Income Maintenance Programs and Multidimensional Screening*

Joel Shapiro†
Universitat Pompeu Fabra

April 9, 2001

Abstract

This paper examines properties of optimal poverty assistance programs under different informational environments using an income maintenance framework. To that end, we make both the income generating ability and the disutility of labor of individuals unobservable, and compare the resulting benefit schedules with those of programs found in the United States since Welfare Reform (1996). We find that optimal programs closely resemble a Negative Income Tax with a Benefit Reduction Rate that depends on the distribution of population characteristics. A policy of workfare (unpaid public sector work) is inefficient when disutility of labor is unobservable, but minimum work requirements (for paid work) may be used in that same environment. The distortions to work incentives and the presence of minimum work requirements depend on the observability and relative importance of the population's characteristics.

JEL Classifications: D82,H21,H53

Keywords: welfare programs, optimal taxation, multidimensional screening

*I thank Patrick Bolton, Harvey Rosen, and Robert Shimer for excellent comments and suggestions. I have also benefitted from discussions with Catherine Brown, Ernst Schaumburg, Diane Whitmore, and seminar participants at various universities. Financial support from the Center of Domestic and Comparative Policy Studies is appreciated.

†Contact: Joel Shapiro, Departament D'Economia i Empresa, Universitat Pompeu Fabra, Ramon Trias Fargas 25-27, 08005 Barcelona SPAIN. Email: joel.shapiro@econ.upf.es

1 Introduction

Understanding how the design of poverty assistance programs accomplishes stated goals has gained new importance since the overhaul of the welfare system in the United States in 1996. Currently, under the Temporary Assistance for Needy Families (TANF) program, states have been given mandates to experiment with programs that reduce poverty and decrease welfare rolls by promoting work. However, the goals of reducing poverty and promoting work may not be completely aligned. There is evidence that due to increasing wage inequality, low skill workers may not be able to earn enough to support their families at a minimal level[1]. There is also a growing body of evidence that as the welfare rolls diminish, the participants remaining have a combination of low ability, responsibilities at home, and other factors that make them increasingly difficult to move into employment[2]. This raises the question of how states should design programs to address populations with these characteristics.

Besley and Coate (1992, 1995) introduce a new type of static model to the analysis of welfare programs. They employ the tools of mechanism design to solve for optimal income maintenance programs. Such programs incorporate policy goals of reducing poverty and decreasing program size by ensuring that all participants are above a minimum income level and that the cost of the program is minimized. By bringing the design issue into a general optimization framework, some new policy questions have been addressed. Specifically, Besley and Coate characterize such programs when income generating ability is unobservable, and describe the trade-off of introducing workfare into the program.

This paper substantially extends the Besley and Coate framework in order to examine significant issues in program design. The main question we address is how income maintenance programs depend on the informational environment in which they will be used. We investigate the implications of having both income generating ability *and* disutility of labor as unobservable characteristics of individuals. Disutility of labor is clearly an important factor in individuals' labor decisions, especially among welfare populations. It can arise from several factors, such as the need to care for children, physical handicaps, living far away from work, or strict preference for leisure. A survey by The Urban Institute of the participants in the TANF program in 1997 indicated that 77% reported at least one obstacle to working.

Disentangling the effects of ability and disutility of labor is very important, as we show that the optimal program depends substantially on which effect dominates. Consider the following two populations that a planner could face: the elderly and disabled, and poor young single mothers. For the first group, it is clear that income generating ability is low and observable, but actual difficulty of working may be noisy. For the second group, the

[1]For an analysis, see Blank (1997).

[2]For an extreme example, The New York Times (September 15, 1998) describes how Hispanic immigrant women's language difficulties, inadequate job training, and traditional values prevent them from being able to find jobs.

burden of working is clear, but actual ability may not be. And even within these groupings, many would argue that neither characteristic can be observed. Should a program designed to assist one group look the same as one designed for the other? In what follows, we examine the solution when each characteristic is unobservable[3] and, using multidimensional screening techniques, an environment where *both* characteristics are unobservable[4].

The solutions show distinct differences between the environments that have strong policy implications. These differences arise from the way each characteristic affects the trade-off between reducing informational rents and maintaining the earned income of workers. We find that a workfare (unpaid public sector work) policy may be used for screening when ability is unobservable, but is inefficient when disutility of labor is unobservable. However, a program with minimum work requirements (for paid work) is optimal when disutility of labor is unobservable. The optimality of work requirements in the two dimensional screening model depends on the relative importance of each characteristic. In addition, all of the programs look like variants of the Negative Income Tax - they have an income guarantee for those at the bottom of the distribution and a Benefit Reduction Rate (BRR, the implicit marginal tax rate) which depends on population characteristics. Given the wide variance in state population characteristics and in state TANF programs[5], these solutions provide a framework for understanding how the incentive effects and institutional design interact with different populations.

This work provides an application of several advances in mechanism design theory. Voluntary participation of agents in the income maintenance program represents an example of a "countervailing incentive", introduced by Lewis and Sappington (1989a, 1989b), and further explored by Maggi and Rodríguez-Clare (1995) and Jullien (1997). With reservation utility allowed to vary with type, the participants have the incentive to claim that they have a 'better' type and thus a higher opportunity cost that needs to be compensated, which conflicts with the usual informational incentive to claim that one has a 'worse' type. Multidimensional screening problems and their difficulties have been examined since Laffont, Maskin, and Rochet (1987) solved for a specific parameterization of a non-linear pricing model. We are able to use their results (generalized in McAfee and McMillan (1988)) and reduce the dimensionality of the problem, allowing us to find the optimal control solution.

The paper is organized as follows: In Section 2 we construct an example that explicitly shows how differences in what information is observable can substantially change policy recommendations. In Section 3, the setup for the general model appears. In Section 4, we

[3]Our examination of the environment where only income generating ability is unobservable is essentially a continuous version of Besley and Coate (1995).

[4]Beaudry and Blackorby (1998) investigate a utilitarian program where value of market and non-market time are unobservable using a framework in which consumption and leisure are perfect substitutes.

[5]For example, in 1997 states' Benefit Reduction Rates were between .33 to 1, with quite a lot of variance. For a more detailed comparison of state TANF programs see Gallagher et al. (1998).

look at the case when just ability is unobservable, and Section 5 then finds the optimal program when just disutility of labor is unobservable. Section 6 solves the model for when both dimensions are unobservable, and Section 7 concludes.

2 An Example: Workfare

Suppose there are two types of individuals in the population that a planner might include in her income maintenance program, one who earns \$6 per hour and one who earns \$12 per hour. The planner knows that the population contains 40% of the first group and 60% of the second. She has also determined that the minimum income level necessary to live is \$200 a week, and would like to ensure that everyone can obtain this at a minimal cost to the government. She has three instruments to use: hours of work, transfers and hours of workfare. Workfare is defined as unpaid labor.

If we define the utility of an individual as $wage * hours\ worked - .15 * (hours\ worked)^2$, we find that in the absence of the program (which we call the status quo) the first type of individual falls below the minimum income level (they earn \$120) while the second type is far above it (\$480). The incentive problem arises since types are unobservable. If the planner just offered \$80 conditional on the amount of work that the \$6 type had been working in the status quo, both groups would accept the deal. The planner must therefore minimize costs subject to incentive constraints as well as income maintenance constraints. We also assume that the planner sets a fixed reservation utility for both types, in order to not decrease their utility too much in the pursuit of guaranteeing them income. We set this equal to 60, the utility of the low type in the status quo.

Besley and Coate (1992) consider whether workfare can be optimal as a screening mechanism. Workfare can help screen by imposing a burden on those who announce that they are the low type. The planner must bear a cost for this, which comes in the form of higher transfers to the low types (since they will have reduced earnings). In the above problem the costs to the government are lower than the benefits, and a workfare scheme is implemented. The low types, who worked in the private sector for 20 hours in the status quo, now only work in the unpaid public sector for 30.6 hours and receive a transfer of \$200. The high types work the same amount as they did before, but must pay a tax of \$180. From this, we see that workfare is optimal for at least some parameter values[6].

The analysis of Besley and Coate *does not* allow for any unobservable characteristics besides ability. If we redefine the problem as one where disutility of labor is unobservable, will the policy recommendation of implementing a workfare scheme hold? Again, the rea-

[6] The condition for optimality of workfare is $w_l - (1-\alpha)w_h < 0$, where w_l is the wage of the low type, w_h is the wage of the high type, and α is the fraction of low types. Note that Besley and Coate (1992) employ a somewhat different formulation (they only allow for two instruments, workare and transfers and use a voluntary participation constraint) but find the same restriction for workare to be the optimal allocation.

soning for implementing workfare would be that by imposing a burden on those who declare themselves high disutility types, the cost saved from screening would be more than the cost of augmenting the income of the high disutility types. Suppose the planner targets a population of equal ability individuals who earn a wage of \$6. In the previous problem the index of disutility of labor was fixed at .15. Now we assume there are two types: 40% of the relevant group has an index of .15 and 60% has an index of .075. Therefore, without any program, the high disutility types earn \$120 and the low disutility types earn \$240, \$40 more than the minimum income level. The solution to the workfare problem in this environment has the high disutility individuals working in the private sector for 30.6 hours and receiving a transfer of \$16.4, while the low disutility individuals work in the private sector at their status quo level of 40 hours and pay taxes of \$9.8. Workfare levels, in contrast with the previous scenario, are set to zero for both individuals. Indeed, for all parameter values we find this will be true.

Lemma 1 *In the income maintenance problem with disutility of labor unobservable (and there are two types) there is no solution where a planner implements a workfare requirement.*

The proof of this (using a more general quasilinear utility function) can be found in the appendix. The intuition is straightforward: in addition to the cost of compensating high disutility types for wages lost by displacing their private sector work, the planner must deal with the fact that the cost to the low disutility type of accepting workfare is less than the cost to the high disutility type. This increases the incentive of the low disutility type to deviate, and eliminates workfare's role as a screening device[7]. The above result clearly implies that the environment in which disutility of labor is unobservable markedly differs from that where just income generating ability is unobservable. Policy recommendations should therefore be sensitive to what environment is being discussed.

3 Environment

Income maintenance is a critical element of poverty reduction programs. Policy initiatives such as the Negative Income Tax, wage subsidies, Earned Income Tax Credit, and conditional income subsidies all embed the implicit goal of increasing the income of the poor. In this paper income maintenance represents an explicit goal of the planner and is incorporated as the constraint that income be greater than or equal to some minimum prescribed level z. Since z is not defined in terms of other parameters, the program is general enough to solve for any amount of poverty reduction.

[7]Cuff (1999) also examines the efficiency of workfare as a screening device when both ability and disutility of labor are unobservable from a welfarist perspective. Her findings are analogous to the ones here.

