Structural Differencing Estimation with Nonlinear Budget Sets*

Che-Yuan Liang#
September 5, 2013

Abstract: I develop a reform evaluation method for evaluating labor supply in nonlinear budget sets which is both structural and quasi-experimental. The model only requires preferences to be convex on the budget frontier and results in a simple three-dimensional labor supply function. The estimation exploits only the between-individual variation provided by tax reforms. I apply the method to evaluate earned income effects of the Swedish earned income tax credit introduced between 2007 and 2010. I find a positive extensive margin effect for married individuals but also negative intensive margin effects due to the absence of a phase-out region.

Keywords: nonlinear budget sets, reform evaluation, structural models, quasi-experimental methods, earned income tax credit

JEL classification: D04, H24, J22

---

* I thank Sören Blomquist, Mikael Elinder, Lennart Flood, Eva Mörk, Håkan Selin, Whitney Newey, and seminar participants at Uppsala Center for Fiscal Studies, Department of Economics, Uppsala University, the Workshop on Public Economics and Public Policy in Copenhagen, and the 28th Annual Conference of the European Economic Association in Gothenburg for valuable comments and suggestions. The Jan Wallander and Tom Hedelius Foundation, the Swedish Research Council for Health, Working Life and Welfare (FORTE), and the Uppsala Center for Fiscal Studies (UCFS) are acknowledged for their financial support.
# Uppsala Center for Fiscal Studies, Department of Economics, Uppsala University, P.O. Box 513, SE-75120, Uppsala, Sweden; E-mail: che-yuan.liang@nek.uu.se.

1. Introduction

Optimal design of tax systems requires knowledge about the behavioral effects of taxation. There are, however, many challenges in estimating the effects of taxes on, e.g., labor supply in the form of hours of work or earned income.[1] A major difficulty is that real-world tax systems are complicated and create nonlinear budget sets. Structural methods model behavior in such environments and can fully account for the complications. Another difficulty is that budget sets are not randomly assigned. Quasi-experimental methods exploit variation in budget sets provided by tax reforms, which is likely random.

In this paper, I develop a combined structural and quasi-experimental reform evaluation method, structural differencing. The method structurally accounts for the complications created by nonlinear budget sets while imposing only a weak restriction on behavior *and* exploits the variation provided by reforms for identification. As an empirical application, I evaluate the earned income effects of the Swedish earned income tax credit (EITC) reform implemented between 2007 and 2010, which is a reform that has not previously been successfully evaluated using quasi-experimental methods (Edmark et al., 2011).

The static labor supply model that I develop only requires the utility maximization assumption that the preferences on the budget frontier are convex in labor supply. This assumption implies that there is a uniquely preferred (but not necessarily practically available or chosen) option for every individual. The restriction allows for but does not limit behavior to utility maximization with convex preferences in convex budget sets. The resulting labor supply function consists of a three-dimensional integral that depends on the net income[2] levels, the net income derivatives (with respect to labor supply), and the labor supply levels on the budget frontier. This function represents as an intuitive and parsimonious extension of a labor supply function on a linear budget set, which would be two-dimensional and depend on the net income level and derivative (at some fixed labor supply level) on the budget frontier.

The expected labor supply function conditional on the budget set can be approximated flexibly with polynomial terms and estimated while imposing only the additivity and equality restrictions implied by the model, thereby resulting in transparent terms such as the area under the budget frontier. I further develop a quasi-experimental differencing estimation method to estimate the model and for evaluating reforms. By only using between-individual variation in reform effects, identification can become as credible as in reduced-form quasi-experimental evaluations. I develop this method by constructing and estimating the effects of reform-induced differences in the labor supply function and budget sets. I also develop a method for conducting placebo tests to validate the identifying assumptions when several pre-reform years are available. From the reform-evaluation point of view, complementing the quasi-experimental method with structure allows a more precise characterization of the variation in reform effects between individuals. This makes it possible to exploit and account for *all* differences in reform effects between individuals in different regions of the budget set.[3]

[1] I use the term labor supply, broadly defined as potentially including the effort dimension of work, which was demonstrated to be potentially important by Feldstein (1995, 1999).
[2] In static labor supply models, individuals consume all of their net income.
[3] As opposed to, e.g., only using differences in reform effects on the marginal tax rate.

Traditional structural models of labor supply typically start with utility maximization in piecewise-linear continuous convex budget sets. Estimation can be performed using maximum likelihood methods by introducing distributional assumptions, as in Burtless and Hausman (1978) and Hausman (1985). A modern method that discretizes the labor supply choices like in van Soest et al. (1995), Hoynes (1996), and Keane and Moffit (1998), can handle more complicated, nonconvex budgets. This method focuses on estimation of the utility function but could be sensitive to optimization or measurement errors in labor supply (van Soest et al., 2002).

For the estimation of expected labor supply, Blomquist and Newey (2002) developed a nonparametric method based on least squares that is easier to implement and that does not require distributional assumptions, which they demonstrate to be often too restrictive. The method presented here is related to theirs. However, some nonconvexities in the budget sets are allowed here. Still, the derivation and resulting labor supply function are more intuitive and parsimonious.[4] This result is achieved by imposing additional equality restrictions implied by utility maximization. Both labor supply functions are three-dimensional and approximated flexibly using polynomials before estimation. My labor supply function, however, pools all the variations in different regions of the budget set. This pooling results in the need of fewer approximating terms, weaker identification condition, less sensitivity to the functional form and outliers, and better precision.[5]

The maximum likelihood methods typically use the within-year variation in budget sets, which often depends on a host of observable and unobservable characteristics. In principle, nothing prevents pooling data across years, and Blomquist and Newey (2002) do this to also use the between-year variation created by tax reforms. I proceed one step further by *only* using quasi-experimental *between-individual variation* in *reform effects*. This is achieved by first isolating reform effects on the budget sets and then removing the common reform effect across individuals by accounting for time-fixed effects of, e.g., technological growth and macroeconomic conditions on expected labor supply.[6]

For the purpose of reform evaluation, in some cases, quasi-experimental reduced-form methods can provide satisfactory evaluations. When a reform affects the budget sets of a treated group and there is an untreated group, difference-in-differences estimation like in Feldstein (1995) and Eissa (1995) can determine the full reform effect of interest. However, putting more structure on the analysis by characterizing the first-stage effects of the reform on the budget sets is often desirable for two reasons. First, such knowledge *may* be useful for several things, such as for within or out-of-sample predictions; for coherency with economic theory; or for assessing welfare effects of the reform. Second, we may want to exploit the first-stage variation in reform effects between treated individuals. Often, there is no completely untreated group after the reform, in which case the variation between treated individuals may be the *only* variation provided by the reform, such as for the Swedish EITC.

[4] For instance, my function contains one symmetric term that sums over points on the budget frontier, whereas their function contains two terms of different types, where one sums over segments interacted with kinks.

[5] This could be important for implementability in many applications, especially if some variation is differenced away.

[6] For a labor supply function that depends on the budget set, what time-fixed effects accounts for is transparent, unlike, e.g., for a utility function (which is estimated in the discrete-choice approach), which depends on consumption and labor supply.

