4 Welfare Effects of Tax Reform and Labor Supply at the Intensive and Extensive Margins

Nada Eissa, Henrik Jacobsen Kleven, and Claus Thustrup Kreiner

4.1 Introduction

A central finding of the modern empirical labor literature is that labor supply responses tend to be concentrated more along the extensive margin (labor force participation) than along the intensive margin (hours of work). Evidence from expansions of tax-based transfers in the United States and the United Kingdom seems to indicate substantial effects on labor force participation, but only modest effects on hours of work for those who are working (Eissa and Liebman 1996; Meyer and Rosenbaum 2001; Blundell 2001). These recent findings support direct and indirect evidence from earlier work on labor supply. Direct evidence from the Negative Income Tax (NIT) experiments in the United States shows that the average participation response from all experiments was slightly larger than the annual hours-of-work response for both single female heads and married women (Robins 1985). For married women, Mroz (1987) and Triest (1990) estimate larger wage and income elasticities when including all (working and nonworking) women than when including only working women, suggesting that the participation margin is more sensitive to taxes than hours worked by working women.

This chapter argues that the concentration of labor supply responsiveness along the extensive margin carries important implications for the welfare evaluation of tax reform. The theoretical public finance literature has largely ignored the participation margin and instead has focused on labor supply at the intensive margin. An important exception is Saez 2002, which extends the theory of optimal taxation to account for both margins of labor supply response. Saez demonstrates that if participation elasticities are relatively high at the bottom of the earnings distribution—as indeed seems to be the case—the optimal

policy subsidizes low-income earners. Interestingly, this result is broadly in line with the recent tax reforms in the United States and the United Kingdom, which have incorporated large subsidies for low-income workers. The optimality approach is limited in the evaluation of actual tax reforms, however. First, this approach deals with the properties of tax systems where all taxes and transfers have been optimized. This never occurs, since actual reforms change specific elements of the tax code and leave other, potentially inefficient, elements unchanged. Second, optimal taxation depends on the magnitude of elasticities at the optimal point, which is presumably different from the current position, or anything else previously observed. The evaluation of actual tax reforms, on the other hand, depends only on behavioral responses in the current equilibrium.

To understand the implications of the participation decision for evaluating tax reform, we embed the extensive margin in a simple welfare theoretic framework. A central issue is whether to model the extensive response within a standard convex labor supply model or within a model incorporating non-convexities in budget sets or preferences. In the convex framework, labor force participation is determined by a reservation wage and is nondiscrete at the individual level. Thus, a tax reform lowering the marginal tax rate by a small amount at or around the reservation wage level induces some individuals to enter the labor market at few working hours. In the chapter we show that extensive labor supply responses of this type create only second-order welfare effects, and are therefore inconsequential for the welfare evaluation of (small) tax reforms. Moreover, the kind of participation responses predicted by the standard convex model is inconsistent with the empirical evidence. Empirical evidence shows that almost no workers choose to enter the labor market at very small hours of work.

A realistic model must account for discrete participation behavior, where people enter the labor market at some minimum hours of work, say, twenty or thirty hours per week. This type of behavior is typically explained by non-convexities created by work costs (Cogan 1981; Heim and Meyer 2003). Based on Eissa, Kleven, and Kreiner (2004), we outline a simple framework incorporating fixed costs of work, and we show that extensive labor supply responses generate first-order welfare effects within such a framework. The extensive welfare effect from tax reform depends on the size of the participation elasticity, and it depends on the pre-reform levels of tax rates and on the changes in tax rates introduced by the reform. In particular, the welfare effect

from participation responses is related to the effective *average* tax rate, including the average benefit reduction rate for welfare benefits. Due to nonlinearities and discontinuities in taxes and transfers, the average tax rate may be significantly different from the marginal tax rate associated with the intensive margin of labor supply.

It is exactly because of nonlinearities in the tax transfer system and the implied difference in average and marginal tax rates that it is necessary to distinguish between the intensive and extensive margins in the welfare analysis. In the case of linear taxes and transfers, we argue that there is no loss from conflating the two margins in a total labor supply response. But nonlinearities are significant in practice, especially at the lower end of the distribution where welfare benefits and tax-based transfers play an important role. In the United States, for example, the Earned Income Tax Credit (EITC) implies a marginal tax rate equal to −36 percent in the phase-in range (in 2002), while the marginal tax rate implied by the phase-out of the credit is +21 percent. Since most individuals eligible for the EITC have income beyond the phase-in range, the credit tends to reduce incentives to supply labor along the intensive margin. At the same time, the program reduced average tax liabilities and therefore average tax rates, thereby improving the returns to work. In the welfare evaluation of an EITC reform, it is important to distinguish explicitly between intensive and extensive responses and to account for the different incentives created along the two margins. Moreover, even if one were to consider a linear tax reform (say, a proportional tax change), one needs to distinguish between the two margins of response as long as the pre-reform tax system involves nonlinearities.

Finally, we consider an empirical application to the Tax Reform Act of 1986 (TRA86) in the United States. This reform constituted the most fundamental change in the federal income tax system in nearly forty years. By design, the reform had its largest impact on high-income earners, but it also included substantial benefits at the lower end of the distribution. Our concern is with the welfare effects for single mothers, a group characterized by low earnings and long spells of non-employment. This group experienced large tax cuts from the reform due to an expansion of the EITC, and because of increases in the standard deduction, personal exemptions, and a more favorable tax rate schedule. The combination of all these elements implied substantial improvements in the incentives to supply labor along both the intensive and extensive margins. We show that the reform created

substantial welfare gains for single mothers and that most of the effect was generated by positive participation responses. Hence the application serves to confirm that large errors can be made by omitting the extensive response in the welfare analysis of tax reform.

The chapter is organized in the following way. Section 4.2 discusses the theoretical analysis of tax reform. Section 4.3 describes a number of tax reforms passed in the United States over the past two decades, while section 4.4 reviews the empirical evidence on the effects of tax reform on labor supply. Section 4.5 presents microsimulation results for the welfare effects of TRA86, and section 4.6 concludes.

4.2 A Theoretical Analysis of Tax Reform

4.2.1 Labor Supply Responses

In general, tax reforms may affect labor supply along both the intensive and the extensive margins. Along the intensive margin, a tax reform that changes the marginal net wage induces employees to adjust their weekly hours of work from, say, forty to thirty-nine hours. At the same time, a tax reform may create extensive responses by affecting the incentive to participate in the labor market. This effect is likely to be particularly relevant for certain subgroups of the population such as married females, single mothers, low-educated individuals, the young, and the elderly (retired). For these individuals, higher tax burdens may make it worthwhile to leave the labor market entirely, going from, say, thirty or twenty hours per week to zero hours per week.

The bulk of the literature on taxation and labor supply is based on the standard static labor supply model, where preferences and budget sets are assumed to be convex. This model underlies Mirrlees' (1971) exploration into optimal income taxation. The framework has subsequently been used to examine the welfare cost of taxation by Ballard, Shoven, and Whalley (1985), Ballard (1988), Triest (1994), Browning (1995), and Dahlby (1998). In the convex model, optimal labor supply is given by the tangency of the budget line and the indifference curve, implying that marginal changes in prices and endowments give rise to marginal changes in individual behavior. In other words, the model is one of continuous choice.

Figure 4.1 illustrates the choice of labor supply in the standard model. In the figure, we consider two individuals facing the same budget constraint, where y_0 is nonlabor income and $(1-t)w$ is the mar-

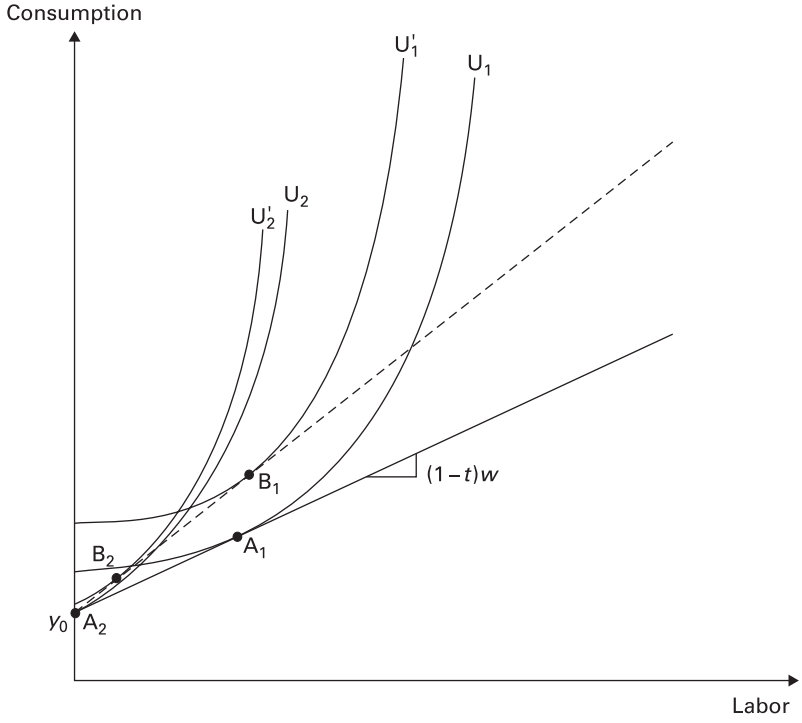

Figure 4.1
Intensive versus extensive responses in the convex model

ginal net-of-tax wage. The indifference curves of the two individuals are drawn such that individual 1 has a relatively low valuation of leisure, while individual 2 has a relatively high valuation of leisure. Consequently, it is optimal for the first individual to work many hours, while the second individual chooses to stay outside the labor market since there is no point of tangency at positive hours. Consider now a reduction in the tax rate causing an increase in the slope of the budget line. This induces individual 1 to increase his or her hours of work a little bit (an intensive response). For individual 2, on the other hand, the higher net wage gives rise to an interior solution such that he or she decides to enter the labor market (an extensive response). However, the extensive response involves a change in labor supply from zero hours to some small (infinitesimal) number of hours. Thus, the type of participation response predicted by this framework is a marginal one, just like the change in hours of work for those who are working.

As already indicated, this may be misleading in the case of labor supply behavior, where discrete jumps seem to be empirically important. In particular, discreteness is significant in connection with the labor force participation decision. We do not observe that people enter the labor market at infinitesimal hours of work, but that they do so at, say, twenty or forty hours. Therefore, to be able to realistically capture labor supply behavior at the extensive margin, we need to employ a different framework. In particular, some type of non-convexity is required.