The analysis operates in a static, deterministic environment. Individuals' types are exogenous and observable to the individual but not the planner. Income generating ability is represented by w and is assumed to be perfectly correlated with the real wage. We represent the disutility of labor by an index θ. The types w and θ have supports of $[0, b]$ and $[\theta_a, \theta_b]$respectively[8], and are known to be jointly distributed according to the pdf $g(w, \theta)$. We make use of the conditional pdfs, which for notational simplicity are called $f(w)$ and $p(\theta)$[9]. The instruments available to the planner are hours of work (l) and transfers (t).

There are a few assumptions that are implicitly embedded in the model. First, we assume that hours are observable and income is not. This facilitates the analysis and highlights the incentive effects on work. The assumption is harmless, as it yields the same qualitative results as when income is observable instead of hours[10]. An intuitive way to think about this is that each type can actually emulate (almost) everyone else's income through their labor choice, making the incentive constraint essentially the same in both cases.

Second, we assume that the planner can separate by some visible characteristic the population into two groups, the well off and the less well off. The planner constructs the income maintenance plan for the less well off group and funds it from the well off group. An example of such a visible characteristic would be assets or one's placement in defined income ranges[11]. This is similar to how most poverty assistance programs in the U.S. work. The method of funding is not considered in what follows; we just assume that the planner wants to accomplish income maintenance as cheaply as possible. It is important to note that the definition of the less well off group is a construct; not everyone in the group will actually be included in the program (the program size is determined endogenously)[12]. Also, participants do not have the opportunity to gain skills and exit the program. This is a limitation of the model, a result of the static aspect of the analysis. It proves useful, though, in isolating the incentive effects of income maintenance.

In contrast to the examples in the previous section, the model is made more realistic by replacing the individual rationality constraint with a voluntary participation constraint. This means the planner must induce people to participate and does not have the ability to force a program on the population. An important implication of adding a voluntary

[8]We expand beyond the two type model in order to learn more about the shape of the optimal schedule (and under what conditions the constraints bind). As opposed to Besley and Coate (1995), the type space is a continuum. This makes the multidimensional analysis more tractable.

[9]From our definition $f(w) = g_{1|2}(w \mid \theta) = \frac{g(w,\theta)}{\int_0^b g(w,\theta)dw}$, and is associated with c.d.f. $F(w) = \int_0^w f(x)dx$. Also, $p(\theta) = g_{2|1}(\theta \mid w) = \frac{g(w,\theta)}{\int_{\theta_a}^{\theta_b} g(w,\theta)d\theta}$, and is associated with c.d.f. $P(\theta) = \int_\theta^{\theta_b} p(y)dy$..

[10]The results Section 3.2 (unobservable ability) match Besley and Coate's despite the difference in instruments.

[11]Note that given the construction of the model, income itself could reveal information about the ability or disutility of labor of the individual.

[12]This implies that the highest ability b and the lowest disutility θ_a should be thought of as useful measures for how many people are included in the programs rather than cutoffs.

participation constraint is that transfers must be positive, eliminating any taxation aspect of the program.

We assume that an individual's utility is quasilinear in income, $U(w,\theta) = wl - h(l,\theta)$, and that $h_l\left(\right), h_{ll}\left(\right), h_{l\theta}\left(\right) \geq 0$. Two specific functional forms for $h(l,\theta)$ are used throughout the paper, θl^2 and $(l+\theta)^2$. The solutions are not qualitatively affected by the use of either one until the two dimensional case is analyzed. In the two dimensional case the latter is relied upon for reasons which will be made clear.

If no income maintenance program was constructed, people would choose l^* to maximize $U(w,\theta)$ and the maximized $U(w,\theta)$ is labeled $U^0(w,\theta)$. We call the absence of the program the 'status quo' and refer to individuals' choices as 'status quo' allocations. A mechanism consists of (l,t).

4 Income Maintenance When Ability Is Unobservable

4.1 The Full Information Problem

Consider a planner's problem in constructing an Income Maintenance Program when income generating ability, w, is known and θ is fixed. In this case, the planner knows if anyone is above the minimum income level before any program is instituted, and can refrain from giving them any transfers. For those types that deserve aid, the planner must offer a package that gets them above the minimum income level and (weakly) induces them to participate. These two constraints are labelled Income Maintenance and Voluntary Participation. The program is solved as follows.

$$\min_{\{l(w),t(w)\}} \int_0^b t(w)f(w)dw$$

such that

$$wl + t \geq z \qquad (IM)$$

$$wl - h(l,\theta) + t \geq U^0(w) \qquad (VP)$$

To find the solution we first define l^* implicitly from the individual's maximization problem: $h'(l^*,\theta) = w$. We also define $\bar{l}(w)$ as the value of l when both IM and VP bind, giving us $h(\bar{l}(w),\theta) = z - U^0(w)$. Since $\bar{l}(w)$ is the l that minimizes transfers while satisfying both constraints, the full information solution can be characterized in the following manner. If $wl^* \geq z$, $l = l^*$ and $t = 0$. If $wl^* < z$, then $l = \bar{l}(w)$ and $t = z - w\bar{l}(w)$. In words, those who have status quo income above z are not included, and those whose income was less than z receive an income of z while being made indifferent between the program and the status quo.

The labor allocation for program participants is larger than their status quo labor choice. Transfers begin at z and decrease with ability to zero. The labor allocation also decreases with ability. Following the interpretation by Besley and Coate (1995), this schedule resembles a conditional income or wage subsidy scheme. An individual of type w who earns income $w\bar{l}(w)$ receives a benefit of $z - w\bar{l}(w)$; if she earns less, she receives nothing. This will (weakly) induce participation and increase the amount of labor supplied by the targeted population.

Since hours worked increase with ability for the population not admitted to the program, the first best allocation is not incentive compatible. Once ability becomes unobservable, this part of the population could claim low ability, work the same number of hours that they did in the status quo, and receive a transfer. We must accordingly focus on a second best world.

4.2 Solving the Income Maintenance Problem (IMP$_w$)

Now we explore the planner's problem when w is unobservable. We add to the problem considered above the standard incentive compatibility constraint to induce truth-telling:

$$wl(w) - h(l(w), \theta) + t(w) \geq wl(\hat{w}) - h(l(\hat{w}), \theta) + t(\hat{w}) \ \forall w, \hat{w} \qquad (IC)$$

Note that the single crossing property holds and using methods from Myerson (1981) allows us to express utility as $U(w) = U(0) + \int_0^w l(x)dx$. We require that $l(w)$ be non-decreasing and continuous. We assume that $\frac{F(w)-1}{f(w)}$ is increasing in w, as is standard. In addition, we define $h(l, \theta) = \theta l^2$ for notational convenience. All qualitative results hold for any $h(l, \theta)$ that satisfy our assumptions from section 2. With this functional form, $U^0(w) = \frac{w^2}{4\theta}$. Since income (defined as real earnings plus transfers) is monotonically increasing in w, IM must only be satisfied at $w = 0$ and can be replaced with $t(0) \geq z$.

Using a solution technique suggested by Maggi and Rodríguez-Clare (1995), we redefine the problem and analyze it in an optimal control setting. Let $R(w)$ represent rent, that is, type w's utility above her reservation utility (notationally, $R(w) = U(w) - \frac{w^2}{4\theta}$). The objective function can now be written as:

$$\int_0^b \{-R(w) - \frac{w^2}{4\theta} + wl(w) - \theta l(w)^2\} f(w)dw \qquad (IMP_w)$$

The main constraints are VP (which now takes the form $R(w) \geq 0$), IC ($R_w = l(w) - \frac{w}{2\theta}$), and IM ($R(w) + \frac{w^2}{4\theta} + \theta l(w)^2 - z \geq 0$). Additionally, the implementability constraint $\frac{\partial l}{\partial w} \geq 0$ and a non-negativity constraint $l(w) \geq 0$ must be taken into account. We form the Hamiltonian:

$$H(R, l, u, \lambda, \alpha, w) = \{-R - \frac{w^2}{4\theta} + wl - \theta l^2\} f(w) + \lambda(l(w) - \frac{w}{2\theta}) + \alpha u$$

where $R(w)$ and $l(w)$ are the state variables, $\lambda(w)$ and $\alpha(w)$ their respective costate variables, and $u(w) = \frac{\partial l}{\partial w}$ is a control. We add the non-negativity constraints to form a Lagrangian.

The VP constraint presents us with a case of "countervailing incentives" (analyzed in Lewis and Sappington (1989a and 1989b) and generalized in Maggi and Rodríguez-Clare (1995) and Jullien (2000)). A potential applicant for IMP_w has an incentive to understate his actual income generating ability, since the program intends to bring individuals with low earnings above a minimum income level. As seen in the first best solution, this involves higher transfers for lower types. The voluntary participation constraint, however, shows that higher types have higher reservation utilities, and must be appropriately compensated for participating in the program. This provides applicants with a conflicting incentive to overstate their types. It seems natural to have participants' outside options vary with their types. Since the program is not mandatory, a planner can't impose a minimum utility level on applicants. In fact, participation (take-up) represents an important issue for welfare researchers[13]. The solution, which determines how to allocate informational rents, depends on how the incentives interact. The following lemma lends insight into the problem.

Lemma 2 *If VP binds at $w' > 0$, it binds for all $w > w'$.*

We first rewrite VP using the envelope theorem to gain some intuition.

$$U(0) + \int_0^w l(x)dx \geq U^0(0) + \int_0^w l^*(x)dx \qquad (VP')$$

Say that VP' binds at some w'. Then $l(w'+\epsilon)$ must equal $l^*(w'+\epsilon)$. If it were less than $l^*(w'+\epsilon)$, VP' would be violated. If it were equal to $l^*(w'+\epsilon)$, the transfer would equal zero, the best that can be achieved. By continuity of $l^*(\cdot)$, $l(w'^+) = l^*(w')$. The proposition now follows. We don't assume continuity of $l(\cdot)$ to get this result, although we will assume it for the optimal control problem and verify it using the solution. Note that the proposition doesn't hold for $w' = 0$. If VP' binds at 0, it must bind at $l(0) > l^*(0)$. This is because IM binds before VP' does at $w = 0$ for a large interval of l, which can be seen in the analysis of the first best solution (using the notation of that solution, IM binds first if $l(0) < \bar{l}(0)$). With $l(0) > l^*(0)$, there is built in slack in the VP' constraint for low types. The incentive to overstate (actually, to understate less) one's type increases with ability. The opportunity cost to workers of joining the program and receiving less working income than they would in the status quo begins to outweigh the extra transfers received for pretending that they are low types. This provides help in constructing a solution to the problem. The solution found satisfies both the necessary and sufficient conditions[14]. An overview of all solutions can be found in the appendices.

[13] For example, see Moffitt (1992) and Hoynes (1996).