Many quasi-experimental methods, such as those presented in Blundell et al. (1998) and Gruber and Saez (2002), add structure by linearizing the budget set where the individual is located.[7] This approach leaves some nonlinear budget set issues unsolved, such as (the possibly counterfactual but potential) behavior in budget set regions in which the linearized budget set does not coincide with the actual budget set. The method in this paper can be seen as complementing a quasi-experimental method with a structural model for the first stage that accounts for nonlinear budget set complications. Structure usually comes at the cost of imposing restrictions on behavior,[8] but the restriction made here is weak. In contrast, linearizing the budget set implies strong functional form assumptions by ruling out other parts of the budget set than where the individual is located to affect behavior. From a precision perspective, the method presented here does require the estimation of more parameters than the linearization method. On the other hand, it also exploits variation from the entire budget set and not only at the point of linearization, which is useful when the between-individual reform variation is complicated and spread out over different budget set regions, such as for the Swedish EITC.

The empirical application presented here evaluates the earned income[9] and participation[10] effects of the Swedish EITC introduced between 2007 and 2010. In-work tax subsidies, lowering tax rates at low and medium income levels, are popular and have been adopted by many countries in the last decade. The Earned Income Tax Credit (EITC) in the U.S. and the Working Families Tax Credit (WFTC) in the U.K. are two leading examples. The proponents believe that such programs encourage participation, especially among disadvantaged groups that contain a significant number of individuals who are out of the labor market, such as single mothers. There is a substantial literature that evaluates these programs,[11] and these papers tend to support the proponents' beliefs.

The Swedish program differs from those in other countries in two important ways. It is universal and does not target particular groups, and the credits that are phased in at low and medium income levels are not phased out, thereby resulting in tax cuts also for individuals at high income levels. The program represents the largest reform in Sweden over the last 20 years and mechanically decreased overall labor-related tax revenues from individuals by approximately 15% in 2010.

A few papers have evaluated the reform effects on labor supply. Professor Flood and coauthors apply the structural discrete-choice method on pre-reform samples of individuals combined with out-of-sample simulations.[12] The results are in line with the evaluation results

[7] When employment participation is the outcome of interest, the typical method is to restrict the individual's choice set to two options, nonparticipation and one more or less arbitrarily chosen/predicted participation option (e.g., Eissa and Hoynes, 2004). This approach could, in a broad sense, be seen as linearizing the budget set between these two points.

[8] Meyer and Rosenbaum (2001) add significant structure to their quasi-experimental method to examine participation effects by accounting for expected taxes over a distribution of labor supply options but need to make functional form assumptions on the utility function (linear) and distributional assumptions on the error term (normality).

[9] The data set used does not contain information regarding hours of work.

[10] I investigate the probability of having a positive earned income and the probability of having income above 100,000 SEK, which is slightly less than a third of the average earned income.

[11] See, e.g., Eissa and Liebman (1996), Eissa and Hoynes (2004), and Meyer and Rosenbaum (2001) for the U.S. and, e.g., Blundell et al. (2005) and Blundell and Shephard (2012) for the U.K.

[12] Flood et al. (2007) and Aaberge and Flood (2008) investigate effects for single mothers, and Ericson et al. (2009) investigate effects for broader groups.

of similar programs in other countries. Edmark et al. (2012) apply the quasi-experimental difference-in-differences method exploiting reform-induced variation in changes in the average net-of-tax rates between individuals in different municipalities and with different predicted income levels to identify the effects on participation. They conclude, however, that the isolated variation used is either too small or not exogenous and cannot provide a credible quasi-experimental evaluation.

In this paper, I exploit previously unused sources of between-individual variation in reform effects on budget sets. Because the (Swedish) EITC formula interacts with the rest of the pre-reform tax and transfer system, which differs between individuals with different (observable) demographic background characteristics, the reform effects on budget sets differ between individuals. Furthermore, this variation in reform effects is different in different reform years, which have different EITC formulas. For married or cohabiting[13] individuals, the variation in the spouses' gross incomes provides additional variation in two ways. First the credits the spouses receive differ, which results in different reform effects on the individuals' unearned incomes in the standard secondary-earner model that I use. Second, the individuals' credits interact with other credits and transfers that depend on family income. All this variation can be harnessed and exploited by the developed method illustrating its usefulness.

I use a combined annual register and survey data set, HEK, that spans 2003-2010, including four pre-reform years in addition to the four reform years. The combination of these data and FASIT, a micro simulation model of the Swedish tax and transfer system, enables a detailed characterization of the sample individuals' budget sets that, in principle, account for all aspects of the Swedish tax code that are relevant for individuals; this characterization is essential for the method presented here. In contrast with previous evaluations, I evaluate all years of the reform, including both the intensive and extensive margin responses, for married men, married women, and singles.[14]

I find that the reform had positive participation effects for couples. The effects on the probability of having positive income are 1.7 percentage points for married men and 2.9 percentage points for married women. Most of the participation effects are, however, at low incomes. Half of the "newly supplied" jobs for married men and most of the jobs for married women provide incomes less than 100,000 SEK (which is slightly less than a third of the average earned income). I also find negative earned income effects for married working men, which lead to overall negative earned income effects of 2.5% for married men. These effects imply an additional 30% revenue loss from married men for the government, in addition to the mechanical revenue loss of the credits. The reform effects on earned incomes for married women and the reform effects for singles are not statistically significant but are imprecisely estimated.

The estimated effects are consistent with the ambiguous theoretical predictions. Extensive margin effects can be positive because of positive substitution effects. Intensive margin effects can be negative because of negative income effects. The estimated effects are short-run effects. The long-run effects may be different if adjustment takes time and because of general equilibrium effects. The results nevertheless suggest that in addition to the program

[13] I use "married" to denote "married or cohabiting" from here on.
[14] Single men and women are pooled for precision reasons.

being costly by also providing medium- and high-income individuals with credits because it lacks a phase-out region, this lack also results in significantly negative intensive margin effects.

The paper proceeds as follows. The next section derives the structural labor supply function. Section 3 describes the quasi-experimental estimation method. Section 4 describes the Swedish EITC and the variation used for identification. Section 5 presents the data and some sample statistics. Section 6 reports the results. The final section concludes.

2. Model

2.1 Assumptions

The individual choice problem at hand is two-dimensional. The individual chooses (c, y) subject to a constraint $c(y)$. Because of the constraint, choosing y uniquely determines c. We are interested in estimating the expectation $E(y)$ conditional on $c(y)$. In the standard static labor supply application, c is consumption or net income, y is some measure of labor supply, and $c(y)$ is the function that describes the consumption possibilities on the budget frontier. In the empirical application of this paper, y is earned income. The problem is complicated by the fact that $c(y)$ could be nonlinear. Real tax and transfer systems typically produce piecewise-linear continuous convex budget sets. Such budget sets can be characterized by the slopes and intercepts (at some y) of each of the linear segments, which gives $c(y)$ a dimensionality of two times the number of segments. At this point, I do not put any restrictions on $c(y)$ except that it is continuous. The continuity assumption is for convenience but can be relaxed.

The individual maximizes the following utility function that is assumed to be continuous, again for convenience:

$$\max_{c,y} U(c, y, \nu) \quad \text{s.t.} \quad c = c(y) \Leftrightarrow \tag{1}$$

$$\max_{y} u(y, \nu) = U(c(y), y, \nu). \tag{2}$$

where ν is an individual heterogeneity term that represents differences in taste for labor supply and consumption. I work with a one-dimensional ν, but it is straightforward to allow ν to be multi-dimensional. The optimization problem results in the preferred labor supply choice:

$$y^d = y_j \quad \text{if } u_j > u_k \quad \text{for all } 0 \leq j, k \leq J, \tag{3}$$

where $u_j = u(y_j, \nu) = U(c_j = c(y_j), y_j, \nu)$ is the utility function on the budget frontier and $0 = y_0 \leq y_j \leq y_J$.