A way to introduce a non-convexity into the analysis would be to allow for fixed costs of working. In a well-known paper, Cogan (1981) proposed a model with fixed costs of working to explain discrete labor supply behavior. In Cogan's analysis, the fixed costs may be monetary costs (say, child care and transportation expenses) or they may take the form of a loss of time (e.g., commuting time). In figure 4.2, we extend the analysis of labor supply choice along these lines. An individual who chooses to stay outside the labor market receives nonlabor income y_0. If he decides to enter the labor market, he loses Δy_f in income and Δl_f in leisure time upon entry, thereby creating a discontinuity in the budget set. In the initial situation, it is still the case that individual 1 works relatively long hours, while individual 2 does not work at all. Now, if we reduce the tax rate a little bit, individual 1 responds again with a marginal change in working hours. In contrast, individual 2 now reacts by making a discrete jump from not working at all to working nearly as many hours as individual 1. Thus, the incorporation of fixed costs of working in the budget constraint seems to provide a more realistic model of participation behavior.[1] Indeed, the studies by Cogan (1981) for the United States and by Blundell, Ham, and Meghir (1987) for the United Kingdom show that fixed costs are empirically important for the labor supply of married women.

The design of the tax transfer system may in itself be part of the explanation for the observed discontinuity in labor supply. For example, in some countries individuals face a U-shaped pattern of marginal tax rates. Effective marginal tax rates are high for low-income individuals due to the phasing out of welfare benefits, while marginal tax rates are high at the top of the earnings distribution where the progressivity of the tax system kicks in. Such a system gives rise to an S-shaped budget line as illustrated in figure 4.3. In the figure, the individual initially stays out of the labor market. But a small reduction in tax rates, creat-

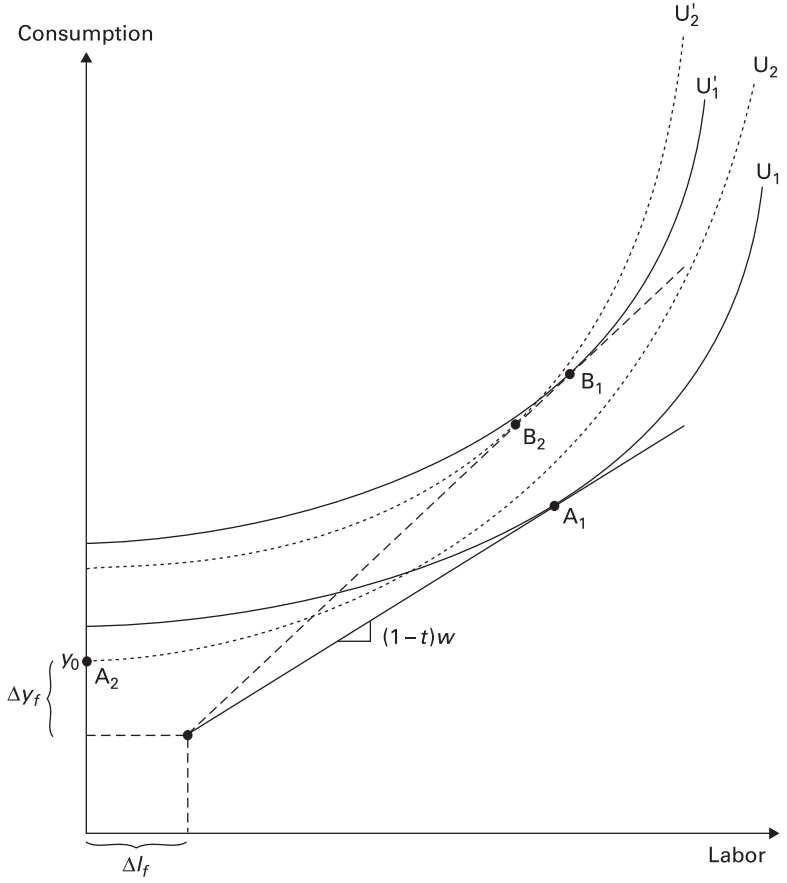

Figure 4.2
Intensive versus extensive responses with fixed costs of working

ing an upward shift in the budget line, induces the individual to jump in at l^* hours.

So far we have considered labor supply problems involving only one candidate for an interior solution, which is then compared to the corner solution with non-participation. In this situation, labor supply is discontinuous at the point of entry and exit, whereas hours of work for those who are working is a continuous choice variable. Such a model is fine as long as the non-convexity in the budget set is concentrated at the lowest earnings points as in the case of fixed work costs or welfare benefits that are (quickly) phased out upon entry. But if

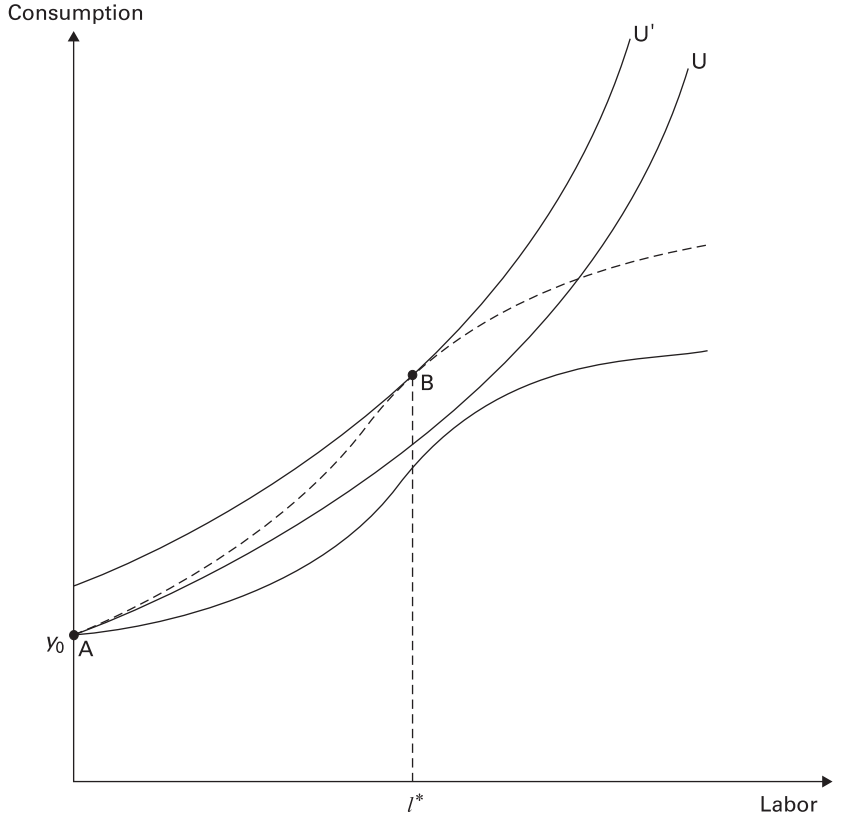

Figure 4.3
Discrete extensive response with a U-shaped pattern of marginal tax rates

there are discontinuities or other non-convexities further into the interior of the budget set, even the intensive labor supply response may become discrete. Such a situation is depicted in figure 4.4, where the budget line is discontinuous at l_1^* hours. The discontinuity could reflect that a benefit is lost completely once income exceeds a certain threshold or that some tax is imposed on the entire income above the threshold.[2] In the figure, it is initially optimal for the worker to locate at the notch l_1^* so as to avoid the discontinuous tax payment. But a tax reform creating a small upward shift in the budget line induces the worker to move discretely from l_1^* (part-time) to l_2^* (full-time).

To summarize, we note that tax reforms entail intensive as well as extensive labor supply responses, and to account for the observed

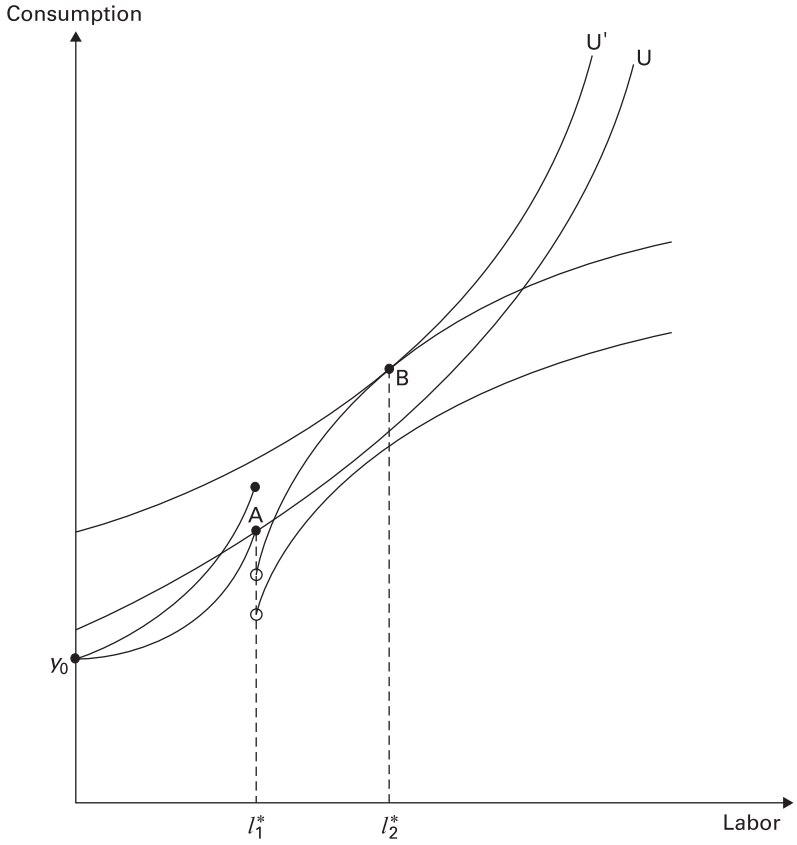

Figure 4.4
Discrete intensive response with a non-convex budget set

discreteness of responses, they have to be modeled by introducing non-convexities into the standard framework. The presence of non-convexities and discreteness is presumably more important for the extensive response, because of fixed work costs.