[14] Since the IM constraint is convex in l, the standard sufficiency proof (see Seierstad and Sydsaeter (1987)) does not directly apply. A proof that applies to this specific case is available from the author upon request.

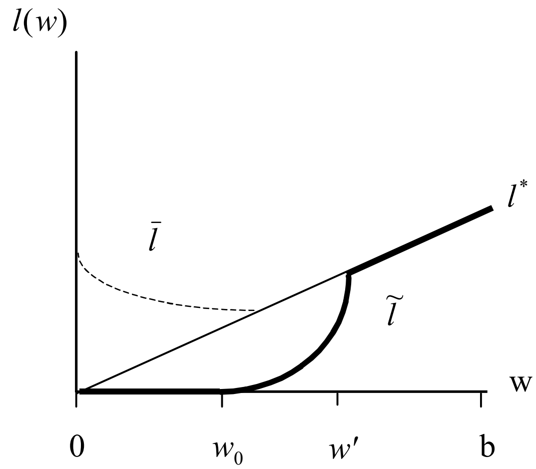

Figure 1: Income Maintenance With θ Fixed: Status Quo, First Best, and IMP_w

Proposition 1 *The solution to IMP_w such that IC, VP', and IM hold is:*

i) for $w \in [0, w_0]$, $l(w) = 0$, $t(w) = z$

ii) for $w \in [w_0, w']$, $l(w) = \frac{w}{2\theta} - \frac{F(w')-F(w)}{2\theta f(w)}$, $t(w) = z + \int_0^w l(x)dx - wl(w) + \theta l(w)^2$

iii) for $w \in [w', b]$, $l(w) = \frac{w}{2\theta}$, $t(w) = 0$

where we define w_0 by $\frac{w_0}{2\theta} - \frac{F(w')-F(w_0)}{2\theta f(w_0)} = 0$ and w' by $R(w') = 0$ (if there is no w'such that $R(w') = 0$, $w' = b$).

The results of the above proposition can be seen in Figure 1 (in which the solution is in bold and labelled as $\tilde{l}(w)$). The status quo allocation l^* and the first best allocation ($\bar{l}$ up until $wl^*(w) = z$, and l^* from then on) are placed alongside the results for comparison.

The objective is clearly to allocate transfers to low types. High types value work more than low types, so by distorting the labor allocation of the low types downward and further away from the status quo allocation, incentive compatibility can be achieved. Of course, the trade-off of reaching incentive compatibility is that the reduction in informational rents are balanced by the costs of distorting the allocations away from the full information solution. The allocation of zero hours to the bottom of the distribution comes from the fact that the $l(w)$ that maximizes the virtual surplus for low types is negative, running into our non-negativity constraint. The countervailing incentive to overstate (again, to understate less) one's ability becomes an issue as the reservation utility increases more quickly for the high types. Transfers decrease from z for the bunched interval to 0 starting at w'. The people at the top end of the distribution are offered their status quo allocation, so they essentially opt out of the program. Income maintenance is satisfied for this interval since l is increasing and IM was satisfied at the bottom of the distribution. This implies that a necessary (but not sufficient) condition for having VP' bind ($w' < b$) is that the higher types were earning more than z in the status quo.

Besley and Coate (1995) present a solution using discrete types similar to the one above. They find three regions as we do: low types are bunched and have IM bind, middle types work less than in the status quo and receive transfers that decrease with ability, and high types receive their status quo allocation.

The solution sheds light on how the optimal income maintenance program relates to actual poverty assistance programs. First, the transfer schedule is analogous to a Negative Income Tax. Nonworking participants are guaranteed the minimum income level. Participants who work receive a transfer that decreases as work increases (but at a lower rate, as it can be shown that earned income plus transfers is increasing with work) until the point where transfers equal zero, and earnings revert to the status quo amount. The Benefit Reduction Rate, the rate at which earnings are taxed, is not constant and depends on the distribution of the population. It is bounded below one, distinguishing this program from the precursor to TANF, the Aid to Families with Dependent Children program (AFDC) , which had a marginal tax rate of 100% for a substantial length of time. It is intuitive to make the BRR depend on population characteristics: if there is a large mass of very low types, providing some incentive to work is beneficial.

Second, hours worked are distorted downward. While a standard mechanism design result, this provides a rationale for work disincentives in a transfer program - to eliminate the adverse selection problem. Lastly, like any of these actual programs, it is sufficient to report status quo income for eligibility and transfers. Since the planner can generate the ability level/wage from status quo income, she can provide a menu of hours and transfers linked to income. Therefore, IMP_w can be considered a means-tested program.

4.3 Mandatory Participation and Typical Social Welfare Functions

Now we fix a common reservation utility (adding a standard Individual Rationality constraint) instead of allowing people to revert to their status quo option (dropping the VP' constraint) in order to explore the implications of an income maintenance program without the outside option of the standard labor market - a mandatory program. If IR takes the form $U(w) \geq r$, we find the following:

Corollary 1 *The solution to IMP_w such that IC, IR, and IM hold is:*
i) for $w \in [0, w_0]$, $l(w) = 0$, $t(w) = z$
ii) for $w \in [w_0, b]$, $l(w) = \frac{w}{2\theta} - \frac{1-F(w)}{2\theta f(w)}$, $t(w) = z + \int_0^w l(x)dx - wl(w) + \theta l(w)^2$
where we define w_0 by $\frac{w_0}{2\theta} - \frac{1-F(w_0)}{2\theta f(w_0)} = 0$.

The solution is similar, and we find that in cases where VP' does not bind for high types, the problem corresponds exactly to one with an IR constraint (where the reservation value is equal to the status quo reservation utility of the lowest type). As a result, it is clear that the cost of a program with VP' is at least as great as, if not greater than, the cost with the IR constraint (with an appropriately chosen r). This implies that the middle types are receiving higher transfers under a VP' scheme. It is also possible that transfers are negative for high types in the mandatory program (as can be seen in the example in the introduction).

The above results further accentuate the contrast with common views on welfare reform. Creating a mandatory program does not give an incentive to the planner to increase hours of work by participants. In fact, for all cases, the planner either keeps hours the same or lowers them. The planner achieves lower costs and satisfies the objectives of income maintenance and full participation. What this environment ignores is the dynamic argument that work may build skills and independence, allowing participants to increase their income generating ability and possibly remove themselves from the program. With this argument, it is possible that under-allocating work could cost more in the long run[15].

In addition, this mandatory participation model allows a clear comparison with a welfarist optimal taxation model. While the optimization problem (minimizing costs subject to incentive compatibility, individual rationality, and income maintenance) makes sense from the perspective of a policymaker implementing welfare reform, it does not immediately offer comparability with standard social welfare functions (for example, see Mirrlees (1971)). It can be shown, however, that the mandatory participation problem is equivalent to maximizing a Rawlsian objective function (maximizing the utility of the worst off person) with incentive compatibility, income maintenance and a constraint on the size of transfers. This makes sense; the individual rationality constraint, if it binds, binds only for the lowest type - therefore lowest type is being considered in the mandatory participation model. Despite the apparent gap between the Rawlsian objective and a Utilitarian objective (maximizing expected/average utility of the population), the Utilitarian objective yields a qualitatively similar result, although labor is not distorted downwards as much. This comes from the fact that the "average" type values labor more than the lowest type.

5 Income Maintenance When Disutility of Labor Is Unobservable

One of the main challenges facing welfare reform is how to decrease the long-term dependency of some program participants. Why are these people so dependent? One answer is that they

[15]The dynamic argument does not hold for disutility of labor - work can't really decrease someone's tolerance for it. However, other possible instruments and transfers outside of the model (such as subsidized child care or better access to transportation) could change someone's disutility of labor over time.

lack the skills that employers value. Another answer is that some people find that working full time is not possible given their personal situation. Individuals may have responsibilities to their families or disabilities that make it difficult to participate in the labor force. Both answers make sense and some may argue that they are very closely related; that is, the ability of an individual is correlated with their disutility of labor. Given our underlying framework of rational choice by individuals though, it is important for the planner to take into account all relevant factors on how the participants make decisions. For example, an independent woman with a high school diploma will make very different choices than a single mother of similar educational attainment who must take care of family members in poor health.

In practice, we find that under TANF, states do not require that all welfare recipients work and often target disutility of labor. Adults who have disabilities or care for someone who is disabled are generally exempted. Adults with infants also often receive temporary exemptions. Although there are other programs for people with disabilities, such as SSI (Supplemental Security Income), many disabled people do not qualify for any program except welfare. The Urban Institute[16] estimated that about 28% of families on the AFDC program had either a mother or child with a "functional limitation".

We try to capture the effect of heterogeneity in terms of disutility of labor by a single dimensional index, θ. Large disutility corresponds to a large value of θ, and in the status quo it can be seen that this decreases the hours of work chosen by the individual. The program IMP_θ, in which ability is fixed and disutility of labor is unobservable and allowed to vary is a construct that will isolate the effect of θ on the planner's problem. Later in the paper both unobservables will be allowed to vary, allowing us to understand how a richer program can be designed. In that case, correlations between unobservables are implicitly taken into account.

5.1 Full Information

The problem is structured in the same way as IMP_w, with w fixed.

$$\min_{\{l(\theta),t(\theta)\}} \int_{\theta_a}^{\theta_b} t(\theta)p(\theta)d\theta$$

such that

$$wl + t \geq z \tag{IM}$$

$$wl - h(l,\theta) + t \geq U^0(\theta) \tag{VP_θ}$$

[16]The Urban Institute. (1996) "Profile of Disability Among AFDC Families"

We define l^* as before and $\bar{l}(\theta)$ as the value of l when both IM and VP_θ bind, giving us $h(\bar{l}(\theta),\theta) = z - U^0(\theta)$. The allocation $\bar{l}(\theta)$ is the l that minimizes transfers while satisfying both constraints. This yields the full information solution: if $wl^* \geq z$, $l = l^*$ and $t = 0$, and if $wl^* \leqslant z$, then $l = \bar{l}(\theta)$ and $t = z - w\bar{l}(\theta)$. By implicit differentiation, we find that $\bar{l}(\theta)$ is decreasing in θ, which contrasts with the IMP_w solution. As in IMP_w, the full information solution can be viewed as a conditional wage or income subsidy.

5.2 The Incentive Problem (IMP$_\theta$)

Again the functional form $h(l,\theta) = \theta l^2$ is used. Here, the assumption of a functional form could possibly make a difference in the solution. We have not generalized the results, but the same qualitative solution was found using the functional form $h(l,\theta) = (l+\theta)^2$. We present the first formulation as it is somewhat simpler notationally.