Now, allow the observed labor supply choice y to contain an additive optimization and/or measurement error ε in addition to the desired labor supply y^d according to

$$y = y^d + \varepsilon. \tag{4}$$

This error term can capture that individuals cannot fine-tune their labor supply because of, e.g., job availability issues, and often have to choose among a limited set of labor supply options.[15] A priori, this component can be large and important for many individuals.[16] Assuming that ν and ε are independent of $c(.)$, and letting $F(\nu,\varepsilon)$ be the distribution function, $E(y) = Y\big(c(.)\big) = \iint y(c(.),\nu,\varepsilon)dF(\nu,\varepsilon)$.

Structural methods focus on estimating parameters of Y or U. This can be done by imposing a functional form on $F(\nu,\varepsilon)$ and using maximum likelihood estimation, in which case the parameters of $F(\nu,\varepsilon)$ can also be estimated (e.g., Burtless and Hausman, 1978; Hausman, 1985; van Soest et al., 1995; Hoynes, 1996; Keane and Moffit, 1998). Another way to proceed is to not make any distributional assumptions and to directly and only estimate parameters of Y like in Blomquist and Newey (2002). This can be done by deriving an expression for $Y\big(c(.)\big)$ under certain assumptions on behavior, approximating terms in this expression with polynomials, and estimate the parameters using least squares. They also demonstrate that distributional assumptions may severely bias the results. I follow their approach.

If the budget set is linear, $c(.)$ could be completely characterized by two numbers, such as the unearned income, c_0, (the intercept with the consumption axis) and the slope, c' (the derivative with respect to labor supply) which is the net-of-tax rate when y is earned income. $Y(c_0,c')$ would then be two-dimensional, easy to approximate flexibly using, e.g., polynomials, and easy to estimate by least squares. In nonlinear budget sets, $c(.)$ is complicated and $Y\big(c(.)\big)$ has high dimensionality. Because the number of polynomial terms increases exponentially with dimensionality, reducing dimensionality is crucial for feasible approximation and estimation. To achieve low dimensionality, I only make one assumption on behavior, which is described in Assumption 1.

Assumption 1. Preferences on the budget frontier are convex in labor supply:

$$u_j > u_k = u_l \text{ for all } y_k < y_j < y_l \text{ where } u_k = u_l \tag{5}$$

This assumption requires utility to be increasing in labor supply up to a certain level (the level could be the endpoints) and decreasing thereafter. The assumption guarantees that there is a uniquely preferred (but not necessarily practically available or chosen) level for every individual given her preferences and budget set. Because the requirement is on the budget-constrained preferences, it is a joint assumption on the shapes of preferences and the budget set.[17] The requirement neither implies nor is implied by convex preferences. Two consequences of the assumption are crucial for the derivation of the labor supply function. These consequences are stated in Lemmas 1 and 2.

[15] If we know the options for the individuals, such restrictions can be specified and absorbed by $c(y)$.
[16] Structural models for piecewise-linear budget sets typically produce a disproportional number of individuals who want to locate at kink points. In practice, typically little or no bunching is observed (Saez, 2010), which implies either small behavioral responses or large optimization errors.
[17] Whereas convex preferences are equivalent to the utility function being quasi-concave, convex budget-constrained preferences are equivalent to the budget-constrained utility function and the Lagrangian of the problem in Equation (1) being quasi-concave.

Lemma 1. Assumption 1 implies that there is only one local maximum:

$$y^d = y_j, \text{ if and only if } \frac{du_j}{dy_j} = 0 \quad \left(\Leftrightarrow \frac{\partial U_j}{\partial c_j} c_j' + \frac{\partial U_j}{\partial y_j} = 0\right). \tag{6}$$

Proof. This is the standard first-order condition of optimization. ■

Lemma 2. Assumption 1 implies that there is no interior local minimum:

$$(c_j, y_j)\colon \frac{du_{j-\epsilon}}{dy_{j-\epsilon}} \leq 0 \text{ and } \frac{du_j}{dy_j} > 0, \epsilon \to 0^+, \text{ is empty, for all } y_j, \tag{7}$$

Proof. Suppose not; then, there is a $u_j \leq u_{j-\epsilon}$, and $u_j \leq u_{j+\epsilon}$. Then, there are a $k = j - \epsilon$ and an $l = j + \epsilon$ such that Equation (5) does not hold. ■

Lemma 1 guarantees that any local maximum must also be a global maximum. We can therefore check whether a labor supply level is the preferred one by checking whether it is locally optimal by comparing it with neighboring labor supply levels and ignoring all other levels. Lemma 2 rules out interior budget set points that are worse than their neighbors on both sides. Lemma 2 is a restriction on the convexity of the budget set in relation to the convexity of preferences. In Corollary 1, I further prove that Assumption 1 allows, but does not limit behavior to, utility maximization with convex preferences in convex budget sets.

Corollary 1. Assumption 1 allows for utility maximization with convex preferences in convex budget sets.

Proof. Suppose not, i.e., that we have utility maximization with convex preferences in convex budget sets but that Equation (5) does not hold. Then, there is a $u_j < u_k = u_l$ such that $y_k < y_j = ay_k + (1-a)y_l < y_l$, $0 < a < 1$, and $u_k = u_l$. Convexity of preferences implies that $U(c_a = ac_k + (1-a)c_l, y_j) > u_k = u_l$. Convexity of the budget set implies that $c_j > c_a$ and $u_j > U(c_a, y_j)$. However, then $u_j > u_k = u_l$, which is a contradiction.[18] ■

Convex budget-constrained preferences can accommodate some nonconvexities in the budget set and preferences. This assumption allows nonconvexities in the budget set where the utility on the budget frontier is monotonic in labor supply in the nonconvex region. This is illustrated in the top-left figure in Figure 1, where utility is increasing in the nonconvex region. It also does not specifically rule out cases where individuals may want to locate in nonconvex regions. One such case is illustrated in the top-right figure in Figure 1. Nonconvexities that create non-monotonic budget-constrained preferences are illustrated in the bottom figures in Figure 1. The two figures have in common that an indifference curve crosses the budget frontier twice in the nonconvex region.[19] This implies that there is a utility minimum in this area,

[18] *Strict* convexity is only required for either preferences or the budget set if we want the budget-constrained preferences to be strictly convex.

[19] Given monotonic and complete preferences, we implicitly require single-crossing in the nonconvex areas. Note that double-crossing outside the nonconvex regions is allowed, such as in the top figures.

which is not allowed by Lemma 2. In the bottom-right figure, this nonconvexity creates two globally optimal labor supply options.