One may question the value of distinguishing explicitly between intensive and extensive labor supply responses for the evaluation of tax reform; that is, will we correctly estimate the welfare effects by combining the two margins into a single elasticity in the standard convex model? If so, Occam's Razor would suggest that we use this simpler method. We show in the following sections, however, that it is indeed important to distinguish between the two types of responses in welfare analysis, except under very special circumstances.

4.2.2 *The Welfare Effect from Intensive Responses*

In this section, we derive the welfare effect from intensive labor supply responses in a simple partial model. We consider a representative individual with earnings wl, where w is an exogenous wage rate and l is the hours of work. Without loss of generality, we disregard nonlabor income. The tax system is described by a function $T(wl, z)$, where z is an abstract parameter that we will use to derive the effects of policy reform. The tax function gives the net payment to the public sector, embodying taxes as well as transfers. The tax transfer schedule may involve nonlinearities and discontinuities, but we will restrict attention to the case of piecewise linearity where individuals face marginal tax rates that are locally constant.

The preferences of the individual are represented by a quasi-linear utility function

$$U = wl - T(wl, z) - v(l), \tag{1}$$

where disutility of work $v(l)$ is increasing and strictly convex ($v' > 0$, $v'' > 0$). The quasi-linear specification excludes income effects on labor supply that simplifies the algebra with no loss of generality for the welfare analysis. In fact, a general utility function would provide exactly the same results (cf. section 4.2.4).

Assuming that an interior solution exists, the optimal choice of working hours is characterized by the familiar condition that the marginal rate of substitution between consumption and leisure equals the marginal net-of-tax wage rate. That is,

$$(1 - m)w = v'(l), \tag{2}$$

where $m \equiv \partial T(wl, z)/\partial(wl)$ is the marginal tax rate.

The consequences of a small tax reform may be approximated by the effects of a marginal change in the parameter z. From (1) and (2), we see that a change in z affects utility directly through the tax payment and indirectly through the effect on hours of work. However, the indirect effect disappears by the envelope theorem, since working hours are initially at the optimum. Accordingly, the change in utility is given by

$$\frac{dU}{dz} = -\frac{\partial T}{\partial z}, \tag{3}$$

which is simply the mechanical change in the tax payment, namely, the tax change exclusive of feedback effects from behavioral responses to the reform.

The utility change does not in itself constitute the welfare effect of tax reform, which is instead related to the change in the deadweight loss (or excess burden) of taxation. With quasi-linear utility, the deadweight loss is given simply by $D = U_0 - U - T$, where U_0 is the (hypothetical) utility level with no taxation. The welfare loss of tax reform is then given by

$$\frac{dD}{dz} = -\frac{dU}{dz} - \frac{dT}{dz} = \frac{\partial T}{\partial z} - \frac{dT}{dz}, \tag{4}$$

where we have inserted the derivative (3). Thus, the welfare loss is equal to the difference between the mechanical change in tax revenue and the total change in revenue accounting for both mechanical effects and behavioral feedback effects of the reform. In other words, the welfare loss of tax reform is given simply by the revenue loss created by changed behavior. Notice that this insight does not rely on the preference specification adopted here; it follows from the definition of the deadweight loss and the application of the envelope theorem.

In this model, the behavioral feedback effect on revenue is given by the change in earnings multiplied by the marginal tax rate, implying that equation (4) may be written as

$$\frac{dD}{dz} = -mw\frac{dl}{dz}. \tag{5}$$

The working hours response dl/dz may be derived from (2) and substituted into (5) so as to get

$$\frac{dD/dz}{wl} = \frac{m}{1-m} \cdot \frac{\partial m}{\partial z} \cdot \varepsilon, \tag{6}$$

where $\varepsilon = v'(\cdot)/[v''(\cdot)l]$ denotes the (compensated) elasticity of hours of work with respect to the marginal net-of-tax wage. This expression is a classic Harberger-type formula for the marginal deadweight burden of taxation. It shows that the marginal deadweight loss depends on the initial level of the marginal tax rate, the increase in the marginal tax rate, and the hours-of-work elasticity.

The previous formula suggests that welfare effects may be substantial even for moderate labor supply elasticities provided that initial

marginal tax rates are high. Tax reforms should therefore aim at reducing marginal tax rates where these rates are relatively high. From this perspective, one would perhaps be skeptical about the recent moves toward "in-work" benefits for low-income families with children (mostly single mothers) such as the American EITC and the British WFTC (Working Family Tax Credit). These policies have raised marginal tax rates on average, since most eligible families are located in the phase-out region of the credit. Moreover, marginal tax rates within the eligible group were initially quite high, because many were receiving other benefits subject to phase-out (cash benefits, food stamps, etc.). This problem with the EITC is reinforced if the existing estimates of a highly elastic female labor supply (Killingsworth and Heckman 1986) are taken to indicate a high intensive elasticity ε. Indeed, this seems to be the underlying interpretation in the strong criticism of the EITC by Browning (1995). It is an incorrect interpretation, however, since high labor supply elasticities for single mothers reflect mostly extensive responses (Mroz 1987; Triest 1990).

4.2.3 The Welfare Effect from Extensive Responses in the Standard Model

A realistic model of extensive labor supply responses requires some type of heterogeneity. With no heterogeneity in preferences or in productivity, either everyone participates or nobody participates. The extensive labor supply response becomes discrete at the macrolevel in that the economy moves from no employment to full employment at some threshold. The standard method of smoothing the participation response is to introduce a continuum of productivities $w \in (0, \overline{w})$ in the convex model from section 4.2.2.

Individual labor supply is described by the first-order condition (2) and a non-negativity constraint on hours. From the first-order condition, non-negative hours correspond to a requirement that the marginal net-of-tax wage is greater than or equal to $v'(0)$. Thus, we may write individual labor supply at productivity w, denoted by l_w, as

$$(1 - m_w)w = v'(l_w) \quad \text{if} \quad (1 - m_w)w \geq v'(0); \qquad \text{otherwise } l_w = 0, \quad (7)$$

where m_w is the marginal tax rate faced by an individual with wage rate w (given the labor earnings chosen by this type). We may define a threshold wage level $\tilde{w}$ for the marginal entrant, that is, $\tilde{w} \equiv v'(0)/(1 - m_{\tilde{w}})$, such that all individuals with $w \geq \tilde{w}$ choose to enter the labor

market while all those with $w < \tilde{w}$ choose not to enter. If we let wages be distributed according to $F(w)$, the participation rate in the economy is given by $1 - F(\tilde{w}) = 1 - F(v'(0)/(1 - m_{\tilde{w}}))$.

It then follows that the tax system affects labor supply at the extensive margin through the marginal tax rate $m_{\tilde{w}}$ for the marginal entrant. Accordingly, a tax reform that reduces the marginal tax rate for people at or around the reservation wage level will induce additional entry. What is the welfare effect of these extensive responses? Notice first that unemployment in this model is voluntary such that the marginal entrant is indifferent between employment and non-employment. By implication, extensive labor supply responses do not create discrete utility changes. As in the previous model, the effect on excess burden is determined solely by the feedback effect on government revenue. But since individuals enter the labor market at infinitesimal working hours in this model, no additional tax revenue is created by more participation, at least not for small reforms. In other words, extensive labor supply responses are irrelevant from the point of view of economic welfare within the standard convex model. For extensive responses to matter for the evaluation of tax policy, the responses have to be discrete as in the nonconvex framework (or non-participation would have to be involuntary).

4.2.4 The Welfare Effect from Extensive Responses in a Nonconvex Model

In this section, we present a stripped-down model to capture participation responses in a realistic way. As suggested in the previous discussion, such a model should incorporate some type of non-convexity as well as heterogeneity across individuals. The non-convexity enables us to explain discrete entry-exit behavior at the individual level, while heterogeneity is needed to get a smooth participation response at the macro level.[3]

Perhaps the simplest way to introduce a non-convexity in the model is through fixed costs of working and, as mentioned in section 4.2.1, such work costs have been shown to be empirically significant. In what follows we lay out a framework where fixed work costs are captured in a single parameter q, which may reflect monetary and time costs or a "psychic" distaste for participation. A heterogeneity may then be introduced through the fixed costs by assuming that the economy consists of a continuum of individuals with $q \in (0, \bar{q})$. For

simplicity, let the productivity level be identical across individuals and, to focus on the extensive margin, let us fix hours of work for those who are working at 1 unit. Thus, individuals are simply making a choice between $l = 0$ and $l = 1$, involving utility levels

$$U = \begin{cases} w - T(w,z) - q & l = 1, \\ -T(0,z) & l = 0. \end{cases} \tag{8}$$

If the individual chooses to work, he obtains wage income net of taxes and the fixed work cost $w - T(w,z) - q$. This has to be compared to the public benefits $-T(0,z)$, which will be received in case of nonparticipation.

For the individual to enter the labor market, the utility from participation must be greater than or equal to the utility from nonparticipation, giving rise to the participation constraint

$$q \leq w - [T(w,z) - T(0,z)] \equiv \tilde{q}. \tag{9}$$

This expression defines an upper bound on the fixed cost, $\tilde{q}$, the size of which reflects the income gain from entry net of taxes and transfers. Individuals with a fixed cost below the threshold value $\tilde{q}$ decide to enter the labor market, while those with a fixed cost above the threshold value stay out. If we let q be distributed according to $P(q)$ with density $p(q)$, the labor force participation rate in the economy becomes equal to $P(\tilde{q})$. The sensitivity of the participation rate P with respect to changes in the net income gain from entry $\tilde{q}$ may be measured by the extensive labor supply elasticity

$$\eta \equiv \frac{\partial P}{\partial \tilde{q}} \frac{\tilde{q}}{P} = \frac{p(\tilde{q})\tilde{q}}{P}. \tag{10}$$

The relationship between taxation and labor force participation in this nonconvex model is different from the relationship in the convex model in section 4.2.3. In the convex model, the participation decision was related to the marginal tax rate at zero labor income whereas, in the model laid out here, the decision is related to the change in total tax transfer payments $T(w,z) - T(0,z)$ following entrance into the labor market. Obviously, this difference is related to the distinction between discrete and continuous labor supply behavior. In the standard model, individuals enter the labor market at infinitesimal hours, creating a change in the total tax liability equal to $T(y,z) - T(0,z) \approx m \cdot y$ since earnings y are small.