We add the incentive constraint to the problem.

$$wl(\theta) - \theta l(\theta)^2 + t(\theta) \geq wl(\hat{\theta}) - \theta l(\hat{\theta})^2 + t(\hat{\theta}) \ \forall \theta, \hat{\theta} \qquad (IC_\theta)$$

Utility for type θ can be written as $U(\theta) = U(\theta_b) + \int_\theta^{\theta_b} l(x)^2 dx$. Note that in this case, we find that $l(\theta)$ must be decreasing in θ (we assume that it is continuous as well). We assume that $\frac{P(\theta)}{p(\theta)}$ is decreasing in θ. Using monotonicity, we simplify the IM constraint to be $wl(\theta_b) + t(\theta_b) \geq z$. We define the utility rent as $R(\theta) = U(\theta) - U^0(\theta)$. We must then maximize

$$\int_{\theta_a}^{\theta_b} \{-R(\theta) - \frac{w^2}{4\theta} + wl(\theta) - \theta l(\theta)^2\} p(\theta) d\theta \qquad (IMP_\theta)$$

subject to IM and VP_θ ($R(\theta) \geq 0$).

Here the applicant has an incentive to overstate her disutility of labor in order to receive higher transfers. A conflicting incentive stems from the potential for the applicant to understate her disutility of labor in order to inflate the price of her outside opportunities. The value of reservation utility is greatest for low types (those with low disutilities of labor) and for them we discover that an analogue to lemma 2 is true.

Lemma 3 *If VP_θ binds for $\theta' < \theta_b$, then it binds for all $\theta < \theta'$*

The solution, however, departs from IMP_w in a significant way.

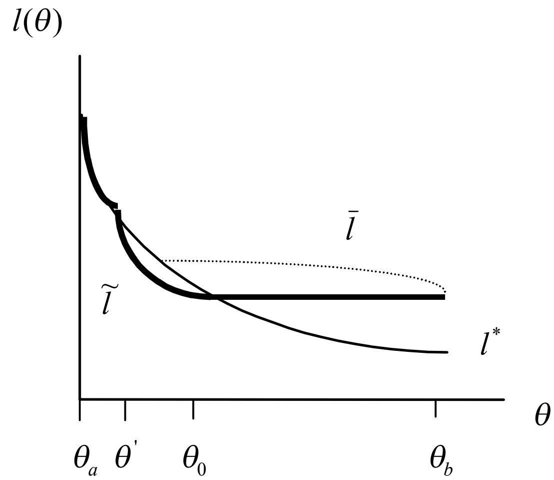

Figure 2: Income Maintenance With w Fixed: Status Quo, First Best, and IMP_θ

Proposition 2 *The solution*[17] *to* IMP_θ *such that* IC_θ, VP_θ, *and* IM *hold*[18] *is:*

i) for $\theta \in [\theta_a, \theta']$, $l(\theta) = \frac{w}{2\theta}$, $t(\theta) = 0$

ii) for $\theta \in [\theta', \theta_0]$, $l(\theta) = \frac{w}{2(\frac{P(\theta')-P(\theta)}{p(\theta)}+\theta)}$, $t(\theta) = -\theta_b \bar{l}(\theta_b)^2 + z + \int_\theta^{\theta_b} l(x)^2 dx - wl(\theta) + \theta l(\theta)^2$

iii) for $\theta \in [\theta_0, \theta_b]$, $l(\theta) = \bar{l}(\theta_b)$, $t(\theta) = z - w\bar{l}(\theta_b)$

where we define θ_0 *by* $\frac{w}{2(\frac{P(\theta')-P(\theta_0)}{p(\theta_0)}+\theta_0)} = \bar{l}(\theta_b)$ *and* θ' *by* $R(\theta') = 0$ *(if* $\nexists \theta'$ *such that* $R(\theta') = 0$, $\theta' = \theta_a$*).*

Figure 2 shows the solution described in the above proposition (it is in bold and labelled as $\tilde{l}(\theta)$). The status quo allocation l^* and the first best allocation ($\bar{l}$ for θ such that $wl^*(\theta) \leqslant z$, and l^* before that) are placed alongside the results for comparison.

We see from the solution that VP$_\theta$ binds for the lowest type in the distribution as well as for the highest type. The transfer to the highest type is minimized since both VP$_\theta$ and IM bind for it. Notice that for $w > 0$, both status quo labor choice (l^*) and the program allocation are bounded away from 0. This moves l above zero and above l^* to the point where the planner can make both constraints bind. As θ_b increases to infinity, $\bar{l}(\theta_b)$ approaches $l^*(\theta_b)$ and both approach zero. This is reasonable; as disutility of labor becomes extremely high, a person would not work and couldn't really be placed in a job without huge transfers. Nevertheless, for finite θ_b the planner forces high types to work more than in the status quo. Middle types work less than the status quo. All types are assigned less (or equal amounts of)

[17] As with IMP_w, we note that the standard sufficiency proof does not hold because o f the convexity of IM in l. For IMP_θ our proof of sufficiency is much more restrictive - we need the condition $P(\theta') - 2Y > 0$ (where Y is defined as in the appendix) to hold.

[18] Note that this schedule assumes that $l^*(\theta_a) > \bar{l}(\theta_b)$. If otherwise (generally this can occur for low values of w), the solution is $l(\theta) = \bar{l}(\theta_b)$, $t(\theta) = z - w\bar{l}(\theta_b)$ for all θ. Both solutions are discussed in Appendix C.

work than they would receive in the first best allocation, due to the informational rent that they must be given. While this is qualitatively the same for IMP_w, having the program begin at $l = \bar{l}$ is impossible in IMP_w because monotonicity would conflict with the decreasing first best allocation (compare Figure 1 and Figure 2). It is also interesting to observe that the usual "no distortion at the top" result holds, as does no distortion for the "worst" type. This kind of result has also been found in specific formulations of the optimal income taxation model[19].

Both IMP_w and IMP_θ differ with respect to the status quo in a surprising way. The program IMP_w makes people with low ability lie idle when, in the absence of a program, they would choose to work. On the other hand, IMP_θ makes those with high disutility of labor work more than they would if there were no program. Low ability people need large transfers whether they work or not, so IMP_w focuses more on decreasing transfers to others (reducing the incentive problem). High disutility people still earn the same amount as low disutility people in IMP_θ, making it more costly to reduce their labor.

The solution for IMP_θ departs from the Negative Income Tax structure of IMP_w. There is no guaranteed amount of transfers for no work. In fact, a guarantee of transfers begins at a designated minimum amount of work. As labor increases, a Benefit Reduction Rate kicks in, again dependent on the distribution of the population. This plan combines the NIT with minimum work requirements. Like IMP_w though, IMP_θ is a means-tested program and the hours and transfers can be conditioned on status quo income.

6 The Two Dimensional Problem

Multidimensional screening models, while presenting large opportunities for understanding beyond single dimensional models, are scarce due to their general unwieldiness. A uniform approach has only very recently been suggested and solved (see Armstrong (1996), Rochet and Choné (1998), and Rochet and Stole (2000)).

In our setting, the question of what a two dimensional solution looks like is a natural one. Welfare programs often discriminate between groups in ways that involve both ability and disutility of labor. In the AFDC program primary recipients were single mothers. If we think of poverty assistance in general, the elderly, people with disabilities, and those unable to get employment are targeted. The programs IMP_w and IMP_θ can't completely address this issue, so we need to expand our scope. One point that can be made is that welfare programs do distinguish between groups because some characteristics related to ability and disutility of labor are observable. While this is true to a certain extent, within categories, it is difficult to observe exact types, and one can think of the following as an approach to addressing each category. In addition, the framework allows for correlation between types.

[19] See Mirrlees (1997) for a discussion of such results.

Laffont, Maskin, and Rochet (1987) solve a specific multidimensional model of nonlinear pricing. Their methodology was generalized in McAfee and McMillan (1988) and will be used in this paper. The approach relies on the observation that there will be indifference curves across type space, where different type profiles receive the same allocation. If a "generalized" single crossing property holds (analogous to the single crossing property of one dimensional problems), then we have a knowledge of the form that incentive compatible indifference curves take and can potentially reduce the problem to an analysis of allocations to different indifference groupings, a single dimensional problem.

We redefine the problem slightly to both illustrate important comparative statics and simplify calculations. We set the lowest wage as $a \geq 0$ (as opposed to 0 for IMP_w) to use as a parameter to compare the single dimensional and multidimensional solutions. As a approaches b, the problem collapses from the two dimensional case to the one dimensional IMP_θ. We set θ_a equal to 0 for added tractability, though this also may be thought of as a normalization of disutility of labor. IM is defined as before. The problem can be stated as:

$$\max_{\{l,t\}} \int_0^{\theta_b} \int_a^b -t(w,\theta)g(w,\theta)dwd\theta$$

such that IM holds, as well as:

$$U(w,\theta \mid w,\theta) \geq U(\hat{w},\hat{\theta} \mid w,\theta)\ \forall w,\theta,\hat{w},\hat{\theta} \qquad (IC_{w\theta})$$

$$U(w,\theta) \geq U^0(w,\theta) \qquad (VP_{w\theta})$$

This formulation satisfies the Generalized Single Crossing Property defined by McAfee and McMillan (1988). Using this property, we can describe incentive compatible isolabor curves parameterized by the variable s.

Here we must make a few more simplifying assumptions. We assume a functional form for $h(l,\theta)$ of $(l+\theta)^2$. This satisfies all previous assumptions on $h(l,\theta)$. Note that both the parameterizations $h(l,\theta) = (l+\theta)^2$ and $h(l,\theta) = \theta l^2$ yield the same qualitative results in both single dimensional cases, but that the latter makes the analysis more difficult in the two dimensional case. Essentially, this results from the fact that the cross partial derivative ($h_{l\theta}$) varies with l, making the Spence-Mirrlees single crossing condition endogenous. One can think of θ in the former as a shift parameter, while in the latter θ changes the curvature[20].

[20] Another way to think about the difference between the two functional forms is to consider two types θ_l and θ_h ($\theta_l < \theta_h$). For the functional form $(l+\theta)^2$, the difference in disutilities between types is somewhat similar irrespective of the allocation l, while for θl^2, the difference is small for low allocations and large for high allocations. This implies that a planner should consider carefully the impact of assigning full work weeks more if utility takes on the second form, making our results potentially sensitive to specification for the better off types.