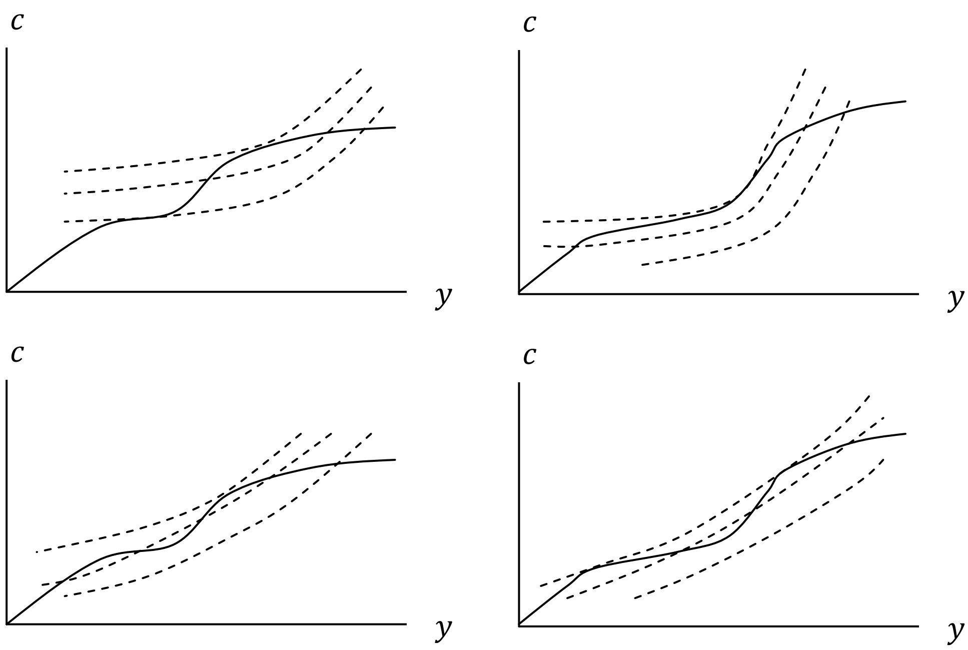

Figure 1. Nonconvexities allowed (top figures) and not allowed (bottom figures)

2.2 Labor supply function

Lemma 1 reduces the global optimization problem of where to locate into many local problems of whether to locate at a specific labor supply level. Each local problem, however, only depends on the first-order condition in Equation (6). This condition only depends on three numbers, c_j, c_j', and y_j. We still need to work over all possible labor supply levels. Because the problem at each level has the same structure, the problems depend on the three numbers in the same way. However, the starting and ending points must be accounted for. In Theorem 1, I derive an expected labor supply expression that imposes as many equality restrictions implied by Assumption 1 as possible. This is achieved by exploiting the repetitive structure implied by Lemma 1. End points are incorporated into the repetitive structure using a decomposition implied by Lemma 2.

Theorem 1. Given Assumption 1, by the law of total expectation, expected labor supply is

$$E(y) = \int_0^{J-\epsilon} \mu\left(c_j, c_j', y_j\right) dy_j,$$

$$\mu\left(c_j, c_j', y_j\right) = \iint\limits_{\frac{du_j}{dy_j}\left(c_j, c_j', y_j\right)>0} \left(y_j + \varepsilon\right) dF(\nu, \varepsilon) - \iint\limits_{\frac{du_j}{dy_j}\left(c_j, c_j', y_j\right)>0} \left(y_j + \epsilon + \varepsilon\right) dF(\nu, \varepsilon). \tag{8}$$

Proof. Lemma 2 implies

$$\begin{aligned}\left\{(c_j, y_j): \frac{du_j}{dy_j} = 0\right\} &= \left\{(c_j, y_j): \frac{du_{j-\epsilon}}{dy_{j-\epsilon}} > 0 \text{ and } \frac{du_j}{dy_j} \leq 0\right\} \\ &= \left\{(c_j, y_j): \frac{du_{j-\epsilon}}{dy_{j-\epsilon}} > 0\right\} - \left\{(c_j, y_j): \frac{du_j}{dy_j} > 0\right\}.\end{aligned} \tag{9}$$

The law of total expectation gives

$$E(y) =^{\text{L1}} \iint\limits_{\frac{du_0}{dy_0} \leq 0} (y_0 + \varepsilon) dF(\nu, \varepsilon) + \iiint\limits_{\substack{\frac{du_j}{dy_j} = 0, \\ 0 < y_j < y_J}} (y_j + \varepsilon) dF(\nu, \varepsilon) dy_j + \iint\limits_{\frac{du_J}{dy_J} \geq 0} (y_J + \varepsilon) dF(\nu, \varepsilon)$$

$$=^{\text{Eq.(9)}} - \iint\limits_{\frac{du_0}{dy_0} > 0} (y_0 + \varepsilon) dF(\nu, \varepsilon) - \iiint\limits_{\substack{\frac{du_j}{dy_j} > 0, \\ 0 < y_j < y_J}} (y_j + \varepsilon) dF(\nu, \varepsilon) dy_j$$

$$+ \iiint\limits_{\substack{\frac{du_j}{dy_j} > 0, \\ 0 \leq y_j < y_{J-\epsilon}}} (y_j + \epsilon + \varepsilon) dF(\nu, \varepsilon) dy_j + \iint\limits_{\frac{du_{J-\epsilon}}{dy_{J-\epsilon}} > 0} (y_{J-\epsilon} + \epsilon + \varepsilon) dF(\nu, \varepsilon)$$

I used Lemma 1 (L1) in the first equality and Equation (9) and the implication of Lemma 2 for the second equality. Collecting terms gives Equation (8).

The labor supply expression in Equation (9) has a very intuitive interpretation. It is an integral over terms at all labor supply levels, where the term at each level depends only on (c_j, c_j', y_j), which are the net income level, the net income derivative, and the labor supply level on the budget frontier. Each term in the integral is three-dimensional. Because the integral is over terms of the same functional form, the dimensionality does not increase.

In practice, $c(y)$ can be obtained by identifying c at different fixed y varied with an interval of δ. This provides a set of discretized points, (c_j, y_j), $j = 0, \dots, J$, on the budget frontier that approximates $c(y)$ increasingly well as the number of points is increased. The integral can then be approximated by summation over the different points j. It is possible to derive a similar version of Theorem 1 that only allows for discretized labor supply options, which would not require continuity of $c(y)$ or $U(c, y)$. The convexity assumption in Assumption 1 would then be replaced by single-peakedness. In fact, only "weak" single-peakedness would be required allowing two adjacent labor supply options on the budget set to be utility-maximizing. The equivalent of Lemma 1 would be that an option is optimal if it is weakly preferred to the two adjacent options. Lemma 2 would become that no interior option can be less preferred to both its two adjacent options. The resulting labor supply function equivalent to Theorem 1 would be

$$E(y) = \sum_{j=0}^{J-1} \mu\left(c_j, c_j{}', y_j\right),$$

$$\mu\left(c_j, c_j{}', y_j\right) = \iint\limits_{\substack{U(c_j, y_j, \nu) \geq \\ U\left(c_j + c_j'\delta, y_j + \delta, \nu\right)}} \left(y_j + \varepsilon\right) dF(\nu, \varepsilon) - \iint\limits_{\substack{U(c_j, y_j, \nu) < \\ U\left(c_j + c_j'\delta, y_j + \delta, \nu\right)}} \left(y_j + \delta + \varepsilon\right) dF(\nu, \varepsilon), \tag{10}$$

where $c_j' = \left(c_{j+1} - c_j\right)/\delta$. Of course, real budget sets may also be discrete rather than continuous.

The derivation here is related to but simpler than the one presented in Blomquist and Newey (2002), which also derives an expected labor supply expression that depends on the entire budget set. A difference is that they assume utility maximization with convex preferences on piecewise-linear continuous convex budget sets. For many tax systems with only minor nonconvexities, such as the Swedish system, they show, however, that the influence of the nonconvexities is negligible. Although both expressions fully describe how the budget set influences expected labor supply for most real budget sets, the expression here is more intuitive and parsimonious.