As in the previous model, non-employment is voluntary and the marginal entrant is indifferent between participating or not. Accordingly, the effect of extensive labor supply responses on aggregate utility envelopes out, and the marginal excess burden is therefore given by the revenue implications of changed participation. Since aggregate government revenue R is given by

$$R = T(w,z)P(\tilde{q}) + T(0,z)(1 - P(\tilde{q})), \tag{11}$$

the marginal excess burden may be written as

$$\frac{dD}{dz} = \frac{\partial R}{\partial z} - \frac{dR}{dz} = -[T(w,z) - T(0,z)]\frac{dP(\tilde{q})}{dz}. \tag{12}$$

The welfare effect of a tax reform is determined by its impact on the labor force participation rate multiplied by the increase in the net tax liability created by entry into the labor market.

By using equations (9) and (10), it is straightforward to derive the participation effect as a function of the extensive labor supply elasticity and parameters of the tax transfer system and the reform. By defining a tax rate on labor force participation, $a \equiv [T(w,z) - T(0,z)]/w$, we may write the marginal deadweight burden in proportion to aggregate income in the following way:

$$\frac{dD/dz}{wP} = \frac{a}{1-a} \cdot \frac{\partial a}{\partial z} \cdot \eta. \tag{13}$$

This deadweight formula reflects the same basic form as the traditional one in equation (6), but it is related to different policy parameters and a different elasticity. In particular, the welfare cost is no longer related to the *marginal* tax rate. It is instead related to the tax rate applying at the extensive margin. This tax rate is an effective *average* tax rate, which includes the loss in benefits from labor market entry. Finally, the welfare effect depends on the sensitivity of entry-exit behavior as measured by the elasticity of labor force participation with respect to the net-of-tax income gain from entry, η.

The policy conclusions that follow from the preceding formula are quite different from those implied by the standard analysis. Again, the EITC is a case in point. As mentioned in section 2.2, the standard analysis leads to some skepticism of the EITC, because of the implied increase in effective marginal tax rates in the phase-out region of the tax credit. The standard analysis, however, misses an essential problem with the welfare system, which is that it keeps some individuals

completely out of the labor force. A model with discrete participation, on the other hand, allows the introduction of tax credits to low-wage earners to improve welfare, since lower average tax rates encourage labor force participation. This conclusion is reinforced by empirical evidence indicating high participation elasticities among those targeted by in-work benefits in both the United States and the United Kingdom (Eissa and Liebman 1996; Blundell et al. 2000).

4.2.5 *A General Formula for the Evaluation of Tax Reform*

Tax reforms give rise to both intensive responses and (discrete) extensive responses. The effects on welfare from these two types of responses have been illustrated in sections 2.2 and 2.4. In order to focus on the central mechanisms at work, we have kept the theory as simple as possible. To quantify the welfare effects of a reform in the tax benefit system, we would need to combine intensive and extensive responses in a more realistic theoretical setting. The theory should build on a more general specification of preferences, so as to allow for income effects in labor supply, and should also account for individual heterogeneity in productivities and preferences. In Eissa, Kleven, and Kreiner 2004, we show that the more general model gives rise to the following formula for the marginal excess burden[4]

$$\frac{dD/dz}{\sum_{i=1}^{N} w_i h_i P_i} = \sum_{i=1}^{N} \left[\frac{m_i}{1-m_i} \frac{\partial m_i}{\partial z} \cdot \varepsilon_i + \frac{a_i}{1-a_i} \frac{\partial a_i}{\partial z} \cdot \eta_i \right] s_i, \tag{14}$$

where $i = 1, \ldots, N$ is the index parameter for individuals and $s_i \equiv w_i l_i P_i / (\sum_{i=1}^{N} w_i l_i P_i)$ is the wage share of individual i. In this analysis, each individual works with a given probability P_i. Accordingly, the participation elasticity, η_i, measures how much the probability of working increases when the income gain of entry rises.

Formula (14) shows how to calculate the impact on aggregate welfare of a tax reform. The first term in the bracket measures the welfare effect of intensive margin responses, while the second term captures the welfare effect of extensive margin responses. Notice the resemblance to the welfare effects obtained along the two margins in the simple models, namely, equations (6) and (13). The total welfare effect is obtained by adding intensive and extensive margin effects for each individual, then calculating a weighted sum of the individual welfare effects with wage shares used as weights.

When using the previous formula to evaluate actual tax reforms, it is important to distinguish between compensated and uncompensated labor supply elasticities. The simple models are silent about this because of the absence of income effects. The general theory, however, reveals that we should apply compensated elasticities. For the participation elasticity, however, it may be shown that compensated and uncompensated responses to tax reform are identical (Eissa, Kleven, and Kreiner 2004).

A priori one might have wondered whether the standard convex framework could be saved by a reinterpretation of the labor supply elasticity. Following this interpretation, one would introduce extensive responses into the framework simply by using estimates of the *total* labor supply elasticity including both margins of response.[5] The previous formula demonstrates that, in general, this approach is not correct. This is because the extensive welfare effect is related to the effective average tax wedge (including average benefit reduction), which is different from the effective marginal tax wedge (including marginal benefit reduction) associated with the intensive effect. For tax reform in practice, the two rates generally differ in terms of both pre-reform levels and the reform-induced changes. In the case of the EITC, for example, the changes in marginal and average tax rates have opposite signs for most people eligible for the credit.

There is one special case, however, where the conventional model is valid. This is the case of a *linear* Negative Income Tax, which grants a lump-sum transfer B to all individuals in the economy (participants and non-participants) and then imposes a constant marginal tax rate on labor income, $m_i = m\ \forall i$. In this case, the tax burden on labor market entry for individual i becomes $T(w_i l_i, z) - T(0, z) = (mw_i l_i - B) - (-B)$, which implies a participation tax rate $a_i = m\ \forall i$. Moreover, if the tax reform is simply a change of tax transfer parameters within the framework of the NIT, we would have $\partial a_i/\partial z = \partial m_i/\partial z\ \forall i$. Inserting these relationships in equation (14), we get a standard Harberger formula with the intensive and extensive elasticities lumped together:

$$\frac{dD/dz}{\sum_{i=1}^{N} w_i h_i P_i} = \frac{m}{1-m}\,\frac{\partial m}{\partial z}\sum_{i=1}^{N}(\varepsilon_i + \eta_i)s_i.$$

The practical applicability of this special case is clearly limited since it requires that the entire welfare system be a linear NIT. Public benefits tend to be non-universal, targeted to low- and middle-income

classes through earnings or work tests (creating discontinuities) or through gradual phase-outs. Consider as examples low-income support, in-work benefits, housing and education subsidies, child benefits, medical aid, food stamps, and public pensions. While some benefits may be universal in some countries, they are never collectively so. A description of the actual welfare systems prevailing in a number of European countries is provided by Immervoll et al. (2004), while the U.S. tax transfer system will be discussed in section 4.5.1.

4.3 Tax Reforms in the United States and the EITC

Over the past two decades, a series of tax acts—passed in 1981, 1986, 1990, 1993, 2001, and 2003—have substantially changed the federal income tax structure and federal income tax liabilities in the United States. To outline the major features of the tax changes, table 4.1 presents federal income tax parameters from 1984 to 2002. The table highlights the dramatic changes to the federal income tax schedule over the period. In 1984, the federal (non-EITC) tax schedule consisted of fifteen brackets, with marginal rates ranging from 0 to 50 percent. It now stands at five rates, ranging from 15 percent to nearly 40 percent. The table also highlights the central role of the Earned Income Tax Credit in altering the shape of the tax schedule.

Since the EITC is important in the empirical application that we discuss later on, it is useful to start by describing the EITC program. It began in 1975 as a modest program aimed at offsetting the social security payroll tax for low-income families with children. Since its introduction the credit was little changed, increasing from $400 to $500 (nominal dollars) by 1985 at its maximum level. Following the 1986 and subsequently the 1990 and 1993 expansions, the EITC has become the largest cash transfer program for low-income families with children at the federal level. By 2000, total EITC expenditures (tax expenditures and direct outlays) amounted to about $30 billion.

An important feature of the EITC is that eligibility is conditional on the taxpayer having positive earned income. Moreover, the size of the credit to which the taxpayer is entitled depends on the amount of earned income (adjusted gross income, or AGI) and, since 1991, the number of qualifying children who meet certain age, relationship, and residency tests. In tax year 2002, taxpayers with two or more children must have earnings no higher than $33,150 to qualify for a credit. Children must be under age 19 (or 24 if a full-time student) or permanently

disabled, and must reside with the taxpayer for more than half the year.

There are three regions in the credit schedule. The initial phase-in region transfers an amount equal to the subsidy rate (currently 36 percent for larger families) times earnings. In the flat region, the family receives the maximum credit ($4,204 in tax year 2002), while in the phase-out region, the credit is phased out at a fixed phase-out rate (21 percent in 2002). The credit is refundable so that a taxpayer with no federal tax liability, for example, would receive a tax refund from the government for the full amount of the credit.

The 1986 expansion of the EITC, passed as part of the Tax Reform Act of 1986 (TRA86) increased the subsidy rate for the phase-in of the credit from 11 percent to 14 percent and increased the maximum income to which the subsidy rate was applied from $5,000 to $6,080. This resulted in an increase in the maximum credit from $550 to $851 ($788 in 1986 dollars). The phase-out rate was reduced from 12.22 percent to 10 percent. The higher maximum credit and the lower phase-out rate combined to expand the phase-out region from $11,000 in 1986 to $18,576 by 1988.

The impact of the EITC expansion on the tax liability of eligible taxpayers was reinforced by other elements of TRA86. TRA86 increased the standard deduction for a taxpayer filing as head of household from $2,480 in 1986 (included in the zero bracket) to $4,400 in 1988. TRA86 further reduced the tax liability of taxpayers with children by increasing the deduction per dependent exemption from $1080 in 1986 to $1950 in 1988. Finally, the tax schedules were changed. The tax schedule changes were particularly beneficial to head-of-household filers because the increased standard deduction and exemption amounts meant that in 1988 the typical head-of-household filer did not jump from the 15 percent tax bracket to the 28 percent tax bracket until her adjusted gross income (AGI) exceeded $33,565. Together, all these elements of the 1986 reform reduced substantially the tax liabilities for single women with children.