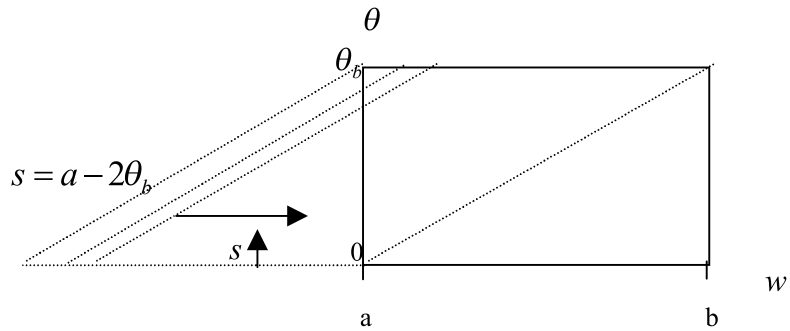

Figure 3: Isolabor curves (indexed by s) in (w, θ) space

When solving for l^* for this functional form, we notice that it will take negative values, so it is redefined as $l^* = \max\left[\frac{w}{2} - \theta, 0\right]$. This tells us that some people choose not to work in the status quo because of some combination of low ability and high disutility of labor. Within the context of the model, this implies that they earn no income in the status quo. Since the utilities have been normalized, this does not imply that they have zero resources[21]. We strictly define the income maintenance constraint to be *earned* income greater than or equal to z [22]. This population in the U.S. is not counted as unemployed because it is not seeking work, but it does exist and there is evidence that it is growing. The upper bound on disutility of labor θ_b is set equal to $\frac{b}{2}$ to make the results readable (but does not affect results qualitatively). This assumption tells us that even the highest ability person, if they have the highest disutility of labor, will not work in the status quo.

Given the above formulation we can now define a 'type aggregator' s, which indexes incentive compatible isolabor curves, using the equation $w = s + 2\theta$. Appendix D reviews how this is derived and Figure 3 depicts s in (w, θ) space. This means s is in the interval $[a - 2\theta_b, b]$ and allows us to transform variables from (w, θ) to (s, θ). Utility becomes $U(s, \theta) = sl(s) - l(s)^2 - \theta^2 + t(s)$; and since θ does not interact with the allocation we can simplify $U(s, \theta)$ to $V(s) = sl - l^2 + t$. The allocation depends only on the relevant isolabor curve s, and

[21] For a study of how single mothers survive without earned income, see Edin and Lein (1997).

[22] Remember that some criteria has been established to designate the population $[a, b] \times [0, \theta_b]$ less well off and under consideration for inclusion in the income maintenance program. Therefore, even though retired millionares may choose not to work, we assume that they will be ineligible for this program.

this formulation makes it clear that in terms of incentive compatibility, the planner only cares about the reported s, not the separate w or θ. The way the model has been constructed, the isolabor curves are exogenous with respect to allocation choice, simplifying the derivations. Therefore, using the same arguments from the single dimensional approach[23], it must be that $\frac{\partial}{\partial s}U(s,\theta,l(s),t(s)) = U_s$ and $\frac{\partial}{\partial s}V(s,l(s),t(s)) = V_s = l(s)$. We also require $l(s)$ to be increasing in s.

The constraints must now be transformed to (s,θ) space. Reservation utility depends on whether $\frac{w}{2} - \theta > 0$ (notice that this is equivalent to $s > 0$). It equals $\frac{w^2}{4} - \theta w$ when positive amounts of labor are chosen, $-\theta^2$ otherwise. This permits us to transform $VP_{w\theta}$ into two constraints that depend solely on s:

$$V(s) > 0 \text{ for } s \in [a - 2\theta_b, 0) \qquad (VP1)$$

$$V(s) > \frac{1}{4}s^2 \text{ for } s \in [0, b] \qquad (VP2)$$

The voluntary participation constraints are so simple because status quo isolabor curves have the same slope as the incentive compatible isolabor curves.

The Income Maintenance constraint is more difficult to handle. Transformed, it becomes $V(s)+l^2+2\theta l \geq z$. Its dependence on θ does not conform to our method. However, noticing that it is satisfied for all points encompassed by a given s when it is satisfied for the minimum θ along that s provides a way to eliminate θ. For $s < a$, the minimum θ occurs where $w = a$. At these points $s = a - 2\theta$, and the constraint is $V(s) + l^2 + (a-s)l \geq z$ $(IM1)$. For $s \geq a$, the minimum θ is 0, reducing the constraint to $V(s) + l^2 \geq z$ $(IM2)$.

The objective function transforms to (letting $T(s,l(s)) = V(s) - sl(s) + l(s)^2$):

$$\max -\int_0^{\theta_b}\int_{a-2\theta}^{b-2\theta} T(s,l(s))g(s+2\theta,\theta)dsd\theta$$

θ can be eliminated by first permuting the integrals:

$$-\int_{-2\theta_b}^{0}\int_{\frac{a-s}{2}}^{\theta_b} T(s,l(s))g(s+2\theta,\theta)d\theta ds - \int_0^a\int_{\frac{a-s}{2}}^{\frac{b-s}{2}} T(s,l(s))g(s+2\theta,\theta)d\theta ds -$$
$$\int_a^b\int_0^{\frac{b-s}{2}} T(s,l(s))g(s+2\theta,\theta)d\theta ds$$

and then integrating out θ :

$$-\int_{a-2\theta_b}^{0} T(s,l(s))g_1(s)ds - \int_0^a T(s,l(s))g_2(s)ds - \int_a^b T(s,l(s))g_3(s)ds \qquad (IMP_{w\theta})$$

[23]For example, see Guesnerie and Laffont (1984).

where $g_1(s) = \int_{\frac{a-s}{2}}^{\theta_b} g(s+2\theta,\theta)d\theta$, $g_2(s) = \int_{\frac{a-s}{2}}^{\frac{b-s}{2}} g(s+2\theta,\theta)d\theta$, and $g_3(s) = \int_0^{\frac{b-s}{2}} g(s+2\theta,\theta)d\theta$.

We must make the following assumption to guarantee a monotone (non-bunched) solution:

- *Assumption:* Both $\frac{G_2(s)-G_2(s')}{g_2(s)}$ and $\frac{G_3(s)-G_3(s')}{g_3(s)}$ are increasing in s (where $G_2(s)$ is defined by $\int_0^s g_2(x)dx + \int_{a-2\theta_b}^0 g_1(x)dx$ and $G_3(s)$ by $\int_a^s g_3(x)dx + G_2(a)$).

This assumption is not easily verified and does not follow from the assumption of the monotone hazard rate for the conditional probability distribution functions $f(w)$ and $p(\theta)$. It can be shown, however, that $\frac{s}{2}$ plus either hazard rate is increasing in s for the uniform distribution. This will be used to examine the solution. The assumption essentially guarantees a separating solution, avoiding the necessity to check where the constraint that hours must be weakly increasing binds.

We find the optimal program by maximizing $IMP_{w\theta}$ subject to $VP1$ (for $s < 0$), $VP2$ (for $s \geq 0$), $IM1$ (for $s < a$), and $IM2$ (for $s \geq a$) and non-negativity constraints on $\frac{\partial l}{\partial s}$ and $l(s)$. This type of program, where the objective function and constraints change at some point (here when $s = 0$ and $s = a$), presents a three phase dynamic optimization problem. Each phase must fulfill the standard necessary conditions with an added requirement that the costate variables and the Hamiltonians at each junction be aligned. The specific conditions are listed in the appendix as are explanations of the results. The results that we find are quite intuitive and complement the one dimensional solutions. There exist two possible solutions, each of which satisfies the necessary and sufficient conditions for certain ranges of the parameters. The solution is discussed in the appendix.

Proposition 3 *The follow schedules are solutions (for certain parameter values) to $IMP_{w\theta}$:*

Schedule 1:

i) for $s \in [-2\theta_b, s_0)$, $l(s) = 0$, $t(s) = z$

ii) for $s \in [s_0, s')$, $l(s) = \frac{s}{2} + \frac{G_j(s)-G_i(s')}{2g_j(s)}$, $t(s) = z + \int_{a-2\theta_b}^s l(x)dx - sl(s) + (l(s))^2$

iii) for $s \in [s', b]$, $l(s) = \frac{s}{2}$, $t(s) = 0$

where we define s_0 by $\frac{s_0}{2} + \frac{G_h(s_0)-G_i(s')}{2g_h(s_0)} = 0$, s' by $V(s') = \frac{1}{4}(s')^2$, i as the interval containing s', $j \leq i$, and h as the interval containing s_0.

Schedule 2:

i) for $s \in [-2\theta_b, s_0)$, $l(s) = \bar{l}(a-2\theta_b)$, $t(s) = z - a\bar{l}(a-2\theta_b)$

ii) for $s \in [s_0, s')$, $l(s) = \frac{s}{2} + \frac{G_j(s)-G_i(s')}{2g_j(s)}$, $t(s) = -\theta_b^2 + \int_{a-2\theta_b}^s l(x)dx - sl(s) + (l(s))^2$

iii) for $s \in [s', b]$, $l(s) = \frac{s}{2}$, $t(s) = 0$

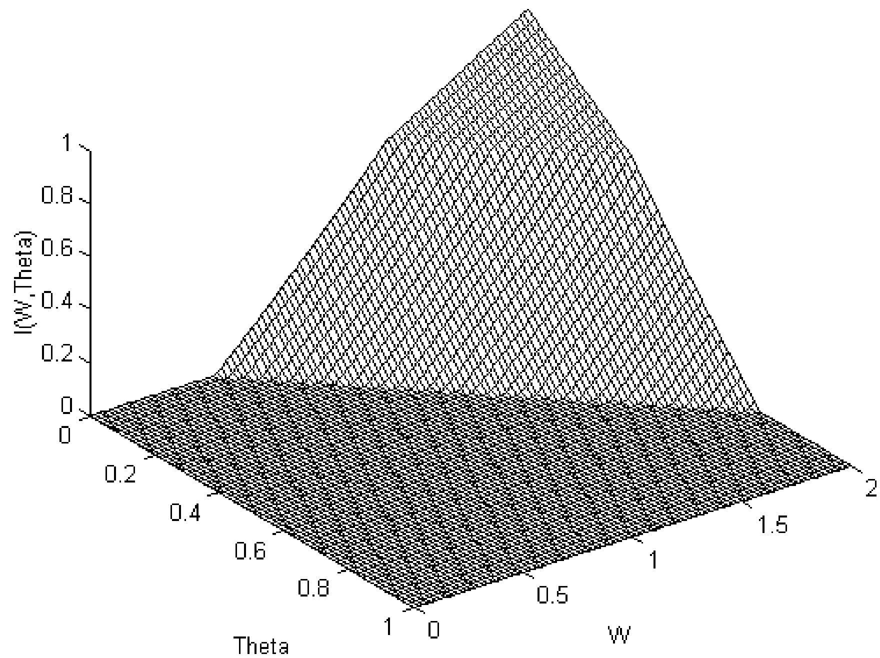

Figure 4: $IMP_{w\theta}$: uniform distribution, $a = 0$ and $z = .2$

where we define s_0 *by* $\frac{s_0}{2} + \frac{G_h(s_0)-G_i(s')}{2g_h(s_0)} = \bar{l}(a - 2\theta_b)$, s' *by* $V(s') = \frac{1}{4}(s')^2$, *and* i, j, h *are the analogues of their definitions for Schedule 1.*

The two possible schedules reflect the one dimensional solutions. Schedule 1, much like IMP_w, has an allocation of zero work and a transfer of the minimum income level for the lowest types (where lowest here means individuals with low ability and high disutility of labor) with transfers decreasing and work increasing as s increases until $VP2$ binds. Schedule 2, like IMP_θ, has minimum work requirements: the lowest types must work for $\bar{l}$ hours in order to receive a transfer. As s increases, the amount of work required increases and transfers decrease, until $VP2$ binds.