Both labor supply expressions have a sum term that is three-dimensional. The summation in their case, however, goes over characteristics of the segments and kinks of the piecewise-linear budget set. Because budget sets have different numbers of kinks and segments of different lengths that are located at different labor supply options, the interpretation of the sum term is less intuitive in their case. Their expression also contains another two-dimensional term that depends on characteristics of the last segment. This term can be interpreted as the labor supply function of a linear budget set, which is identified under some conditions, in which case the sum term represents a correction term for nonlinearities in the budget set. [20] The reduction in the number of terms here and the resulting "symmetric" (in each labor supply level) expression is due to the imposition of more equality restrictions implied by utility maximization.

Although convex budget-constrained preferences are assumed in the derivation, the resulting dimensionality structure of the labor supply expression can accommodate more general data-generating processes. The essential ingredient is that the contribution of a labor supply level in a budget set to the labor supply *function* only depends on the level and derivative characteristics of that option.[21] The convexity assumption, however, must hold for the labor supply estimates to be coherent with the behavioral utility maximization model. I will, however, not empirically impose the assumption by constraining the parameters in the empirical specification.

[20] Their labor supply expression, for a piecewise-linear budget set with J segments indexed by j, characterized by $\left(w_j, y_j, k_j\right)$, $j = 1, \dots, J$, where w is the slope, y is the intercept with the consumption axis, and k_j is the kink point, has the form $E(Y) = \bar{Y}\left(w_J, y_J\right) + \sum_{j=1}^{J-1}\left[\theta\left(w_j, y_j, k_{j+1}\right) - \theta\left(w_{j+1}, y_{j+1}, k_{j+1}\right)\right]$. Liang (2012) and Kumar (2012) demonstrate that an additional term $K(w_1, y_1)$ is needed to handle censoring at zero labor supply (which I allow for here).

[21] Individuals may, however, make nonlocal comparisons, e.g., because they have a limited set of options, which can be captured by the optimization error term.

It is possible to derive expected values of functions of y using the same procedure as in Theorem 1. In the empirical application, we are also interested in the probability that y is greater than certain threshold values, i.e., $Pr(y > k)$. The traditional measure of participation probability, where y is hours worked, has the threshold value $k = 0$. When y is some other measure of labor supply, such as taxable income or earned income, other threshold values may be more interesting. We can derive those functions by replacing y with $D(y > k)$, where $D(.)$ is an indicator that takes the value one if the argument is greater than k. This replacement would result in a function with the same three-dimensional structure. Such an extension is analogous to the extension of the method of Blomquist and Newey (2002) presented in Liang (2012). This result holds, however, for any transformation of y, such as a smoothened function of $D(y > k)$.

3. Empirical Strategy

3.1 Polynomial Approximation

The terms in the labor supply functions in Equations (8) and (10) can be approximated flexibly using polynomials and estimated by least squares. Polynomials are convenient because the additivity and equality constraints implied by the model are easily incorporated into the estimation by excluding interaction terms (Stone, 1985) and by imposing equality of coefficients (Porter, 1998). The term is approximated according to

$$\sum_{j=0}^{J-1} \mu\left(c_j, c_j', y_j\right) \sim \boldsymbol{x}_P = (x_1, \dots, x_P), \;\; x_p = \sum_{j=0}^{J-1} c_j^{p_c(p)} c_j'^{p_{c'}(p)} y_j^{p_y(p)}. \tag{11}$$

P denotes the number of approximating terms, p indexes them, and p_c, $p_{c'}$, and p_y are integers. $p_c + p_{c'} + p_y$ is the order of the polynomial approximation.

The approximating terms have intuitive interpretations. The first-order term in c is $\sum c$, which is the area of the budget set below the budget frontier. The other first-order term is $\sum c'$, which is the sum of the net income derivatives between adjacent labor supply levels. This is also the increase in net income between the starting and ending points, $c_J - c_0$. Whereas $\sum c$ is a measure of the average consumption possibilities in the budget set, $\sum c'$ measures how much of these possibilities are due to labor supply rather than unearned income. $\sum y$ is a constant and is dropped.

The second-order terms are $\sum c^2, \sum c'^2, \sum cc'$, $\sum cy$, and $\sum c'y$. The squared terms give different weights to net incomes and net-income derivatives at different net-income and derivative levels. The interaction term between net income and its derivative accounts for derivatives possibly having different importance at different net income levels. The interaction terms with the labor supply level give different weight to net incomes and net-income derivatives at different labor supply levels. The second-order terms characterize budget sets beyond their average levels and starting and ending points by providing measures of *where* net incomes appear in the budget sets. The same holds true for terms of even higher order.

I use terms up to the fourth order. Because many of the higher-order terms describe similar features of the budget set, some of them may be collinear or close to collinear and are dropped in the estimation. The results are, however, quite insensitive to the exact set of higher-order terms included, beyond the second order.

To illustrate how the approximating terms differ for different budget sets, consider the budget sets in Figure 2. In the top-left figure, there are two budget sets with a different $\sum c$. In the top-right figure, there are two budget sets with the same $\sum c$ but a different $\sum c'$. In the bottom-left figure, there are two budget sets with the same $\sum c'$ but a different $\sum c$. Finally, in the bottom-right figure, there are two budget sets with the same $\sum c$ and $\sum c'$ but different second-order terms. We can see how the characterization depends on the properties of different regions of the budget set.

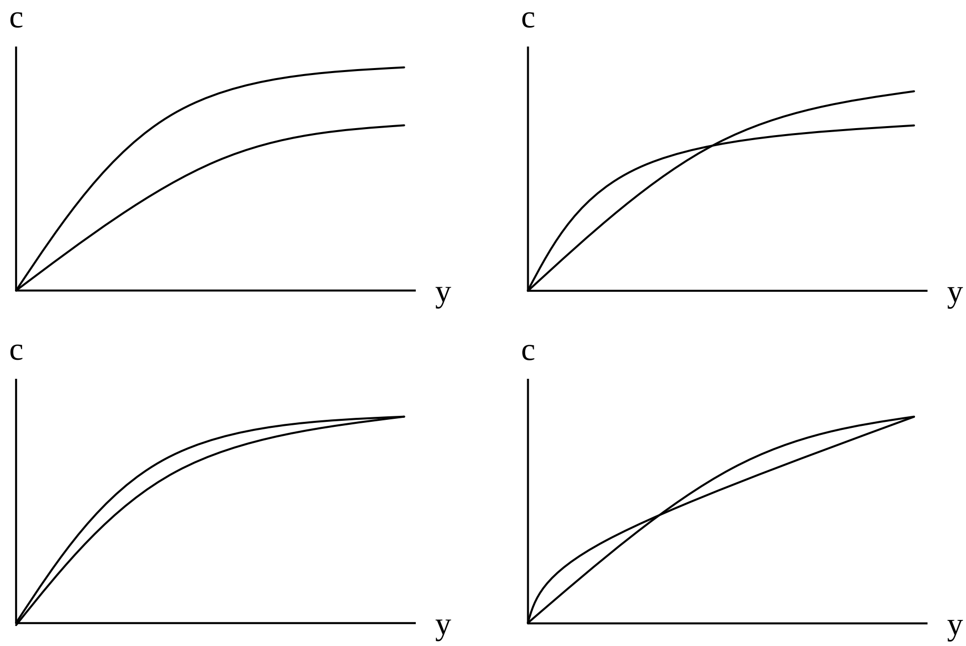

Figure 2. Budget sets with different approximating terms

Notes: The left figures have different budget set areas; the right figures have the same area.