The Omnibus Budget Reconciliation Act of 1990 (OBRA90) further expanded the EITC for all eligible families and introduced a different EITC schedule for families with two or more children. The phase-in rate of the EITC was increased from 14 percent to 18.5 for taxpayers with one child and 19.5 percent for taxpayers with more children. OBRA90 also led to a larger (nominal and real) increase in the maximum benefit, phased in over three years.

Table 4.1
Federal income tax and EITC parameters, 1984–2002

	Federal income tax parameters		EITC parameters (family with one child; family with two or more children)			
Year	[lowest, highest marginal tax rate] (number of brackets)	Personal exemption, standard deduction[a,b]	Phase-in rate	Maximum credit	Phase-out rate	Maximum earnings
1984	[0.000; 0.500] (15)	$1,000; $0	0.100	$500	0.125	$10,000
1985	[0.000; 0.500] (15)	$1,040; $0	0.110	$550	0.122	$11,000
1986	[0.000; 0.500] (15)	$1,080; $0	0.110	$550	0.122	$11,000
TRA86						
1987	[0.110; 0.390] (5)	$1,900; $2,540	0.140	$851	0.100	$15,432
1988	[0.150; 0.330] (2)	$1,950; $4,400	0.140	$874	0.100	$18,576
1989	[0.150; 0.330] (2)	$2,000; $4,550	0.140	$910	0.100	$19,340
1990	[0.150; 0.330] (2)	$2,050; $4,750	0.140	$953	0.100	$20,264
OBRA90[c]						
1991	[0.150; 0.310] (3)	$2,150; $5,000	0.167	$1,192; $1,235	0.119; 0.124	$21,250
1992	[0.150; 0.310] (3)	$2,300; $5,250	0.176	$1,324; $1,384	0.126; 0.130	$22,370
1993	[0.150; 0.396] (5)	$2,350; $5,450	0.185	$1,434; $1,511	0.132; 0.139	$23,050
OBRA93						
1994	[0.150; 0.396] (5)	$2,450; $5,600	0.263	$2,038; $2,526	0.160; 0.177	$23,755; $25,296
1995	[0.150; 0.396] (5)	$2,500; $5,750	0.340	$2,094; $3,110	0.160; 0.202	$24,396; $26,673
1996	[0.150; 0.396] (5)	$2,550; $5,900	0.340	$2,152; $3,556	0.160; 0.211	$25,078; $28,495
1997	[0.150; 0.396] (5)	$2,650; $6,050	0.340	$2,210; $3,656	0.160; 0.211	$25,750; $29,290
2000	[0.150; 0.391] (5)	$2,900; $6,650	0.340	$2,353; $3,888	0.160; 0.211	$27,450; $31,152
EGTRRA						
2001	[0.100; 0.386] (5)	$3,000; $6,900	0.263	$2,428; $4,008	0.160; 0.211	$28,250; $32,100
2002	[0.100; 0.386] (6)	$3,050; $7,000	0.340	$2,547; $4,204	0.160; 0.211	$30,200; $33,150

Source: The Green Book and authors' calculations from Internal Revenue Service (IRS) forms and publications.
[a] The standard deductions are given for head of household tax return.
[b] In 1984–1986, there were no standard deductions because of the zero bracket. The fifteen brackets include the zero bracket.
[c] Basic EITC only. Does not include supplemental young child credit or health insurance credit.
[d] Introduced a small benefit for taxpayers with no qualifying children, phased in at 0.0765 up to a maximum credit of $306.

The budget reconciliation act of 1993 (OBRA93) further increased the *additional* maximum benefit for taxpayers with two or more children to $1,400 by 1996 ($3,556 vs. $2,152 in 1996), and doubled the subsidy rate for the lowest-income recipients from 19.5 to 40 percent for larger families (18.5 to 34 percent for families with one child). These changes combined to dramatically expand eligibility for the EITC, such that by 1996 a couple with two children would still be eligible at incomes of almost $30,000. Other than the expansions to the EITC there was little in the way of changes to the federal income tax for lower-income individuals in OBRA93.

Finally, the 2001 Economic Growth and Tax Relief Reconciliation Act (EGTRRA) further reduced taxes at the lower end of the distribution, although to a much lesser extent than the previous reforms. This time the main change in tax liabilities did not come through the EITC but through a reduction in the lowest income tax bracket from 15 to 10 percent.

4.4 Intensive and Extensive Labor Supply Responses: A Review of the Empirical Literature

This chapter argues that recent empirical evidence on labor supply behavior has important implications for the evaluation of tax reforms. In particular, we are interested in the emerging consensus that labor supply responses tend to be concentrated more along the extensive margin (labor force participation) than along the intensive margin (hours of work). This consensus emerged over an extended period of time from three different sources of evidence: direct evidence from recent tax reforms and expansions to tax-based transfers in the United States and the United Kingdom, evidence from the Negative Income Tax Experiments, and indirect evidence from structural models of labor supply. We discuss the evidence in that order.

Perhaps the most compelling evidence on the responsiveness of the participation margin arises from the evaluation of the labor supply effects of recent reforms to the tax and to the transfer system in the United States. We focus on that evidence because it is most relevant for the empirical exercise in the chapter. The literature evaluating expansions to the Earned Income Tax Credit has generally employed both quasi-experimental (difference-in-differences) methods (Eissa and Liebman 1996, hereafter EL; Eissa and Hoynes 2004, hereafter EH; Hotz, Mullin, and Scholz 2002, hereafter HMS) and structural methods

(Dickert, Houser, and Scholz 1995, hereafter DHS; and Meyer and Rosenbaum 2001, hereafter MR).

The first set of results come from studies that use quasi-experimental methods to examine the labor supply effects of the Tax Reform Act of 1986 on female heads (EL), the 1993 EITC expansion on married women (EH), and the EITC expansions in the 1990s on welfare recipients in California (HMS).[6] EL compare the change in labor force participation and hours worked by single mothers to that of single women without children, and find a sizeable labor force participation response of 2.8 percentage points (out of a base of 74.2). Their data (the Current Population Survey, or CPS) also show no evidence of an hours-of-work response.

HMS extend this work by accounting for labor supply incentives due to state-level welfare to work programs. Using administrative data on welfare recipients in four California counties, HMS evaluate the labor supply response of larger welfare families to the marginal second child credit. Their findings are dramatic and show an increase in the employment rate of larger families between 6 and 8 percentage points relative to families with one child, implying an elasticity with respect to net income of up to 1.7.

Overall the evidence based on the difference-in-differences model is consistent and suggests fairly strong participation effects, especially for female household heads. One limitation of the reduced-form labor force participation methods (as applied in these papers) is the use of group-level variation in taxes and transfers. This approach assumes that all relevant wage and income changes are captured by group-level variation in family type and size (presence and number of children) and time.

DHS and MR exploit the fact that tax reforms typically have heterogenous effects within group to estimate the effect of EITC expansions on labor force participation. DHS use cross-sectional data from the 1990 Survey of Income and Program Participation (SIPP) and estimate a joint program and labor force participation model, identified by variations in the returns to part-time (or full-time) employment in different states. They estimate a labor force participation elasticity of 0.35. MR extend this work by using time variation in tax and welfare policies. The use of time variation allows MR to eliminate any bias that results from different state characteristics that are unobserved and correlated with labor supply incentives. MR use a discrete participation model based on comparisons of utility in and out of the labor force, and CPS

data from 1985 to 1997. Their finding that the EITC accounts for about 60 percent of the increase in the employment of single mothers over the period implies a labor force participation elasticity of about 0.7.

The second major source of direct evidence comes from the Negative Income Tax experiments. The NIT experiments were randomized experiments conducted in various sites in the United States from 1968 to 1982 to study the responsiveness of labor supply to taxes (see review by Moffitt and Kehrer 1981). Program participants received over $60 million in direct benefits over the four sites (New Jersey; Gary, Indiana; rural Iowa and North Carolina; and Seattle/Denver). The experiments typically lasted three years, but a small proportion of the Seattle/Denver experiment lasted for five to twenty years. The NIT experiments varied across sites in several features, including sample size (809 to 4,800), income eligibility cutoff (1.5 to 3.25 times poverty), and in the tax benefit parameters. Robins (1985) presents a useful comparison of the labor supply findings in the NIT experiments. His summary concludes that participation responses are small for men but sizeable for females (single and married) as well as young individuals. In fact, the average employment effects across the four experiments are stronger than the annual hours worked response for both single and married females. One may argue that these findings are not relevant because of the dramatic changes to the labor market and the labor market participation of women over the past three decades. We are sympathetic to this view but present these results to show that recent findings regarding the participation responses to tax reforms should not be surprising.

Finally, indirect evidence comes from two studies that are largely focused on the statistical properties of labor supply models (Triest 1990; Mroz 1987). Triest examines labor supply behavior using nonlinear budget set models. These models allow for jumps as well stickiness in labor supply with nonproportional tax schedules and therefore are particularly useful for evaluating the effects of taxes. Using Panel Study of Income Dynamics (PSID) data, Triest estimates censored and truncated hours of work models and finds larger elasticities for married women when using the censored data. He concludes that the difference is likely due to stronger labor force participation than hours worked responsiveness to taxes. In his sensitivity analysis of married women's hours of work, Mroz (1987) finds similar results for married women using a standard model of labor supply that does not explicitly

account for budget set nonlinearities. Both Triest and Mroz finds small hours worked responses by working women.

Overall, the evidence is convincing that the labor force participation margin is sensitive to taxes, and in fact more sensitive than are hours worked.

4.5 Welfare Effects from Tax Reform: The Tax Reform Act of 1986 as a Case Study

4.5.1 Methodology and Data

As described earlier, a succession of tax acts have changed work incentives in the United States quite substantially. In particular, the EITC expansions strongly improved returns to entry in the labor market, and they involved large changes on the intensive margin too because of the phasing in and out of the credit. Based on the empirical findings of large participation responses, we have argued that it is crucial to incorporate this margin of response in the welfare analysis. More specifically, we showed theoretically that it is necessary to make an explicit distinction between the extensive and intensive margins whenever nonlinearities are significant, either in the pre-reform taxes and transfers or in the reform itself. Since the EITC is a highly nonlinear transfer, our point seems to be very relevant for the recent tax reforms.