Clearly, it is important to understand the influence of the parameters in this problem. The main parameter of interest is a, the lower bound of wages. In the appendix we show:

Lemma 4 *For* a *close to* 0*, only Schedule 1 satisfies the necessary conditions. As* a *approaches* b*, only Schedule 2 satisfies the necessary conditions.*

This proves that both one dimensional solutions are robust to increased dimensionality and directly relates the two dimensional solution to the one dimensional ones. For a close to 0, i.e. the range of wages is large, the solution is similar to that of IMP_w. As a approaches

b, the range of wages decreases, collapsing the problem into the one dimensional problem of IMP_θ. It also provides substantial intuition about why the work requirements will arise - with little variance in ability/wages, distorting work downwards to save on incentives can be very costly. This leads to the work requirements type solution of IMP_θ. With a large distribution of wages the incentive problem of keeping high ability people from deviating dominates, leading to the absence of work requirements. This offers a sense that first, different economies should utilize different poverty assistance programs, and second, that depending on the population targeted within one country, different programs could and should be constructed in different ways.

Figure 4 plots $IMP_{w\theta}$ given a uniform distribution, $z = .2$, and ranges of w and θ of $[0, 2]$ and $[0, 1]$ respectively. The program itself resembles a Negative Income Tax, with a guaranteed amount of income for low types, and a variable BRR that depends on population characteristics.

7 Conclusion

Large-scale poverty assistance programs in the United States focus on raising the income of participants. Temporary Assistance to Needy Families (which replaced AFDC), Supplemental Security Income, and the Earned Income Tax Credit augment participants' income, while Medicaid and Food Stamps decrease the cost of goods that are deemed important to participants. The push to guarantee some minimum standard of living has been linked to purchasing power (utility has been neglected for both philosophical and feasibility reasons) and therefore income. This paper explores properties of programs with an explicit objective of income maintenance.

Several conclusions emerge from the analysis. First of all, *program construction should be sensitive to the environment it is created in.* When types can be observed, we find that a conditional income subsidy scheme is optimal. If only disutility of labor can be observed, guaranteed amounts should be offered to those who don't work. If only ability can be observed, then a minimum work requirement may be added in order for a participant to receive any transfer. If both are unobservable, the solution may use work requirements or may not, depending on the range of abilities of program participants and the consequent trade-off between incentive savings and costly distortion of work. Such programs as Temporary Assistance for Needy Families and Supplemental Security Income operate among populations where some signals about types are available, making a comparison between them and our solution informative. With the variance in population characteristics and TANF programs across states, our results establish a framework for the comparison of incentive effects between states. Indeed, we observe that many states have work requirements in addition to varying BRRs, and that these work requirements vary between 18 and 40 hours per week.

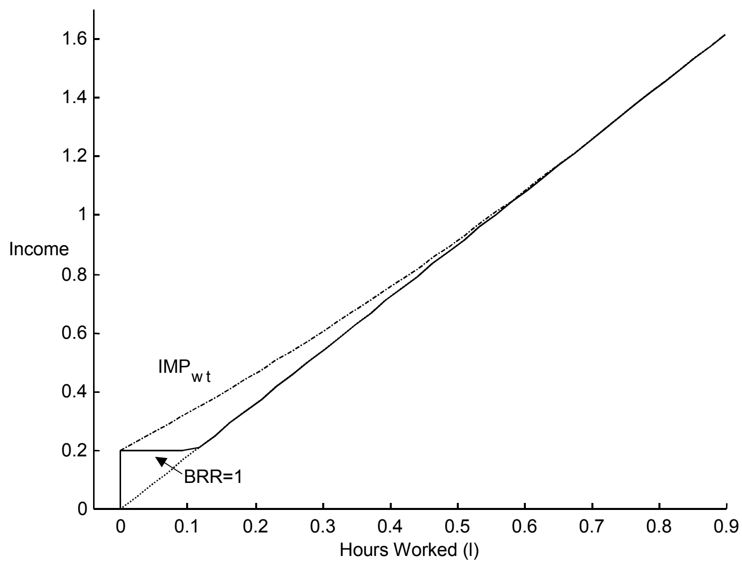

Figure 5: Budget constraint of an individual with type $w = 1.8$ under different programs: A NIT program with the BRR=1, and $IMP_{w\theta}$ (with $a = 0$)

Second of all, *the optimal schedules resemble a Negative Income Tax.* A strict NIT consists of a guaranteed amount of income for no work, with low amounts of income taxed at a fixed Benefit Reduction Rate. The schedules for all three environments designate a BRR that is endogenous to the problem since it depends on population characteristics and changes with hours worked. In Figure 5, we draw the budget constraints that an individual of type $w = 1.8$ faces under a program with a BRR of 1 (such as the old AFDC) and under $IMP_{w\theta}$, with $a = 0$. This emphasizes both the similarities and differences between actual programs and our solution. It may be that fixed BRRs less than one can approximate the optimal schedule, a topic that merits further investigation.

Third, by introducing workfare into two type single dimensional environments, we highlight that *workfare might not be optimal.* The unobservable income generating ability world and the unobservable disutility of labor world provide different incentive structures for imposing work requirements, making workfare's use as a screening mechanism tenuous.

We do not address the dynamic elements of poverty assistance programs in this paper. Such values as building skills or instilling a work ethic can't be included in a static framework. Nevertheless, short-term effectiveness can be a valuable barometer if one's discount rate is relatively high or if one considers that the impact of the dynamic argument is difficult to measure and assess.

The two dimensional model presents an application of the relatively new multidimensional screening theoretical work. While this represents one of the few continuous models solved, simplification hinders its overall generality. The specification of $h(l,\theta) = (l+\theta)^2$ offered a tractable way of eliminating endogenous isolabor curves, but a solution concept that takes into account such curves would increase the robustness of the result. Another possible way to relax the model would be to eliminate the quasilinearity of the utility function.

An important way to extend the model would be to relax the exogeneity of types. This would incorporate the dynamic element lacking in the analysis. Clearly, separate programs that provide day care, medical support, and training can reduce the heterogeneity in ability and disutility of labor, and therefore reduce the incentive problem and costs. Finally, understanding how equilibrium unemployment impacts optimal program design, and vice-versa would be valuable. This is treated simply in the paper, but one can imagine a model in which search frictions determine reservation utility and available choices.

Appendix

A Proof of Lemma 1

The constrained maximization is:

$$\max_{l_L, l_H, c_L, c_H, t_L, t_H} -(\rho t_L + (1-\rho)t_H)$$

where l_i, c_i, and t_i are hours worked, hours of workfare, and transfers for type i (where $i = L, H$) and ρ is the percentage of the population that is type L. w will represent the fixed income generating ability and $\theta_L(\theta_H)$ represents the low (high) disutility of labor type. r will stand for reservation utility and z will be the minimum income level (we assume $z > r$). We use a more general function for disutility of labor than in the example to emphasize the robustness of the result: $h(l,\theta)$ where $h_l\left(\right), h_{ll}\left(\right), h_{l\theta}\left(\right) \geq 0$.

the constraints are:

$$wl_i - h(l_i + c_i, \theta_i) + t_i \geq r \qquad \text{for } i = L, H \ (\lambda_1, \lambda_2) \qquad \text{(Individual Rationality)}$$

$$wl_i - h(l_i + c_i, \theta_i) + t_i \geq wl_{-i} - h(l_{-i} + c_{-i}, \theta_i) + t_{-i} \quad \text{for } i = L, H \ (\lambda_3, \lambda_4\ ) \qquad \text{(Incentive Compatibility)}$$

$$wl_i + t_i \geq z \qquad \text{for } i = L, H \ (\lambda_5, \lambda_6) \qquad \text{(Income Maintenance)}$$

and non-negativity constraints $l_L \geq 0$ (λ_7), $l_H \geq 0$ (λ_8), $c_L \geq 0$ (λ_9), $c_H \geq 0$ (λ_{10}), and . Note that the λs in parentheses denote the Lagrange multipliers for the associated constraint. It is easy to show that $\lambda_1 = 0$ using both Individual Rationality constraints and L's Incentive Compatibility constraint. The necessary conditions for a solution are:

1. $-(1-\rho) + \lambda_2 - \lambda_3 + \lambda_4 + \lambda_6 = 0$
2. $-\rho + \lambda_3 - \lambda_4 + \lambda_5 = 0$
3. $\lambda_2(w - h_l(l_H + c_H, \theta_H)) - \lambda_3(w - h_l(l_H + c_H, \theta_L)) + \lambda_4(w - h_l(l_H + c_H, \theta_H)) + \lambda_6 w + \lambda_8 = 0$
4. $\lambda_3(w - h_l(l_L + c_L, \theta_L)) - \lambda_4(w - h_l(l_L + c_L, \theta_H)) + \lambda_5 w + \lambda_7 = 0$
5. $\lambda_2(-h_l(l_H + c_H, \theta_H)) - \lambda_3(-h_l(l_H + c_H, \theta_L)) + \lambda_4(-h_l(l_H + c_H, \theta_H)) + \lambda_{10} = 0$
6. $\lambda_3(-h_l(l_L + c_L, \theta_L)) - \lambda_4(-h_l(l_L + c_L, \theta_H)) + \lambda_9 = 0$

By combining conditions 4 and 6 we get $(\lambda_3 - \lambda_4 + \lambda_5)w - \lambda_9 + \lambda_7 = 0$. Using condition 2, it is clear that $\lambda_9 > 0$. Similarly, by combining conditions 3 and 5, and using condition 1, we find that $\lambda_{10} > 0$. These results imply that $c_L = c_H = 0$ which proves our lemma.

