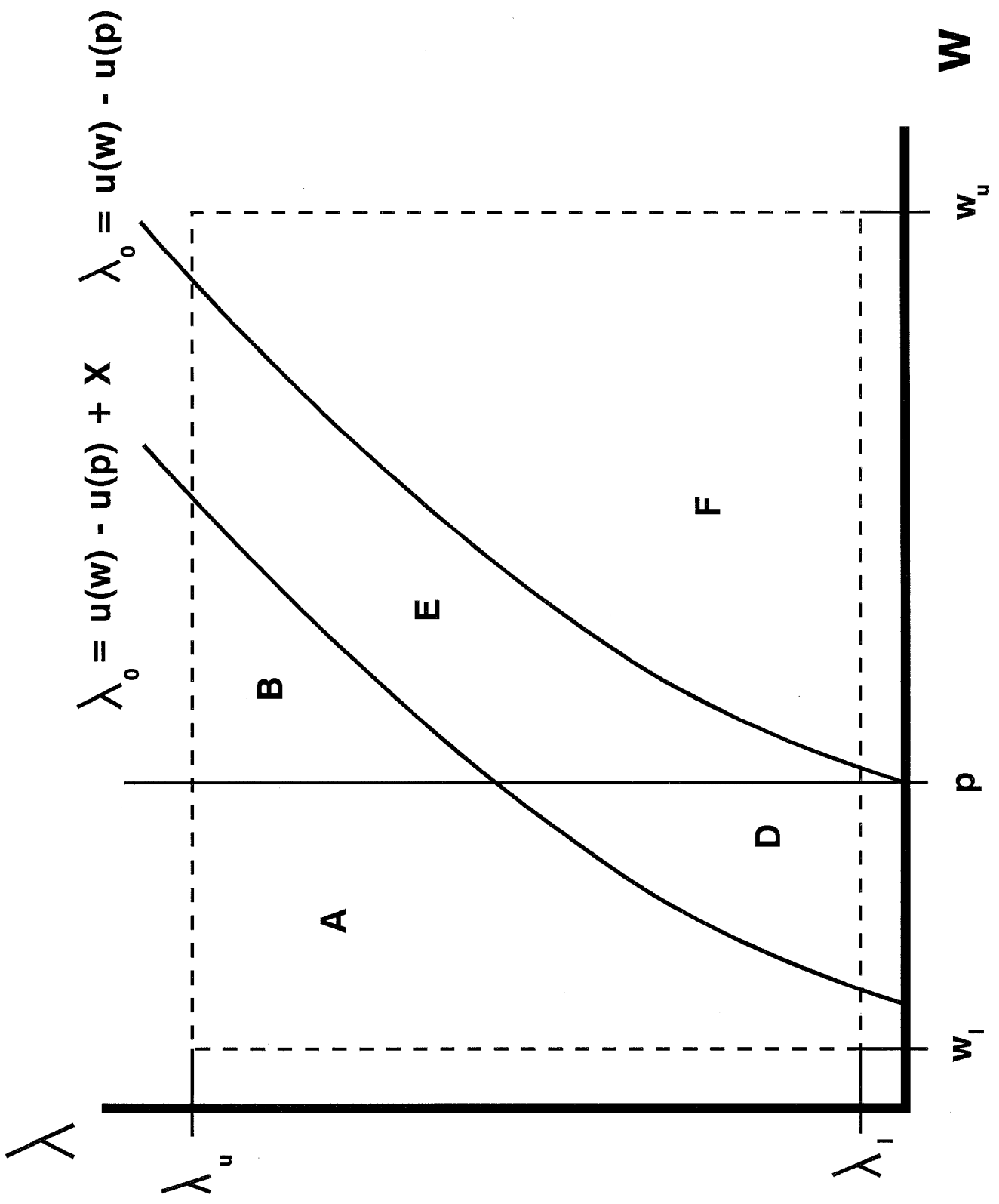


Figure 1

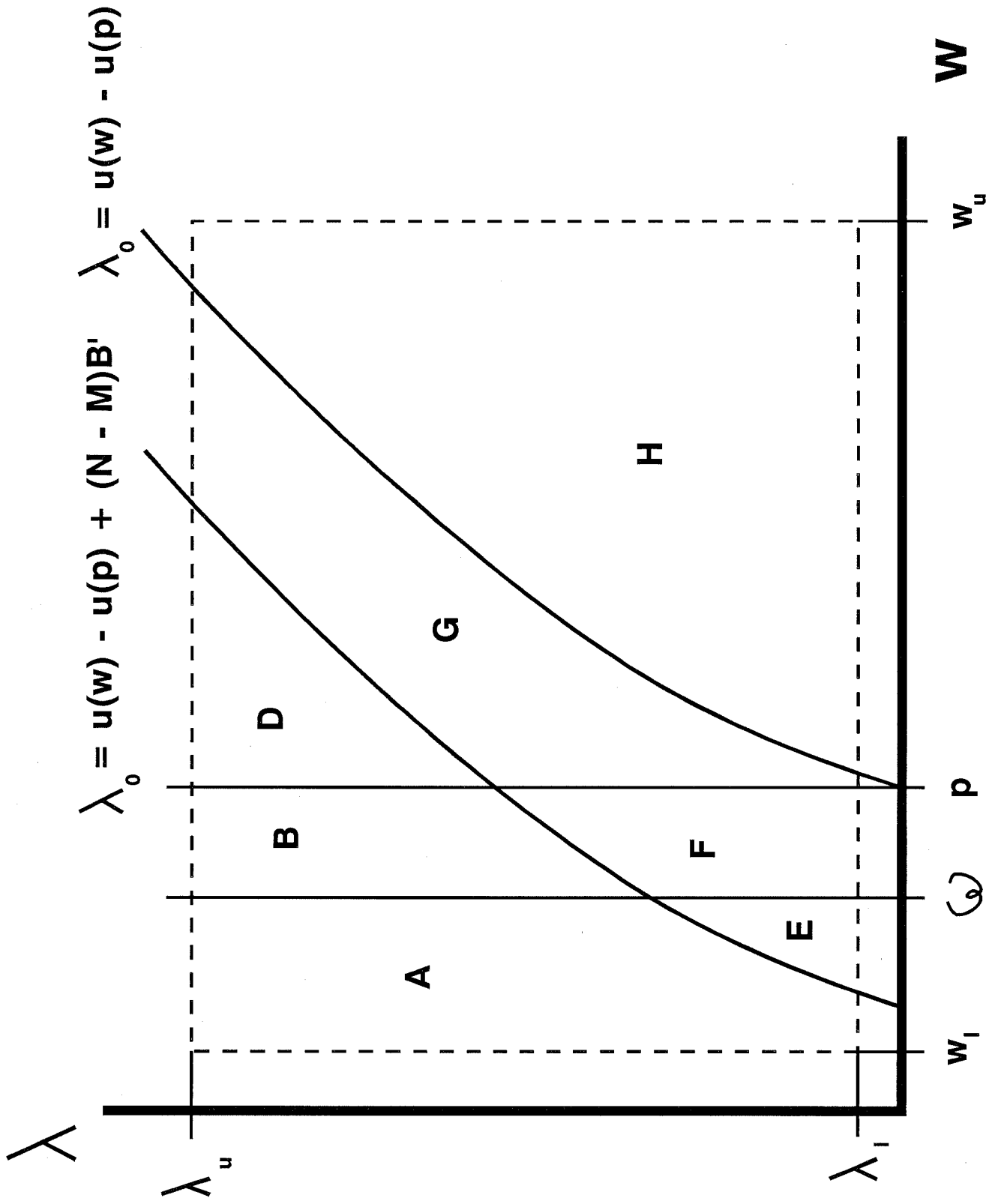


Figure 2

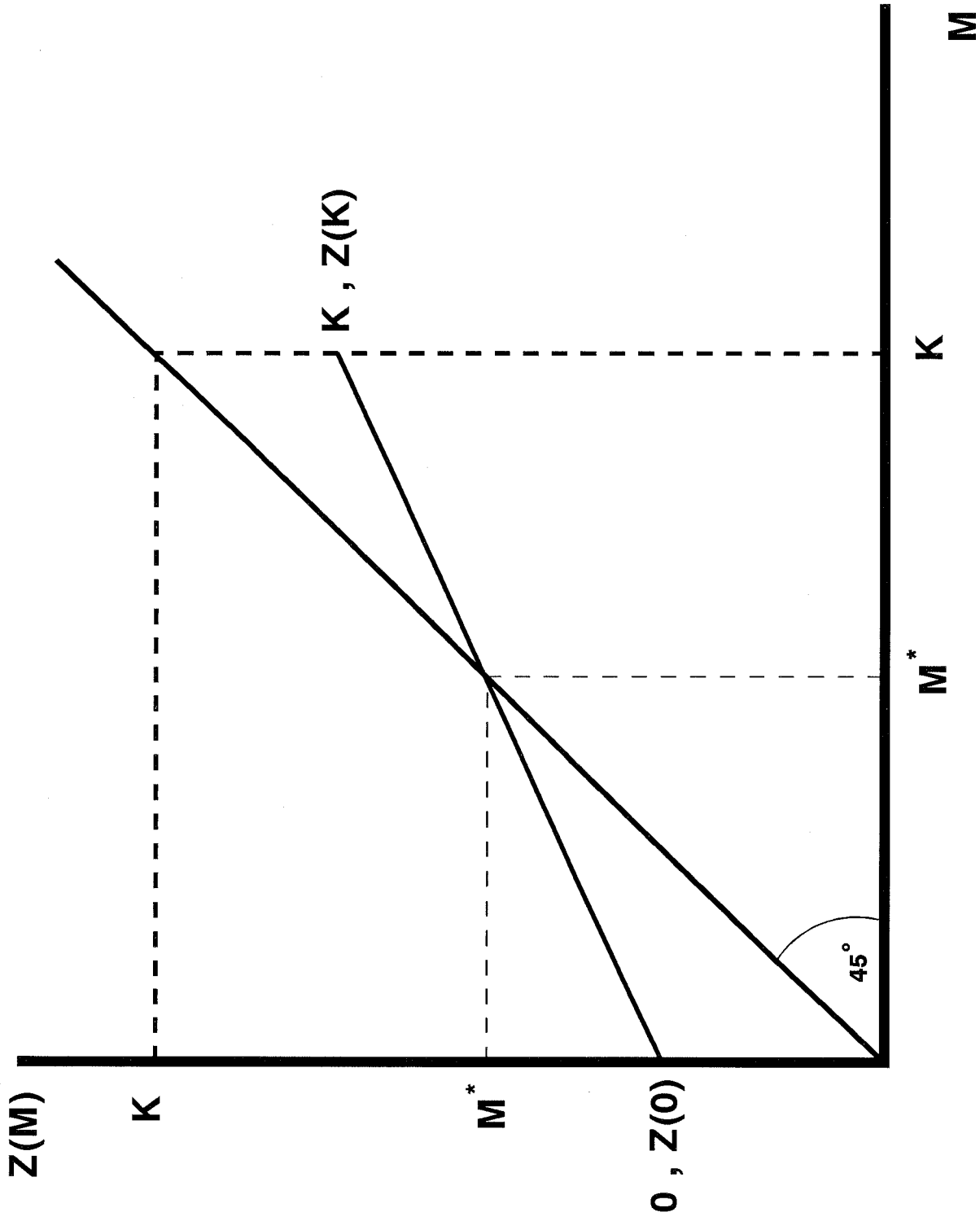


Figure 3

**W. ALLEN WALLIS INSTITUTE OF POLITICAL ECONOMY
1994-95 Working Paper Series**

**University of Rochester
107 Harkness Hall
Rochester, NY 14627**

- WP# 1 "What Constitutions Promote Capital Accumulation? A Political-Economy Approach,"
Krusell, Per and Ríos-Rull, José-Victor, December 1994.
- WP# 2 "Rationalizing School Spending: Efficiency, Externalities, and Equity, and their Connection to Rising Costs,"
Hanushek, Eric A., January 1995.
- WP# 3 "Interpreting Recent Research on Schooling in Developing Countries,"