To understand the quantitative importance of these insights, we apply the theory outlined in section 4.2 to study the Tax Reform Act of 1986 (TRA86). We focus on the welfare effects of the reform on single women with dependent children. This population is particularly interesting for two reasons. First, it were affected more by the EITC than any other group in the labor market. After the full phase-in of the 1986 reform, 95 percent of all single mothers had income below the EITC maximum allowable income. Besides the effect of the EITC, tax liabilities and average tax rates of single mothers were reduced further due to increases in the standard deduction, personal exemptions, and a more favorable tax rate schedule. All these elements led to big improvements in the labor supply incentives on the extensive margin. Second, single mothers are interesting in our context because their labor supply responses to tax reform have been so well documented. In particular, we know that participation elasticities are large for this group.

We use a microsimulation approach to study the welfare effects of the reform. The simulations are based on the formula in (14). Notice that this approach is exact only for small policy reforms. Since the tax changes that were introduced by the 1986 tax act were quite substantial, our calculations should be seen as first-order approximations to the true welfare effects. The potential errors of the approximation will be discussed later.

The implementation of our deadweight formula requires information about labor supply elasticities as well as various tax transfer parameters and wage income shares. The labor supply elasticities will be taken from the empirical literature reviewed in section 4.3. To generate the tax transfer parameters, we first estimate effective marginal and average tax rates (m_i and a_i) for each individual. Calculations of marginal and average tax rates should account for all taxes on labor income (federal taxes, state taxes, and payroll taxes), and they should account for the phase-out of welfare benefits. Moreover, the tax rates should reflect values pertaining to the time immediately before the reform (we use 1985 as the pre-reform year).

The changes in tax rates introduced by the reform ($\partial m_i/\partial z_i$ and $\partial a_i/\partial z_i$) are computed as the difference between the pre-reform rates and imputed post-reform rates, reflecting the changes in the tax code after the reform is fully phased in (but not any behavioral responses). To isolate the impact of TRA86 (which was a federal tax reform), we include only the difference between pre-reform and post-reform federal tax rates. The post-reform federal tax rate is imputed by applying the post-reform federal tax code to pre-reform individual incomes (adjusted for inflation). To account for the phasing-in of the reform, we use 1988 as the post-reform year.

The data for the simulations come from the March Current Population Survey. The March CPS is an annual demographic file of approximately 60,000 households, with information on labor market and income outcomes. From CPS we extract information on unmarried females (widowed, divorced, and never married) who are between 16 and 55 years old and who have dependent children (children under the age of 19 or under 24 if a full-time student). We exclude older women to avoid complications related to modeling retirement decisions. We also exclude any female who was ill or disabled, was in the military, or reported herself retired during the previous year. Finally, we exclude any woman with negative earned income (due to negative

Table 4.2
Summary statistics: Single mothers

Age	32.60 (8.64)
Years of education	12.82 (2.54)
Number of children	1.92 (1.11)
Non-white	0.306 (0.461)
Labor force participation	0.702 (0.457)
Earned income	$7,998 (9.366)
Non-labor income	$2,837 (4,963)
Gross hourly wage	$6.49 (4.49)
Observations	4,732

Note: Authors' tabulations of the 1985 March Current Population Survey. Sample includes unmarried mothers ages 16–55. See text for further sample selection. Earned income includes wage and salary and self-employment income. Non-labor income is calculated as the difference between total income and earnings, and therefore includes income for various sources such as welfare assistance, capital income, social security income, and workers' compensation. The wage is defined for workers only. Standard errors are in parenthesis. All monetary amounts are in nominal dollars.

self-employment income), negative unearned income, or with positive earned income but zero hours of work. The resulting sample size is 4,730 observations. Table 4.2 presents summary statistics of the characteristics of all unmarried females with children used for the analysis. The typical unmarried mother is 33 years old with two children. She has a high school diploma and earns an hourly wage of $6.49 in 1985.

The tax parameters for the sample of female heads are calculated using the Tax Simulation Model (TAXSIM) of the National Bureau of Economic Research (NBER). Since we cannot observe the earnings of nonworkers if employed, we impute earnings for all individuals in the sample using a simple earnings regression.[7] The predicted earnings and CPS data on individual nonlabor income are used as input in TAXSIM to calculate the marginal tax rate and the tax liability of each individual in the sample. Because the CPS does not ask about itemized deductions, we assume everyone takes the standard deduction. The TAXSIM-generated tax rates include the federal, state and payroll tax components but do not include the transfer component.

In the United States, lower-income families are eligible to receive cash assistance from the Temporary Assistance to Needy Families, previously the Aid to Families with Dependent Children (AFDC) program. In addition, eligible families may receive in kind benefits in the form of food vouchers (food stamps) and health insurance (Medicaid).

To incorporate benefits, we augment tax data from TAXSIM with information on AFDC, food stamps, and Medicaid. Because transfer programs have differing eligibility and benefit structures at the federal and state levels, we treat each program separately. We assign unemployed parents the maximum AFDC benefit in their state of residence and apply a tax rate of 25 percent on earned income (Triest 1994). To assign Medicaid benefits, we used the average annual health expenditures on AFDC recipients by state in 1985. Because Medicaid benefits were not reduced on the margin with earned income, Medicaid eligibility and expenditures affect the total tax liability (and average tax rate) but not the marginal tax rate.

To get an impression of the shapes of the individual budget constraints, we display in figure 4.5 the budget set of a woman who is living in Pennsylvania and has two children. The graph shows that effective taxes can be very high at low incomes, especially before the 1986 tax reform. Prior to the reform, if the mother chooses to enter the labor market at $8000, she will lose $6,000 in net tax payments and withdrawal of AFDC and food stamps. In other words, the effective average tax rate relevant for the entry decision is 75 percent for this in-

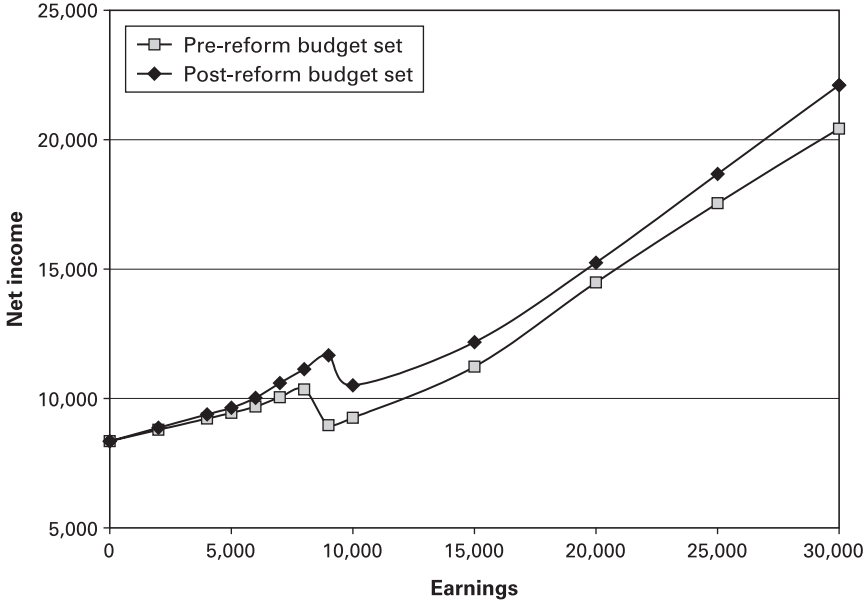

Figure 4.5
Budget set for single mother with two children (Pennsylvania)

dividual. The downward jump in net income as earnings pass $8,000 occurs because Medicaid is lost entirely at this earnings level. Immediately beyond the Medicaid threshold, there is essentially no net income gain from entry; that is, the participation tax rate is almost 100 percent. It will never be optimal for the individual to enter at this earnings level. Finally, notice that the 1986 reform provided a large increase in the incentive to work. At an earnings level of $8,000, the net income gain of working was increased from $2,000 to $3,000.

Figure 4.5 illustrates only an example of a budget set. In the sample, there is substantial variation in the budget sets across the individuals depending on state of residence, number of children, and so forth. In addition, our procedure overestimates on average the true effective tax rates because the take-up rate of benefits has been shown to be much less than 1 (see the review by Moffitt 1992). To correct this bias, we apply an empirical take-up rate, calculated from our CPS data as the share of the eligible population that reports positive benefits during the year. In the 1985 CPS, the empirical take-up rate was 56 percent. We apply this rate to the statutory benefit amount and the statutory benefit reduction rate. Therefore, both the marginal and the average tax rates are adjusted in this procedure.

Table 4.3 presents sample means and standard deviations for the tax parameters of female household heads. What the data shows is that the typical female household head loses 25.4 cents on the next dollar of earnings to federal and state income taxes, and an additional 17.7 cents to the welfare authorities through reduced cash and food benefits. In total, she pays 43 percent on the marginal dollar earned. On average, however, she loses about 52 percent of her total earnings to the tax and welfare authorities with entry into the labor market. As a consequence, the tax rate on the extensive margin is higher than the tax rate on the intensive margin. The table also reports "tax ratios" to be used in the welfare calculations, namely, $t/(1-t)$, where t denotes the marginal or average tax rate. The average tax ratio associated with the extensive welfare effect is higher than the intensive tax ratio important for the intensive welfare effect. Finally, as shown at the bottom of the table, TRA86 had a large impact on the tax treatment of single mothers. On average, their effective marginal and average tax rates were reduced by 6 and 8 percentage points, respectively. The standard deviations indicate substantial heterogeneity, however, especially for the changes in the marginal tax rate due to the phasing in and out of the EITC.

Table 4.3
Tax-benefit parameters for single mothers for TRA86

Tax rate on the intensive margin (pre-reform)	
Marginal tax rate without benefits	0.254 (0.164)
Marginal tax rate with benefits	0.431 (0.156)
Marginal tax ratio	1.129 (3.137)
Tax rate on the extensive margin (pre-reform)	
Average tax rate without benefits	0.144 (0.106)
Average tax rate with benefits	0.518 (0.109)
Average tax ratio	1.359 (3.742)
Effect of the 1986 tax reform act	
Change in the effective marginal tax rate	−0.060 (0.095)
Change in the effective average tax rate	−0.075 (0.029)

Note: The reported parameter values are sample means while numbers in parenthesis are standard deviations. The tax parameters include federal, state, and social security payroll rates. The effective tax rates account for the withdrawal of benefits and include cash assistance (AFDC), food stamps, and Medicaid adjusted for an empirical 54 percent take-up rate. The tax ratios are derived by calculating the average over the individuals of $t/(1-t)$, where t is the relevant tax rate of the individual. The tax ratios include benefits. The pre-reform tax benefit rates are from 1985. The changes in the marginal and average tax rates reflect only changes at the federal level. The data come from the 1985 March Current Population Survey.