It is also easy to see how the labor supply expression and the approximation of this expression simplify for a linear budget set. In this case, $c_j' = c'$ and $c_j = c_0 + c' y_j$, and hence, $E(y) = \sum \mu(c_0, c', y_j)$. Because c_0 and c' are constant across y_j, for a polynomial approximation, we obtain $E(y) = \sum \mu(c_0, c', y_j) = m(c_0, c') \sum n(y_j) = g(c_0, c')$. This is the standard two-dimensional labor supply function, which depends on unearned income and the slope (net-of-tax rate or net wage) of the linear budget set. An extension to nonlinear budget sets is made straightforward by not only involving the net income level and derivative at one point but at all points. Furthermore, each point may contribute differently, which can be described by interaction effects with the labor supply level of that point. This increases the dimensionality by one.

Blomquist and Newey (2002) also approximate their labor supply expression using polynomials.[22] Because the labor supply expression here is more parsimonious, fewer approximating terms are needed for the expression used here. A fourth-order polynomial requires 30 approximating terms here compared with their 41 terms after dropping fully collinear terms. The difference arises because each term here makes use of characteristics of the entire budget set and pools the variation. In their case, the four first-order approximating terms only use characteristics of either the starting or ending points of the budget set, namely, the first and last segment slopes and intercepts with the consumption axis. Independent variation in these parts of the budget sets is therefore needed for identification, unlike here. Pooling the variation also results in more degrees of freedom and less sensitivity to the choice of functional form (i.e., which approximating terms to include). The approximating terms become smoother, which decreases the sensitivity to outliers and makes the problem of poor support in some regions less likely, which are common problems of higher-order terms.[23] These advantages may not matter in some applications but could be decisive for precision in other applications.

3.2 Structural differencing

Previously presented structural methods typically exploit most of the variation between budget sets to identify the labor supply function. Maximum likelihood methods typically use variation between individuals within years, which is possible because tax systems usually treat different individuals differently (and in the case of hours of work, variation in gross wage rates is often used). Blomquist and Newey (2002) instead primarily rely on variation between years and make use of tax reforms (and in the case of hours of work, changes in gross wage rates).

In the context of tax-reform evaluation, it is often desirable to take it one step further and identify the reform effects using *only* the variation created by a reform. Quasi-experimental methods typically also want to partial out time effects between years and isolate variation between years *and* individuals or groups of individuals. This is only possible if the tax reform has different effects on different individuals.

Exploiting tax reforms could be attractive because changes in tax codes could be more exogenous in nature, or it may be easier to isolate exogenous variation from such changes. Using such variation also avoids out-of-sample predictions in the evaluation, which may rely to varying extents on functional forms. Furthermore, budget set effects due to a reform may differ from other budget set effects, e.g., if there exists little knowledge about the reform or if there are adjustment costs.

However, a limitation of exploiting such variation is that the estimates obtained often cannot identify all behavioral parameters and often are not useful for simulations of non-

[22] They also demonstrate that their sum term (see the second term of the labor supply expression in footnote 20) has a convergence rate that is uniform in an unbounded number of segments for a polynomial approximation using results from Newey (1997). This result is applicable here, thereby resulting in a convergence rate that is independent of the number of points used to approximate the budget frontier.

[23] This is similar to how budget-set-induced consumption possibilities are pooled in the utility function in the discrete-choice method, where the amount of variation is seldom a problem, unlike in the method by Blomquist and Newey (2002).

reform-related effects of the tax system. In the typical quasi-experimental setting with a treated and an untreated group, difference-in-differences methods (e.g., Eissa, 1995; Feldstein, 1995) can provide the full reform effect of interest without the need to identify intermediate first-stage effects on the budget set. However, structure is needed if there are no untreated groups and/or one wants to exploit the between-individual variation in reform effects.

The standard approach to impose structure is to linearize the budget set where individuals locate (e.g., Blundell et al., 1998; Gruber and Saez, 2002) or by restricting the individual's choice set to two options predicted in some way when labor supply participation is the outcome of interest. This approach implicitly puts restrictive assumptions on behavior, namely, that behaviors on nonlinear budget sets are similar to those on linearized budget sets. Issues that are ignored include how to predict and account for nonmarginal labor supply responses, where individuals choose an option for which the true budget set does not coincide with the linearized budget set. Another issue is that the point of linearization is endogenous to the labor supply choice.[24] A third issue is that reforms may have complicated effects on the budget set beyond changing characteristics of the point at which an individual is located. These effects can neither be accounted for nor be exploited by methods that rely on linearizing the budget set, and they may be more or less important in different applications.

The structural differencing approach taken here is to start from a structural model that can account for the nonlinear budget set issues and that is less restrictive on behavior compared with methods that require linearization of the budget set. The task is then to isolate and use the quasi-experimental year-individual variation in budget sets due to the between-individual variation in reform effects for the evaluation. Note that typical methods that impose all the structure that is required to account for nonlinear budget sets impose additional distributional assumptions in the estimation step, which is a significant cost for the structure. Such restrictions are avoided here by focusing on the conditional expected labor supply and by using polynomial approximations.

The budget set effects of a specific tax program caused by, e.g., a reform, can be separated from the rest of the program-exclusive budget set effects in the following manner. Let $\boldsymbol{x}$ be the approximating budget vector of regressors that describes the entire budget set according to Equation (11). Let $\boldsymbol{x}^{Pre}$ be the program-exclusive budget set, i.e., the budget vector without the program. Then, $\Delta\boldsymbol{x} = \boldsymbol{x} - \boldsymbol{x}^{Pre}$ characterizes the effects of the program on the budget set. Variation in $\Delta\boldsymbol{x}$ may arise within years if the program affects the budget sets of different (groups of) individuals differently. It may also arise between years because of a reform, in which case $\boldsymbol{x}^{Pre}$ represents the pre-reform budget set (which is a constructed counterfactual in post-reform years). Pre-reform years then provide untreated budget sets with $\Delta\boldsymbol{x} = \boldsymbol{0}$.

To properly identify the effects of $\Delta\boldsymbol{x}$, it is crucial to account for the factors that determine the differential effects of the program. Time heterogeneity in productivity due to, e.g., differences in macro-economic conditions or production technologies between pre- and post-reform years may affect labor supply and can be accounted for using time-fixed effects, η.

[24] Reverse causation of labor supply on marginal tax rates could, e.g., be addressed using instrumental variables, although the results often depend on the choice of instrumental variables, as demonstrated by Weber (2013).

This approach would remove any reform effects that are uniform across individuals. Introducing time-fixed effects in the standard structural discrete-choice method would be more problematic. It is unclear which time-related factors that affect the utility function of focus are separable from the labor supply and consumption inputs. Additionally, removing a time-fixed labor supply and consumption component does not seem to fully remove the common budget set effects of a reform.

Another issue is that individuals may have been selected into obtaining different program effects on the budget set in the post-reform years. Because factors that determine the selection may have their own effects on labor supply, we need to account for them. If these factors are based on individual (or group) fixed characteristics and we have a panel, individual (or group) fixed effects can account for such factors. However, typically, selection is, based on individual- and time-varying factors. In many reforms, including the Swedish EITC reform, the selection into different program effects depends on the program-exclusive budget set and demographics $\boldsymbol{z}$ that are observable to the tax authorities and the researcher. We can then include these variables as controls to correct for the selection. With regard to the program-exclusive budget set, theory provides us with guidance on how they affect labor supply, namely according to $\boldsymbol{\beta}^{Pre}\boldsymbol{x}^{Pre}$.