B Conditions for Proposition 1 (IMP_w)

The Lagrangian is formed as follows:

$$L = H + \gamma R + \beta u + \mu(R + \frac{w^2}{4\theta} + \theta l^2 - z) + \omega l$$

The necessary conditions are:

1. $\lambda'(w) = f(w) - \gamma - \mu$
2. $\alpha'(w) = (-w + 2\theta l)f(w) - \lambda - \mu 2\theta l - \omega$
3. $\alpha + \beta = 0$
4. $\gamma R = 0$; $\gamma \geq 0$; $R \geq 0$
5. $\beta u = 0$; $\beta \geq 0$; $u \geq 0$
6. $\mu(R + \frac{w^2}{4\theta} + \theta l^2 - z) = 0$; $\mu \geq 0$; $R + \frac{w^2}{4\theta} + \theta l^2 - z \geq 0$
7. $\omega l = 0$; $\omega \geq 0$; $l \geq 0$
8. $\lambda(0)R(0) = 0$; $\lambda(0) \leqslant 0$; $\lambda(b)R(b) = 0$; $\lambda(b) \geq 0$
9. $\alpha(0)l(0) = 0$; $\alpha(0) \leqslant 0$; $\alpha(b)l(b) = 0$; $\alpha(b) \geq 0$

Conditions 1 and 2 are the costate equations. Condition 3 is the optimality condition. Conditions 4-7 are the complementary slackness equations. Conditions 8 and 9 are the transversality conditions. Since we have state constraints that may bind, we must allow for jumps in the costate variables. The jump conditions are:

1.$\lambda(w^-) - \lambda(w^+) = \eta_1 + \eta_2$

2. $\alpha(w^-) - \alpha(w^+) = \eta_1 2\theta l + \eta_3$

where η_1, η_2, and η_3 are non-negative and represent the income maintenance constraint, the voluntary participation constraint and the $l \geq 0$ constraint respectively.

Consider the solution in Proposition 1. If $l(0) = 0$, then $\alpha(0^+)$ can equal a number $Y \leqslant 0$. If IM binds at $w = 0$ (and since we know VP will not bind there), $\lambda(0^+) = -\eta_1$.

Let $\eta_1 = F(w')$. So $\lambda(w)$ for low types equals $F(w) - F(w')$. Now denote $\frac{w}{2\theta} - \frac{F(w')-F(w)}{2\theta f(w)}$ by $\tilde{l}(w)$. $\tilde{l}(w)$ makes $\alpha'(w)$ equal 0. Since $\tilde{l}(w) < 0$ for low values of w, the non-negative labor condition makes $\alpha'(w) > 0$ for these values. We can now define w_0 by $\tilde{l}(w_0) = 0$. In effect w_0 is a function of w'. To solve for w', we must first tie up loose ends. By integrating the second costate equation and plugging in the endpoint constraint we get $\alpha(w) = -w(F(w) - F(w')) + Y$. There are two unknowns: w' and Y. We can solve using the two equations $\alpha(w_0(w')) = 0$ and VP'. Note that $\tilde{l}(w') = l^*(w')$. If w' exists, $[w', b]$ is the region that lemma 1 addresses. $\lambda(w) = 0$ and $\gamma = f(w)$ on this region. If w' satisfying the previous equations doesn't exist (VP' doesn't bind) we set $w' = b$.

C Conditions for Proposition 2 (IMP_θ)

The constraints upon the problem are VP_θ, IM, IC_θ (which gives us the law of motion $R_\theta = -l^2 + \frac{w^2}{4\theta^2}$), $\frac{\partial l}{\partial \theta} = u(\theta) \leqslant 0$, and $l \geq 0$. As before, the state variables are R and l, the associated costate variables are λ and α, and the control is u.

The Hamiltonian is:

$$H(R, l, u, \lambda, \alpha, \theta) = \{-R - \frac{w^2}{4\theta} + wl - \theta l^2\}p(\theta) + \lambda(-l^2 + \frac{w^2}{4\theta^2}) + \alpha u$$

The Lagrangian is:

$$L = H + \gamma R - \beta u + \mu(R + \frac{w^2}{4\theta} + \theta l^2 - z) + \omega l$$

The necessary conditions are:

1. $\lambda'(\theta) = p(\theta) - \gamma - \mu$
2. $\alpha'(\theta) = (-w + 2\theta l)p(\theta) + \lambda 2l - \mu 2\theta l - \omega$
3. $\alpha - \beta = 0$
4. $\gamma R = 0$; $\gamma \geq 0$; $R \geq 0$
5. $\beta u = 0$; $\beta \geq 0$; $u \leqslant 0$
6. $\mu(R + \frac{w^2}{4\theta} + \theta l^2 - z) = 0$; $\mu \geq 0$; $R + \frac{w^2}{4\theta} + \theta l^2 - z \geq 0$
7. $\omega l = 0$; $\omega \geq 0$; $l \geq 0$
8. $\lambda(\theta_a)R(\theta_a) = 0$; $\lambda(\theta_a) \leqslant 0$; $\lambda(\theta_b)R(\theta_b) = 0$; $\lambda(\theta_b) \geq 0$
9. $\alpha(\theta_a)l(\theta_a) = 0$; $\alpha(\theta_a) \leqslant 0$; $\alpha(\theta_b)l(\theta_b) = 0$; $\alpha(\theta_b) \geq 0$

Conditions 1 and 2 are the costate equations. Condition 3 is the optimality condition. Conditions 4-7 are the complementary slackness equations. Conditions 8 and 9 are the transversality conditions. Since we have state constraints that may bind, we must allow for jumps in the costate variables. The jump conditions are:

1. $\lambda(\theta^-) - \lambda(\theta^+) = \eta_1 + \eta_2$
2. $\alpha(\theta^-) - \alpha(\theta^+) = \eta_1 2\theta l + \eta_3$

where η_1, η_2, and η_3 are non-negative and represent the income maintenance constraint, the voluntary participation constraint and the $l \geq 0$ constraint respectively.

Assuming that $\bar{l}(\theta_b) \leqslant l^*(\theta_a)$ (the opposite case is examined below), we conjecture that at $\theta = \theta_b$, both VP_θ and IM bind, making $l(\theta_b) = \bar{l}(\theta_b)$. We can say that $\lambda(\theta_b^+) = 0$, $\eta_1 = Y$, and $\eta_2 = P(\theta') - Y$, making $\lambda(\theta) = P(\theta') - P(\theta)$ for high types and $\alpha(\theta_b^-) = Y2\theta_b\bar{l}(\theta_b)$. It is clear from the jumps in the costate variables that $0 \leqslant Y \leqslant P(\theta')$. Denote $\frac{w}{2(\frac{P(\theta')-P(\theta)}{p(\theta)}+\theta)}$ by $\tilde{l}(\theta)$. We can solve for θ_0 as a function of θ' by setting $\tilde{l}(\theta_0) = \bar{l}(\theta_b)$. By integrating the second costate equation and using an endpoint constraint, we can show that $\alpha(\theta) = wP(\theta) - 2\theta l(P(\theta) - P(\theta')) - 2\theta_b\bar{l}(\theta_b)(P(\theta') - Y)$. Since $\tilde{l}(\theta) < \bar{l}(\theta_b)$ for all $\theta \in (\theta_0, \theta_b]$, we know that $\alpha'(\theta)$ is positive on that interval. If we solve $\alpha(\theta_0) = 0$ and $R(\theta') = 0$ simultaneously, we will find a solution for Y and θ'. Note that $\tilde{l}(\theta') = l^*(\theta')$. If we can't find a θ' where VP_θ binds, we set $\theta' = \theta_b$. In a previous version of this paper, we prove that $\alpha(\theta_0)$ can equal 0 for the restricted range of Y.

We now consider the case where $l^*(\theta_a) < \bar{l}(\theta_b)$. It is straightforward to show that the solution is $l(\theta) = \bar{l}(\theta_b)$ for all θ, given the structure put on the problem above. θ_0 is no longer a relevant concept, and θ' is set to θ_a. Setting $Y = 1 - (\frac{w}{2\theta_b})/\bar{l}$ satisfies the constraints on Y (namely $0 \leqslant Y \leqslant 1$) and completes the system.

D Review of the Generalized Single Crossing Property and Its Implications

We first define the marginal rate of substitution between labor and transfers $p(l,(w,\theta)) = \frac{U_l(w,\theta)}{U_t(w,\theta)}$. The Generalized Single Crossing Property (McAfee and McMillan (1988)) is then defined as

$$p(l,(w,\theta)) - p(l,(w',\theta')) = \lambda \begin{bmatrix} p_w(l,(w',\theta')) & p_\theta(l,(w',\theta')) \end{bmatrix} \begin{bmatrix} w - w' \\ \theta - \theta' \end{bmatrix} \qquad (GSC)$$

where $\lambda > 0$. For the quasilinear utility employed in $IMP_{w\theta}$, GSC is satisfied and $\lambda = 1$. The mechanism is described as a pair (l,t), or indirectly, $t(l)$. We rewrite the individual's utility function: $U(l,(w,\theta)) = wl - h(l,\theta) + t(l)$. The individual must choose according to $U_l = 0$, or $p(l,(w,\theta)) + t'(l) = 0$. This implies that l (and therefore t) are constant for all w', θ' such that $p(l,(w',\theta')) = p(l,(w,\theta))$, making the right-hand side of GSC equal to 0 for all such w' and θ'. It is now possible to define isolabor curves in (w,θ) space (with slope equal to $\frac{-p_\theta}{p_w}$). We index these curves by the variable s, which denotes where the relevant isolabor curve $l(s)$ intersects with the w-axis (when $\theta = 0$). Letting $h(l,\theta) = (l+\theta)^2$ yields the simple transformation $w = s + 2\theta$.

E Conditions for Proposition 3 ($IMP_{w\theta}$)

With a three phase optimization program, the necessary conditions are simply the necessary conditions for each phase plus some transition point conditions. These are extended to include jumps in the costate variables.

We first form the Hamiltonians and append the constraints to construct Lagrangians:

$$L_1 = \{-V+sl-l^2\}g_1(s)+\lambda_1(l)+\alpha_1 u+\gamma_1 V+\beta_1 u+\mu_1(V+l^2+(a-s)l-z)+\omega_1 l \text{ for } s \in [a-2\theta_b, 0]$$

$$L_2 = \{-V+sl-l^2\}g_2(s)+\lambda_2(l)+\alpha_2 u+\gamma_2(V-\frac{1}{4}s^2)+\beta_2 u+\mu_2(V+l^2+(a-s)l-z)+\omega_2 l \text{ for } s \in [0, a]$$

$$L_3 = \{-V+sl-l^2\}g_3(s)+\lambda_3(l)+\alpha_3 u+\gamma_3(V-\frac{1}{4}s^2)+\beta_3 u+\mu_3(V+l^2-z)+\omega_3 l \text{ for } s \in [a, b]$$

Without jumps in the costate variables, the transition point conditions are $\lambda_1(0^-) = \lambda_2(0^+)$, $\alpha_1(0^-) = \alpha_2(0^+)$ and $H_1(0) = H_2(0)$ for 0 and $\lambda_1(a^-) = \lambda_2(a^+)$, $\alpha_1(a^-) = \alpha_2(a^+)$ and $H_1(a) = H_2(a)$ for a. With jumps, however, (using Leonard and Long (1992) and Amit (1986)) the new conditions are:

A1. $\lambda_1(0^-) - \lambda_2(0^+) = \psi_1 + \psi_2$
B1. $\alpha_1(0^-) - \alpha_2(0^+) = \psi_1(2l + a - s) + \psi_3$
C1. $H_1(0) = H_2(0)$

A2. $\lambda_2(a^-) - \lambda_3(a^+) = \psi_1 + \psi_2$
B2. $\alpha_2(a^-) - \alpha_3(a^+) = \psi_1(2l) + \psi_3$
C2. $H_2(a) = H_3(a)$

where, with slight abuse of notation, ψ_1 is associated with the IM constraint, ψ_2 is associated with the VP constraint, and ψ_3 is associated with the $l \geq 0$ constraint. In addition, we need jump conditions for the initial point $a - 2\theta_b$. These conditions are exactly analogous to A1 and B1.