4.5.2 Welfare Effects

In this section, we use the tax and benefit calculations described earlier to simulate the welfare effects arising from the behavioral responses of single mothers to the 1986 tax act. Before turning to the simulation results, we attempt to gauge the effects of the tax reform simply by looking at the sample means presented in table 4.3, using a representative agent approach. In this approach, the welfare effect from the intensive margin is calculated by multiplying the means of the marginal tax ratio and the marginal tax rate change, then applying an (compensated) hours-of-work elasticity. To calculate the welfare effect from the extensive margin of response, we repeat the procedure using the average tax ratio, the average tax rate change, and the elasticity of labor force participation. Based on the tax and benefit parameters presented in table 4.3, this approach yields substantial positive welfare effects from both reforms and on both margins even for very small elasticities.

This approach is unlikely to yield precise results, however, because single mothers are quite hetergeneous as a group. More precisely, the shortcoming of the representative agent approach in this evaluation arises from the fact that tax rates, tax rate changes, and wage income

Table 4.4
Welfare effects from TRA86 on single mothers for different elasticity scenarios

Elasticity scenarios	Elasticities		The welfare gain from tax reform			Welfare gain per dollar spent
	Extensive	Intensive	Extensive	Intensive	Total	
Low	0.3	0.1	3.27	0.81	4.08	2.07
		0.2	3.27	1.62	4.89	2.62
		0.3	3.27	2.43	5.70	3.59
Middle	0.5	0.1	5.46	0.81	6.27	4.81
		0.2	5.46	1.62	7.08	9.50
		0.3	5.46	2.43	7.89	352.22
High	0.7	0.1	7.64	0.81	8.45	Laffer
		0.2	7.64	1.62	9.26	Laffer
		0.3	7.64	2.43	10.07	Laffer

Note: The welfare gain is measured in percentage of wage income and is calculated using equation (14) in the text. The total welfare gain is calculated as the sum of the intensive and extensive gains. The welfare gain per dollar spent measures the total welfare gain in proportion to the aggregate reduction in tax burden. A Laffer curve effect arises if the welfare gain is larger than the reduction in tax burden, in which case the reform creates a net tax revenue. The data come from the 1985 March Current Population Survey.

shares are highly correlated. For example, in our sample those initially in the phase-out range of the EITC experience a reduction in the effective average tax rate of 10 percentage points while the effective average tax rate of those in the phase-in interval decreases by only 6 percentage points. Since those in the phase-out range have higher productivities, their behavioral responses have more weight in the welfare calculation. Such heterogeneity is critically important for correctly simulating the welfare effects of tax reforms. As such, one should interpret with caution welfare calculations that do not account fully for the type of heterogeneity just described.

The simulations that we present here exploit all the individual heterogeneity in the sample. The results are shown in table 4.4, where we consider different elasticity scenarios. We examine three different participation elasticities (0.3, 0.5, and 0.7) that are each combined with three different values of the (compensated) hours-of-work elasticity (0.1, 0.2, and 0.3). All these scenarios seem to fall within the range of elasticities estimated in the literature on female labor supply. In particular, the scenarios reflect that the weight of empirical evidence indicates that the extensive elasticity is significantly larger than the intensive elasticity, especially for single mothers.

Columns 4–6 present the estimated welfare gains from the reform in proportion to the aggregate wage income of single mothers, decomposed into effects arising from labor supply responses along the two different margins. The last column shows the welfare gain to single mothers per dollar of revenue spent on their tax relief. This last number corresponds also to the so-called marginal cost of public funds, which is the estimate one would get when thinking about raising additional government revenue by reversing the reform.

Turning to the results in table 4.4, we find that TRA86 generated substantial welfare gains. In the middle scenario, the total gain is about 7.1 percent of wage income, and the welfare effects spanned by the different elasticity scenarios are from 4.1 to 10.1 percent. According to these estimates, the welfare gain per dollar spent is at least $2.00. In fact, for all the high elasticity scenarios—not out of the bounds of empirical estimates—we obtain Laffer curve effects. In these cases, the labor supply responses to the reform create an increase in government revenue that is sufficiently large to finance the initial reductions in the tax liabilities of single mothers. With Laffer effects, the gain per dollar spent is negative and hence difficult to interpret, which is why we do not show it. Notice also that, as we approach the maximum of the Laffer curve, the gain per dollar spent goes to infinity. This explains the very large number in one of the scenarios.

The table shows that the aggregate welfare gain is the result of positive effects along each of the two margins of labor supply. The 1986 tax reform act reduced average tax rates for almost all single mothers (99.7 percent of our sample), thereby increasing labor force participation and creating positive welfare effects along this margin. The effect on marginal tax rates, on the other hand, was not unambiguous due to the phasing in and out of the EITC. In particular, some people experienced higher marginal tax rates due to an expanded phase-out region for the EITC. Yet our tax simulations show that the marginal tax rate was reduced for 81.8 percent of the individuals in our sample, which explains the positive welfare effects created on the intensive margin. Taking a closer look at the size of the welfare effects at the two margins, we observe that for all elasticity combinations, the welfare effect along the extensive margin is greater than that along the intensive margin. In the middle elasticity scenario, almost 80 percent of the overall welfare gain is generated by movements into the labor market.

Of course, one reason for the larger welfare effects along the extensive margin is simply that participation elasticities are higher than

hours-of-work elasticities. However, this difference in labor supply responsiveness cannot account entirely for the difference in welfare effects. As an example, consider the elasticity scenario where both elasticities are equal to 0.3. Even in this scenario, the adjustment in labor supply along the extensive margin creates a larger welfare effect than that occurring on the intensive margin. This result illustrates why it is important to distinguish explicitly between the two margins of labor supply response. The distinction is important due to the difference between marginal and average tax rates created by nonlinearities (and discontinuities). The effect of nonlinearity is present both in the pre-reform tax system and in the reform.

As shown in table 4.3, the 1986 reform reduced average tax rates by more than it reduced marginal rates. In isolation, this generates a larger participation response than hours-of-work response and, by implication, a larger welfare gain is generated along the extensive margin. This effect is reinforced by the fact that effective average tax rates were about 9 percentage points higher than marginal tax rates prior to the reform. This implies that tax distortions were initially higher on the participation margin than on the hours-of-work margin. Therefore, a given increase in labor supply is more beneficial for economic efficiency if it is occurring along the extensive margin.

An important feature of the EITC is its heterogeneous effects on taxpayers at different points in the income distribution. To explore the role of heterogeneity, table 4.5 shows tax-benefit parameters and welfare calculations by income groups for the middle elasticity scenario. We divide the population according to the threshold levels of the EITC prior to the implementation of the 1986 reform. The overall decline in marginal tax rates reflects substantial differences across individuals. The largest reductions were concentrated among taxpayers with incomes in the flat and phase-out ranges before the reform ($5,000–$11,000), representing 49 percent of the total population and 54 percent of wage income. The lowest income taxpayers, representing over one-third of the population of female heads and 10 percent of wage income, had their marginal rates reduced by less than the average individual. Finally, taxpayers beyond the eligibility threshold, representing 36 percent of the aggregate wage income, experienced a slight increase in their marginal tax rate. The small effect on this group is the result of offsetting effects from an expanded phase-out region for the EITC and lower non-EITC federal tax rates. Turning to the average tax rate relevant for participation, we see that it was reduced for all

Table 4.5
Tax-benefit parameters for single mothers by income groups: Participation elasticity equal to 0.5 and hours-of-work elasticity equal to 0.2

	Phase-in <5,000	Plateau 5,000–6,500	Phase-out 6,500–11,000	Beyond >11,000	Aggregate
Group shares					
Population	0.35	0.13	0.36	0.15	1.00
Wage income	0.10	0.09	0.45	0.36	1.00
Tax benefit parameters					
Marginal tax rate	0.45	0.40	0.45	0.37	0.43
Average tax rate	0.52	0.47	0.53	0.52	0.52
Change in marginal tax rate	−0.05	−0.11	−0.08	0.01	−0.06
Change in average tax rate	−0.06	−0.08	−0.10	−0.06	−0.08
Welfare gain					
Intensive	0.12	0.30	1.15	0.05	1.62
Extensive	0.36	0.37	3.27	1.45	5.46
Total	0.48	0.67	4.42	1.50	7.08

Note: The marginal tax rate and the average tax rate incorporate benefits. See notes to tables 4.3 and 4.4 for an explanation of calculations. The decomposition into different income groups is determined by the income thresholds in the Earned Income Tax Credit before the 1986 reform.

income groups. Again, the reduction is largest for those who were located in the plateau or phase-out regions of the EITC prior to the reform.

The last panel in table 4.5 shows that the welfare gains from TRA86 are concentrated among higher-income female household heads. Females with income in the phase-out group represent the largest source of efficiency gains. In fact, the welfare effect along the extensive margin for this group contributes with nearly half of the total welfare gain. The large welfare effect for this group occurs because its members have the highest wage share, face the highest tax rate on labor force participation prior to the reform, and experience the largest reduction in the participation tax rate following the reform. This correlation among wage shares, initial tax rates, and tax rate reductions is important for the size of the total welfare effect. Our results therefore suggest that accounting for heterogeneity in the tax effects can be critical to the proper evaluations of tax reform.

Our evaluation of tax reform includes only first-order welfare effects and is therefore exact only for small reforms. This approach implies that we do not have to apply specific functional forms. On the other

hand, the tax changes introduced by the 1986 tax act were quite substantial and may have created non-negligible second-order effects. One type of second-order effect derives from the fact that the marginal excess burden is positively related to the size of the tax wedge (cf. equation 14). In the simulations, the welfare gain from the increase in labor supply is evaluated at the pre-reform tax wedge. However, as the tax wedge is reduced, so is the marginal welfare gain. By evaluating the entire labor supply response at the pre-reform tax wedge, we tend to overstate the welfare gains. Another type of second-order effect works through the labor supply elasticities. These may go up or down as we move away from the pre-reform equilibrium. The direction and magnitude of the implied welfare effects depend on third derivatives of the utility function of which we have no knowledge.