I then arrive at the following empirical estimation equation:

$$y_{i,t} = \boldsymbol{\beta}^{\Delta}\Delta\boldsymbol{x}_{i,t} + \boldsymbol{\beta}^{Pre}\boldsymbol{x}_{i,t}^{Pre} + \boldsymbol{\beta}^{z}\boldsymbol{z}_{i,t} + \eta_t + e_{i,t}, \tag{12}$$

where i indexes individuals, t indexes years, and e is an idiosyncratic error term. Budget set terms are approximated according to Equation (11). For participation, we can replace $y_{i,t}$ by $D(y_{i,t} > k)$. I report estimates for $k = 0$ and $k > 100$ (in thousands of SEK). The specification is similar to a continuous difference-in-differences specification. Because $\Delta\boldsymbol{x}_{i,t}$ is determined by $\boldsymbol{x}_{i,t}^{Pre}$, $\boldsymbol{z}_{i,t}$, and t, and I control for the direct effects, I employ the interaction between $\boldsymbol{x}_{i,t}^{Pre}$ and $\boldsymbol{z}_{i,t}$ with t to identify the effects of $\Delta\boldsymbol{x}_{i,t}$. We therefore require $\Delta\boldsymbol{x}_{i,t}$ to be conditionally exogenous.

As an example, consider a reform program that occurs between $t = 0$ and $t = 1$. Let us approximate budget sets using the first-order term $x = \sum c$. Now, assume the program gives individuals who would be highly taxed without the program (low $x_{i,1}^{Pre}$) a large tax cut (high $\Delta x_{i,1}$). We would then identify the effect of the tax cut by comparing individuals with large tax cuts in the post-reform year with those receiving small tax cuts (high versus low $\Delta x_{i,1}$). Because these individuals are different (low versus high $x_{i,1}^{Pre}$), we need to remove the effect of heterogeneity between the groups by using the difference between them in labor supply in the pre-reform years (difference in $y_{i,0}$ for low versus high $x_{i,0}^{Pre}$).

Proper identification of the parameters of budget set changes circumvents some of the challenges of identifying parameters of the entire budget set, which is the typical task of structural methods. Identification only requires identifying and accounting for determinants of the changes rather than the entire budget set. There is, however, a cost of credible identification, which is that $\boldsymbol{\beta}^{Pre}$ are nuisance parameters that should not be given any causal interpretation. Predictions of budget set effects on labor supply should therefore stay within the variation in $\Delta\boldsymbol{x}$ provided by the reform. Including the program-exclusive budget set is, of course, still crucial because of its correlation with the reform effects and labor supply.

Because the budget regressors are related in a mutually dependent and complicated manner and because of the large number of budget regressors used in the estimation, the parameter of a certain budget regressor is completely uninformative. Instead, I report the full reform effect as $\widehat{\boldsymbol{\beta}}^{\Delta}\overline{\Delta\boldsymbol{x}}_{i,t}$ for each of the post-reform years. Standard errors are constructed using the delta method.

3.3 Placebo tests

Because identification relies on the same idea as in other quasi-experimental methods based on differencing, the threats to identification are similar. The most common issues are that treatment ($\Delta\boldsymbol{x}$) may depend on individual and time-varying unobservables and/or that there may be trends in the selection variables. Both are related to the fact that something other than the program may have caused different labor supply in the post-reform years, i.e., that there are omitted variables.

It is difficult to completely rule out such omitted variables. However, when there are several pre-reform years, it is possible to perform placebo tests by placing the reform in one of the pre-reform years and estimating the effect there, which should be insignificant. Of course, there was no such program in the pre-reform years, and what we want to do is to place a counterfactual simulated reform there. In the typical reform evaluation case, the reform variable is binary, and it is easy to replace the treatment variable in the placebo year by one rather than zero and run the same regression as before, dropping the post-reform years. In many cases, such as in the empirical application of this paper, this approach is complicated by the fact that the reform effect is multidimensional and described by a vector ($\Delta\boldsymbol{x}$) and that the reform variables are continuous variables.

There are different possible ways to overcome these difficulties to construct some type of placebo tests that serve as specification tests that are similar to the typical case with one binary treatment variable. I proceed by predicting a one-dimensional reform variable using the available selection variables. After estimating Equation (12), I predict the reform effect $Reform_{i,Post} = \widehat{\boldsymbol{\beta}}^{\Delta}\Delta\boldsymbol{x}_{i,Post}$ in a post-reform year $Post$ for which we want to investigate the estimated reform effect by placing a similar reform in one of the pre-treatment years. This provides a one-dimensional measure of the reform intensity in *that* post-reform year. For that post-reform year, I then regress

$$Reform_{i,Post} = \boldsymbol{\beta}^{Pre}_{Selection}\boldsymbol{x}^{Pre}_{i,Post} + \boldsymbol{\beta}^{z}_{Selection}\boldsymbol{z}_{i,Post} + \omega_{i,Post}, \tag{13}$$

where $\omega_{i,Post}$ is an idiosyncratic error term, to obtain $\boldsymbol{\beta}^{Pre}_{Selection}$ and $\boldsymbol{\beta}^{Z}_{Selection}$, which measures the correlation between the selection variables and the reform intensity. I then predict a synthetic measure of reform intensity, $\widehat{Reform}_{i,t}$, for the entire sample, including pre-reform years. We can verify the suitability of this measure by running the main regression in Equation (12) with observations from the post-reform year of interest and the pre-treatment years but with $\widehat{Reform}_{i,t}$ in place of $\Delta\boldsymbol{x}_{i,t}$ for the post-reform year:

$$y_{i,t} = \boldsymbol{\beta}^{Reform}\widehat{Reform}_{i,t} * D(1|t = Placebo) + \boldsymbol{\beta}^{Pre}\boldsymbol{x}^{Pre}_{i,t} + \boldsymbol{\beta}^{z}\boldsymbol{z}_{i,t} + \eta_t + e_{i,t}, \tag{14}$$

where $Placebo$ is the post-reform year $Post$ and D is an indicator variable. Because the synthetic and original reform measures are highly correlated, we will most likely obtain similar estimates of the reform effect $\widehat{\boldsymbol{\beta}}^{Reform}\overline{\widehat{Reform}_{i,Post}}$, although the precision will be different because we may not have all the selection determinants with the right functional form, because any random variation that lead to reform effects are removed, and because the reform effects on the budget set are not one-dimensional.

It is now possible to perform placebo regressions by placing the reform in one of the pre-reform years and running the regression in Equation (14) on the pre-reform sample, where $Placebo$ is now a placebo pre-treatment year. The placebo estimate can then be constructed as $\widehat{\boldsymbol{\beta}}^{Reform}\overline{\widehat{Reform}_{i,Placebo}}$. For the estimated reform effect to be trustworthy, the placebo-estimate of the reform effect should be small compared with the estimated synthetic reform effect and statistically insignificant. The placebo can be placed in each of the pre-reform years available. When there are many post-reform years, we can construct sets of placebo tests for each of the post-reform years.

4. Institutional background

4.1 The Swedish tax system

With the exception of the EITC program introduced between 2007 and 2010 that I want to evaluate, the Swedish tax and transfer system has been fairly constant since the tax reforms of 1990. The structural differencing method that I employ requires the characterization of entire budget sets. I now first describe the main components of the Swedish tax and transfer system that was in place prior to 2007, without the EITC, and then describe the EITC and its effects on budget sets.