Hanushek, Eric A., January 1995.
- WP# 4 "Explaining Social Order: Internalization, External Enforcement, or Equilibrium?"
Calvert, Randall, March 1995.
- WP# 5 "Does Welfare Ruin the Poor?,"
Bird, Edward J., June 1995.

W. Allen Wallis Institute for Political Economy

Working Paper Series

To order copies of the working papers, complete the attached invoice and return to:

Mrs. Terry Fisher
W. Allen Wallis Institute of Political Economy
107 Harkness Hall
University of Rochester
Rochester, NY 14627.

Three (3) papers per year will be provided free of charge as requested below. Each additional paper will require a \$5.00 service fee which must be enclosed with your order.

An invoice is provided below in order that you may request payment from your institution as necessary. Please make your check payable to the W. Allen Wallis Institute of Political Economy.

OFFICIAL INVOICE

Requestor's Name:

Requestor's Address:

Please send me the following papers free of charge:
(Limit: 3 free per year)

WP# _____ WP# _____ WP# _____

I understand there is a \$5.00 fee for each additional paper.
Enclosed is my check or money order in the amount of \$ _____.
Please send me the following papers.

WP# _____ WP# _____ WP# _____
WP# _____ WP# _____ WP# _____
WP# _____ WP# _____ WP# _____
WP# _____ WP# _____ WP# _____