4.6 Concluding Remarks

This chapter has argued that recent empirical evidence on labor supply behavior has important implications for the evaluation of tax reforms. Our interest in particular is the emerging consensus that labor supply responses tend to be concentrated more along the extensive margin (labor force participation) than along the intensive margin (hours of work). To understand the implications of the empirical research for evaluating tax reform, we have outlined a simple welfare theoretic framework incorporating (discrete) participation responses. We show that it is necessary to distinguish explicitly between intensive and extensive responses in welfare analysis whenever nonlinearities are present either in the pre-reform tax transfer system or in the policy reform under consideration. This is because labor force participation is related to a different tax rate than hours worked. While the welfare effects from hours-of-work responses are related to the marginal tax rate, the welfare effects from participation responses are instead linked to the effective average rate of taxation (including the average phase-out for transfers). Differences between these two tax rates are driven by the degree of nonlinearity in the tax transfer schedules, particularly important at the lower end of the income distribution.

To examine the quantitative importance of the theoretical points, we presented a microsimulation exercise to evaluate the welfare effects of TRA86 on female household heads. The results suggest that substantial welfare gains were created by the reform, mostly along the extensive margin of labor supply. The application therefore confirms the

quantitative importance of accounting for the extensive margin of labor supply for a particular tax reform. Eissa, Kleven, and Kreiner (2004) extend the results in this chapter to a wider range of reforms. It is shown that some of the points made here are even more important for some of the more recent reforms in the United States. In particular, we are thinking of the tax acts passed in 1990 and 1993, which introduced very large expansions of the EITC.

In general, our results indicate that policy simulations that do not account for (discrete) participation responses may make significant errors. The error will be larger at the bottom of the earnings distribution for two reasons: participation responses are large, and nonlinearities in taxes and transfers tend to be very important. This criticism seems to apply to, among others, Browning and Johnson (1984), Ballard (1988), Triest (1994), and Browning (1995).

Finally, although we have focused exclusively on tax reform in the United States, the results may carry implications for welfare reform in other countries as well. In particular, many European countries are characterized by large taxes and transfers creating very poor incentives to participate in the labor market for low-wage earners. This implies that in-work benefit reform may be a good policy to adopt in these countries. Indeed, the findings of Immervoll et al. (2004) suggest that this is the case. Interestingly, a number of European countries, including the United Kingdom, Ireland, and France, have introduced various forms of in-work benefits in recent years.

Notes

Comments by Jonas Agell, Isabelle Robert-Bobée, Peter Birch Sørensen, two anonymous referees, and participants at the CESifo Workshop on Tax Policy and Labor Market Performance are gratefully acknowledged. We also wish to thank Emmanuel Saez for helpful discussions on this topic. The activities of EPRU (Economic Policy Research Unit) are supported by a grant from The Danish National Research Foundation.

1. Alternatively, one may introduce the non-convexity directly in the preferences through fixed "psychic" costs of participation.

2. These types of discontinuities in benefits and taxes apply, for example, to Medicaid in the United States (cf. section 5.1), to the national insurance tax in the United Kingdom (see Blundell, Duncan, and Meghir 1998), and to housing benefits in some continental European countries (Immervoll et al. 2004).

3. A general framework for analyzing welfare effects with discrete choice is provided by Small and Rosen (1981).

4. Kleven and Kreiner (2002) discuss the marginal cost of public funds in a similar context.

5. Indeed, it is not uncommon that simulation studies based on the standard convex model employ total elasticities in their calibration. For example, this seems to be the case in Browning and Johnson 1984; Ballard, Shoven, and Whalley 1985; Ballard 1988; Browning 1995; and Bourguignon and Spadaro 2002a, 2002b. In all these studies, high female labor supply elasticities (around 0.5–1.0) are used in the calibration, although elasticity estimates of this magnitude tend to be based on censored specifications including observations with 0 hours of work (Mroz 1987; Triest 1990). By implication, these studies are lumping together extensive and intensive responses in the simulations.

6. More detailed information on eligibility and benefits is provided in Hotz and Scholz 2003.

7. The log of earnings (by workers) is regressed on demographic characteristics, including age, education, age-education interactions, race, and state of residence. We also control for self-selection into the labor force using a propensity score correction. The selection term is identified by the number of children.

References

Ballard, C. L. 1988. "The Marginal Efficiency Cost of Redistribution." *American Economic Review* 78: 1019–1033.

Ballard, C. L., J. B. Shoven, and J. Whalley. 1985. "General Equilibrium Computations of the Marginal Welfare Costs of Taxes in the United States." *American Economic Review* 75: 128–138.

Blundell, R. 2001. "Evaluating Labour Supply Responses to In-work Benefit Reforms for Low Income Workers." In *Taxation, Welfare and the Crisis of Unemployment in Europe*, ed. Marco Buti, Paolo Sestito, and Hans Wijkander, 157–187. London: Edward Elgar.

Blundell, R. W., A. Duncan, J. McCrae, and C. Meghir. 2000. "The Labour Market Impact of the Working Families Tax Credit." *Fiscal Studies* 21: 75–104.

Blundell, R. W., A. Duncan, and C. Meghir. 1998. "Estimating Labor Supply Responses Using Tax Reforms." *Econometrica* 66: 827–861.

Blundell, R. W., J. Ham, and C. Meghir. 1987. "Unemployment and Female Labour Supply." *Economic Journal* 97: 44–64.

Bourguignon, F., and A. Spadaro. 2002a. "Redistribution and Labour Supply Incentives: An Application of the Optimal Tax Theory." Mimeo., DELIA (Department and Laboratory of Applied and Theoretical Economics), Paris.

Bourguignon, F., and A. Spadaro. 2002b. "Tax-benefit Revealed Social Preferences: Are Tax Authorities Non-Paretian?" Mimeo., DELIA (Department and Laboratory of Applied and Theoretical Economics), Paris.

Browning, E. K. 1995. "Effects of the Earned Income Tax Credit on Income and Welfare." *National Tax Journal* 48: 23–43.

Browning, E. K., and W. R. Johnson. 1984. "The Trade-off between Equality and Efficiency." *Journal of Political Economy* 92: 175–203.

Cogan, J. F. 1981. "Fixed Costs and Labor Supply." *Econometrica* 49: 945–963.

Dahlby, Bev. 1998. "Progressive Taxation and the Social Marginal Cost of Public Funds." *Journal of Public Economics* 67: 105–122.

Dickert, S., S. Houser, and J. K. Scholz. 1995. "The Earned Income Tax Credit and Transfer Programs: A Study of Labor Market and Program Participation." In *Tax Policy and the Economy*, 9th ed., ed. J. Poterba, 1–50. Cambridge, Mass.: MIT Press.

Eissa, N., and H. Hoynes. 2004. "Taxes and the Labor Market Participation of Married Couples: The Earned Income Tax Credit." *Journal of Public Economics* 88: 1937–1958.

Eissa, N., H. J. Kleven, and C. T. Kreiner. 2004. "Evaluating Tax Reform: Labor Supply at the Intensive and Extensive Margins." Mimeo.

Eissa, N., and J. Liebman. 1996. "Labor Supply Response to the Earned Income Tax Credit." *Quarterly Journal of Economics* 61: 605–637.

Heim, B. T., and B. D. Meyer. 2004. "Work Costs and Nonconvex Preferences in the Estimation of Labor Supply Models." *Journal of Public Economics* 88: 2323–2338.

Hotz, V. J., C. H. Mullin, and J. K. Scholz. 2002. "The Earned Income Tax Credit and Labor Market Participation of Families on Welfare." Mimeo., University of California, Los Angeles, Vanderbilt, and University of Wisconsin-Madison.

Hotz, V. J., and J. K. Scholz. 2003. "The Earned Income Tax Credit." In *Means-tested Transfer Programs in the United States*, ed. R. A. Moffitt, 141–197. Chicago: University of Chicago Press.

Immervoll, H., H. J. Kleven, C. T. Kreiner, and E. Saez. 2004. "Welfare Reform in European Countries: A Micro-Simulation Analysis." CEPR Discussion Paper No. 4324, March. Forthcoming in *The Economic Journal*.

Killingsworth, M. R., and J. J. Heckman. 1986. "Female Labor Supply: A Survey." In *Handbook of Labor Economics*, vol. 1, ed. O. Ashenfelter and R. Layard, 103–204. Amsterdam: North-Holland.

Kleven, H. J., and C. T. Kreiner. 2002. "The Marginal Cost of Public Funds in OECD Countries: Hours of Work versus Labor Force Participation." CESifo Working Paper No. 935, Münich.

Meyer, B., and D. Rosenbaum. 2001. "Welfare, the Earned Income Tax Credit, and the Labor Supply of Single Mothers." *Quarterly Journal of Economics* 66: 1063–1114.

Mirrlees, J. A. 1971. "An Exploration in the Theory of Optimal Income Taxation." *Review of Economic Studies* 38: 175–208.

Moffitt, R. 1992. "Incentive Effects of the U.S. Welfare System." *Journal of Economic Literature* 30: 1–61.

Moffitt, R., and K. C. Kehrer. 1981. "The Effects of Tax and Transfer Programs on Labor Supply." In *Research in Labor Economics*, ed. R. Ehrenberg, 103–150. Amsterdam: Elsevier.

Mroz, T. A. 1987. "The Sensitivity of an Empirical Model of Married Women's Hours of Work to Economic and Statistical Assumptions." *Econometrica* 55: 765–799.

Robins, P. K. 1985. "A Comparison of the Labor Supply Findings from the Four Negative Income Tax Experiments." *Journal of Human Resources* 20: 567–582.

Saez, E. 2002. "Optimal Income Transfer Programs: Intensive versus Extensive Labor Supply Responses." *Quarterly Journal of Economics* 117: 1039–1073.

Small, K., and H. Rosen. 1981. "Applied Welfare Economics with Discrete Choice Models." *Econometrica* 49: 105–130.

Triest, R. K. 1990. "The Effect of Income Taxation on Labor Supply in the United States." *Journal of Human Resources* 25: 491–516.

Triest, R. K. 1994. "The Efficiency Cost of Increased Progressivity." In *Tax Progressivity and Income Inequality*, ed. J. Slemrod, 137–176. Cambridge: Cambridge University Press.