The other necessary conditions are very similar to the conditions necessary to propositions 1 and 2. We list all conditions except the non-negative ones:

For $s \in [a - 2\theta_b, 0)$

1. $\lambda_1'(s) = g_1(s) - \gamma_1 - \mu_1$
2. $\alpha_1'(s) = (-s + 2l)g_1(s) - \lambda_1 - \mu_1(2l + a - s) - \omega_1$
3. $\alpha_1(s) + \beta_1(s) = 0$

For $s \in (0, a)$

1. $\lambda_2'(s) = g_2(s) - \gamma_2 - \mu_2$
2. $\alpha_2'(s) = (-s + 2l)g_2(s) - \lambda_2 - \mu_2(2l + a - s) - \omega_2$
3. $\alpha_2(s) + \beta_2(s) = 0$

For $s \in (a, b]$

1. $\lambda_3'(s) = g_3(s) - \gamma_3 - \mu_3$
2. $\alpha_3'(s) = (-s + 2l)g_3(s) - \lambda_3 - \mu_3(2l) - \omega_3$
3. $\alpha_3(s) + \beta_3(s) = 0$

There are two candidate solutions that satisfy the necessary conditions for certain parameters. We will go through the derivation for each of them.

Schedule 1: For the first phase $[a-2\theta_b, 0]$, we set $\alpha_1(a-2\theta_b^-) = A \leq 0$ since $l(a-2\theta_b) = 0$. Since $V(0) > 0$, $\lambda_1(a-2\theta_b^-) = 0$. Setting $\psi_1 = G_i(s')$, where i is determined jointly with s', makes $\lambda_1(s) = G_1(s) - G_i(s')$ and $\alpha_1(a-2\theta_b^+) = A - 2\theta_b G(s')$. This makes $\alpha_j'(s) > 0$ and $\alpha_j(s) = s(G_i(s') - G_2(s)) + A - aG_i(s')$ (noting that $j \leq i$), and we set A so that $\alpha_j(s_0) = 0$. We solve for s_0 as a function of s' using $0 = \frac{s_0}{2} + \frac{G_j(s_0) - G_i(s')}{2g_j(s_0)}$ and for s' using the fact that VP binds at s'. Since there are no jumps at the transition points, all of the transition conditions hold. The critical equation necessary for examining optimality is whether there a value of A such that $\alpha_j(s_0) = 0$. The only restriction on A is that it is less than or equal to 0. It is quick to show that the restriction follows from the equation if $a = 0$ and does not if $a = b$. Through implicit differentiation of the cutoff equations ($0 = \frac{s_0}{2} + \frac{G_j(s_0) - G_i(s')}{2g_j(s_0)}$ and VP) we can find the dependence of s_0 and s' on a and show that $\alpha_j(s_0)$ is monotonely decreasing in a. This implies that the necessary conditions are satisfied for low values of a. When the necessary conditions, are satisfied, they are sufficient.

Schedule 2: Since $l(a-2\theta_b) = \bar{l}(a-2\theta_b)$, $\alpha_1(a-2\theta_b^-) = 0$. Since both IM and VP bind, there can be a jump in the costate variables at $a-2\theta_b$. Set $\psi_1 = Y$, $\psi_2 = G_i(s') - Y$ (again i is defined as the interval which includes s' and $j \leq i$). Using these endpoint restrictions and the necessary conditions, we set $\lambda_j(s) = G_j(s) - G_i(s')$ and for $s \leq s_0$, $\alpha_j(s) = (-s + 2\bar{l})G_j(s) + (s - a + 2\theta_b)G_i(s') - Y(2\bar{l} + 2\theta_b)$. The points s_0 and s' can be found from setting $\bar{l}(a-2\theta_b) = \frac{s_0}{2} + \frac{G_j(s_0) - G_i(s')}{2g_j(s_0)}$ and from VP binding at s'. All of the other necessary conditions are satisfied. It remains to prove that there can exist a Y such that $\alpha(s_0) = 0$. From this equation $Y = \frac{(-s_0 + 2\bar{l})G_j(s_0) + (s_0 - a + 2\theta_b)G_i(s')}{2\bar{l} + 2\theta_b}$, and from the conditions for costate jumps, we place the restrictions $0 \leq Y \leq G_i(s')$. It is easy to show that $Y \geq 0$ always holds. Using the fact that $\frac{s}{2} = l^*(s)$, we can also prove that for $a = b$, $Y < G_i(s')$, which demonstrates that collapsing the dimensions does indeed yield the IMP_θ solution. Since the inequality is strict and for a close to b the expression $Y - G_i(s')$ can be shown to be decreasing in a, there is a range of wages close to b where Schedule 2 is the solution. Additionally, we can show that when $a = 0$, $Y > G(s')$. As in IMP_θ, sufficiency for this solution is restricted. A sufficient condition for sufficiency is if $G_i(s')(\bar{l}^2 - (s_0 - a)\bar{l}) - G_j(s_0)(-s_0\bar{l} + 2\bar{l}^2) \geq 0$.

References

[1] AMIT, R. (1986). "Petroleum Reservoir Exploration: Switching from Primary to Secondary Recovery", *Operations Research*, **34**, 534-549.

[2] ARMSTRONG, M. (1996). "Multiproduct Nonlinear Pricing", *Econometrica*, **64**, 51-75.

[3] ARMSTRONG, M. (1999). "Optimal Regulation with Unknown Demand and Cost Functions", *Journal of Economic Theory*, **84**, 196-215.

[4] BEAUDRY, P. and BLACKORBY, C. (1998). "Taxes and Employment Subsidies in Optimal Redistribution Programs", NBER Working Paper #6355.

[5] BESLEY, T. and COATE, S. J. (1992). "Workfare vs. Welfare: Incentive Arguments for Work Requirements in Poverty Alleviation Programs", *American Economic Review*, **82**, 249-261.

[6] BESLEY, T. and COATE, S. J. (1995). "The Design of Income Maintenance Programmes", *Review of Economic Studies*, **62**, 187-221.

[7] BLANK, R. M. (1997). *It Takes a Nation* (Princeton: Princeton University Press).

[8] CUFF, K.(2000). "Optimality of Workfare with Heterogeneous Preferences", *Canadian Journal of Economics,* **33**, 149-74.

[9] EDIN, K. and LEIN, L. (1997). *Making Ends Meet* (New York: Russell Sage Foundation).

[10] GALLAGHER, L. J., GALLAGHER, M., PERESE, K., SCHREIBER, S. and WATSON, K. (1998). "One Year After Federal Welfare Reform: A Description of State Temporary Assistance for Needy Families (TANF) Decisions as of October 1997" (Washington D.C.: The Urban Institute).

[11] GUESNERIE, R. and LAFFONT, J.-J. (1984). "A Complete Solution to a Class of Principal-Agent Problems with an Application to the Control of a Self-Managed Firm", *Journal of Public Economics*, **25**, 329-369.

[12] HOYNES, H. W. (1996). "Welfare Transfers in Two Parent Families: Labor Supply and Welfare Participation Under AFDC-UP", *Econometrica,* **64**, 295-332.

[13] JULLIEN, B. (2000). ´´Participation Constraints in Adverse Selection Models", *Journal of Economic Theory*, **93**, 1-47.

[14] LAFFONT, J.-J., MASKIN, E. and ROCHET, J.-C. (1987). "Optimal Nonlinear Pricing with Two-Dimensional Characteristics". T. Groves, R. Radner, and S. Reiter eds., *Information Incentives and Economic Mechanisms* (Minneapolis: University of Minnesota Press).

[15] LÉONARD, D. and LONG, N. V. (1992). *Optimal Control Theory and Static Optimization in Economics*, (New York: Cambridge University Press).

[16] LEWIS, T. R. and SAPPINGTON, D.E.M. (1989a). "Countervailing Incentives in Agency Problems", *Journal of Economic Theory*, **49**, 294-313.

[17] LEWIS, T. R. and SAPPINGTON, D.E.M. (1989b). "Inflexible Rules in Incentive Problems", *American Economic Review*, **79**, 69-84.

[18] LEWIS, T. R. and SAPPINGTON, D.E.M. (1988). "Regulating a Monopolist with Unknown Demand and Cost Functions", *RAND Journal of Economics*, **19**, 438-457.

[19] MAGGI, G. and RODRÍGUEZ-CLARE, A. (1995). "On Countervailing Incentives", *Journal of Economic Theory*, **66**, 238-263.

[20] McAFEE, R. P. and McMILLAN, J. (1988). "Multidimensional Incentive Compatibility and Mechanism Design", *Journal of Economic Theory*, **46**, 335-354.

[21] MIRRLEES, J. A. (1971) "An Exploration in the Theory of Optimum Income Taxation", *Review of Economic Studies*, **38**, 175-208.

[22] MIRRLEES, J. A. (1997), "Optimal Marginal Tax Rates at Low Incomes", mimeo, University of Cambridge.

[23] MOFFITT, R. (1992). "Incentive Effects of the U.S. Welfare System: A Review", *Journal of Economic Literature*, **XXX**, 1-61.

[24] MYERSON, R. (1981). "Optimal Auction Design", *Mathematics of Operations Research*, **6**, 58-73.

[25] ROCHET, J.-C. and CHONÉ, P.(1998). "Ironing, Sweeping and Multidimensional Screening", *Econometrica*, **66**, 783-826.

[26] ROCHET, J.-C. and STOLE, L. A. (2000). "The Economics of Multidimensional Screening," mimeo.

[27] SEIERSTAD, A. and SYDSAETER, K. (1987). *Optimal Control Theory with Economic Applications*, (New York: North-Holland).

[28] SWARNS, R. L. (September 15, 1998). "Hispanic Mothers Lagging as Others Leave Welfare", *The New York Times*, A1.

[29] THE URBAN INSTITUTE. (1996). "Profile of Disability Among AFDC Families", Policy and Research Report, Summer-Fall, http://www.urban.org/periodcl/26_2/prr26_2d.htm

[30] ZEDLEWSKI, S. (1999). "Work-Related Activities and Limitations of Current Welfare Recipients", The Urban Institute, Washington D.C.