Individuals (as opposed to households) are the basic units of taxation. The labor income tax base varies across different taxes and transfers. The basic local and central government labor income taxes apply to earned income and taxable transfers.[25] The local tax is proportional. The rate varies between 28.9% and 34.2% and it is thickly clustered around the mean of 31.6% in 2010.[26] At a certain threshold, 384,000 SEK in 2010 (all prices are expressed at the 2010 price level), a central government tax rate of 20% becomes effective. At a second threshold, 545,000 SEK in 2010, this rate increases to 25%. The two thresholds have increased to higher income levels over time. Before calculating tax liability, a basic deduction that depends on earned income and taxable transfers is made. The deduction is phased in up to a limit, 131,800 SEK in 2010, and phased out thereafter. This deduction corresponds to introducing tax brackets with different progressive marginal tax rates.

Capital income (included in unearned income that also enters the budget set) is taxed at a rate of 30%. Losses yield (after adjusting for a ratio depending on the type of capital income) tax credits at a rate of 30% for losses up to 100,000 SEK and a rate of 21% for the share of capital income above 100,000 SEK. As with other tax credits, no refunds are pro-

[25] Earned income consists of wage income and income from active self-employment; taxable transfers consist of indirect employment-related incomes such as sickness or parental leave benefits, unemployment benefits, and pension income.

[26] This rate is the sum of the income taxes set by the two lower tiers of government, municipalities and counties. The revenues are used to finance the public services provided by them (primarily childcare, schooling, elder care, and health care).

vided if total tax payments are negative. There is therefore an interaction with labor income taxes, as losses drive marginal labor income tax rates to zero until the credits have been fully exploited.

There are a number of other taxes affecting unearned income or interacting with labor income taxes in a way that differs across individuals. These include taxes on property and wealth, the pension fee, tax credits for the pension fee, household work, and property, and various other minor credits related to labor and capital incomes.

The most important components of the transfer system with respect to their effects on the budget sets of most individuals are the child allowance and the housing allowance. The child allowance is 12,600 SEK for the first child, and the per child allowance increases with each additional child.

The housing allowance is means-tested. Families with children or couples where both partners are under the age of 29 are eligible. The amount depends on the number of children, rents, some other housing characteristics, and the amount and distribution of earned incomes, capital incomes, some nontaxable transfers, and wealth between the spouses. The maximum allowance is 58,800 SEK. It is phased out between a (certain definition of) household income of 58,000 SEK and up to as much as 325,000 SEK for some families.

With respect to indirect taxation, I count it as taxes and include it in the construction of the budget sets. On average, the payroll tax amounted to 30.4% of gross post-payroll-tax earned income in 2010. The value-added tax is 25% of the pre-value-added-tax price on most goods, 12% on some goods, such as food, and 6% on a few goods, such as books. There are also various other consumption taxes. The quotient of aggregate consumption tax revenues divided by aggregate private consumption was 27.7% in 2010.

4.2 The Swedish Earned Income Tax Credit

The Swedish EITC reform was the main component of a broader reform package launched by the center-right coalition government. The key components of this reform were first presented in 2005. The coalition came into power after the parliamentary elections in 2006, after which the details of the reform were established and the budget proposal for 2007 was passed.[27] The EITC is non-refundable and cannot reduce an individual's tax liability below zero. The program has a general character: all individuals below 65 face the same tax credit formula.[28] The credits are automatically implemented in the monthly paychecks and the yearly tax returns, and no special application is needed. This implementation provides individuals immediate and repetitive experiences of the effects of the reform. The program represents the largest reform in Sweden over the last 20 years and mechanically decreased overall labor-related tax reve-

[27] Other reforms between 2007 and 2010 include reforms to the unemployment insurance system and the health insurance system. In general, these reforms made both forms of insurance less extensive. These reforms affected fewer individuals and had only a minor impact on disposable incomes and budget sets in comparison with the EITC. To the extent that they affect the budget sets, these effects are not part of the reform effects on the budget set (and not used for identification of reform effects) but part of the program-exclusive budget set controls (and accounted for).

[28] A more generous tax credit scheme applies to those above 64.

nues from individuals by approximately 15% in 2010.[29] The size of the reform, the automatic implementation, and the pre-election announcement all make it a highly visible program.[30]

The amount of the credit depends on earned income. The credit depends on a basic amount that is inflated each year, the basic deduction, and the local tax rate. The exact formula for 2010 is presented in Table 1. The credit can only be deducted against local taxes. We observe that the amount of credit depends positively on earned income, positively on the local tax rate, and negatively on the basic deduction.

Table 1. The earned income tax credit formula in 2010

Earned income	Earned income tax credit
$< 0.91BA$	$max\big((EI - BD) * t, 0\big)$
$0.91BA - 2.72BA$	$(0.91BA + 0.304 * (EI - 0.91BA) - BD) * t$
$2.72BA - 7.00BA$	$(1.461BA + 0.095 * (EI - 2.72BA) - BD) * t$
$> 7.00BA$	$(1.868\,BA - BD) * t$

Notes: EI is earned income, BD is the basic deduction, BA is the basic amount, and t is the local tax. BA was 42,400 SEK in 2010.

Figure 3 illustrates the extent of the program. The top-right figure illustrates how the amount of the credit varies with gross earned income (net of payroll taxes) in the different reform years for an individual who does not have any taxable transfers, who faces the average local tax rate of 31.6%, and who is not eligible for any other tax credits. The credit is phased in up to an earned income of 334,000 SEK. A specific feature of the Swedish EITC is that there is no phase-out region. The credit is therefore greatest in absolute terms at high income levels. The expansion between 2008 and 2010 was significant and primarily not at the lowest income levels. The maximum credit increased from 12,100 SEK to 21,200 SEK.

To illustrate how the EITC modifies the marginal and average tax rates, tax rates with and without the EITC are plotted in the bottom-left and bottom-right figures in Figure 3. The rates include central and local taxes, the basic deduction, the pension fee, and the pension fee credit, which together are the universal components of earned income taxation. Again, taxable transfers and other individual-varying tax credits are disregarded. We observe that the marginal tax decrease is greatest at low income levels but spans a broader region at medium income levels. The average tax rate changes are smoother across earned incomes and lower at high incomes.

[29] Major expenditure cuts that helped finance the EITC included decreased spending on unemployment insurance, health insurance, higher education and research, and infrastructure investments. On the revenue side, vast privatizations of schools, hospitals, pharmacies, and apartments also generated considerable additional temporary revenues during this period.

[30] Chetty and Saez (2013) and Chetty et al. (2013) demonstrate how information and salience matter for behavioral responses to taxes. Despite the Swedish EITC being highly visible, the exact tax formula is somewhat complicated. It is therefore likely that many individuals did not know the exact details of it in the first reform years, and adjustments may have taken some time. However, this feature is not unique for the Swedish EITC. In fact, many components of the Swedish tax system and of tax reforms in other countries are complicated. No method either requires that the individuals know the tax schedule or reform effects on tax schedules of *other* individuals (i.e., how tax effects vary between individuals). In fact, most methods even allow for optimization errors, which to some degree can capture individuals' lack of information on their own tax schedules. Nevertheless, the behavioral effects identified are conditional on the distribution of knowledge about the tax system. For a tax reform, such knowledge may increase over time. For this reason, long-run effects may therefore be greater than the short-run effects that can be identified using data around the reform years.

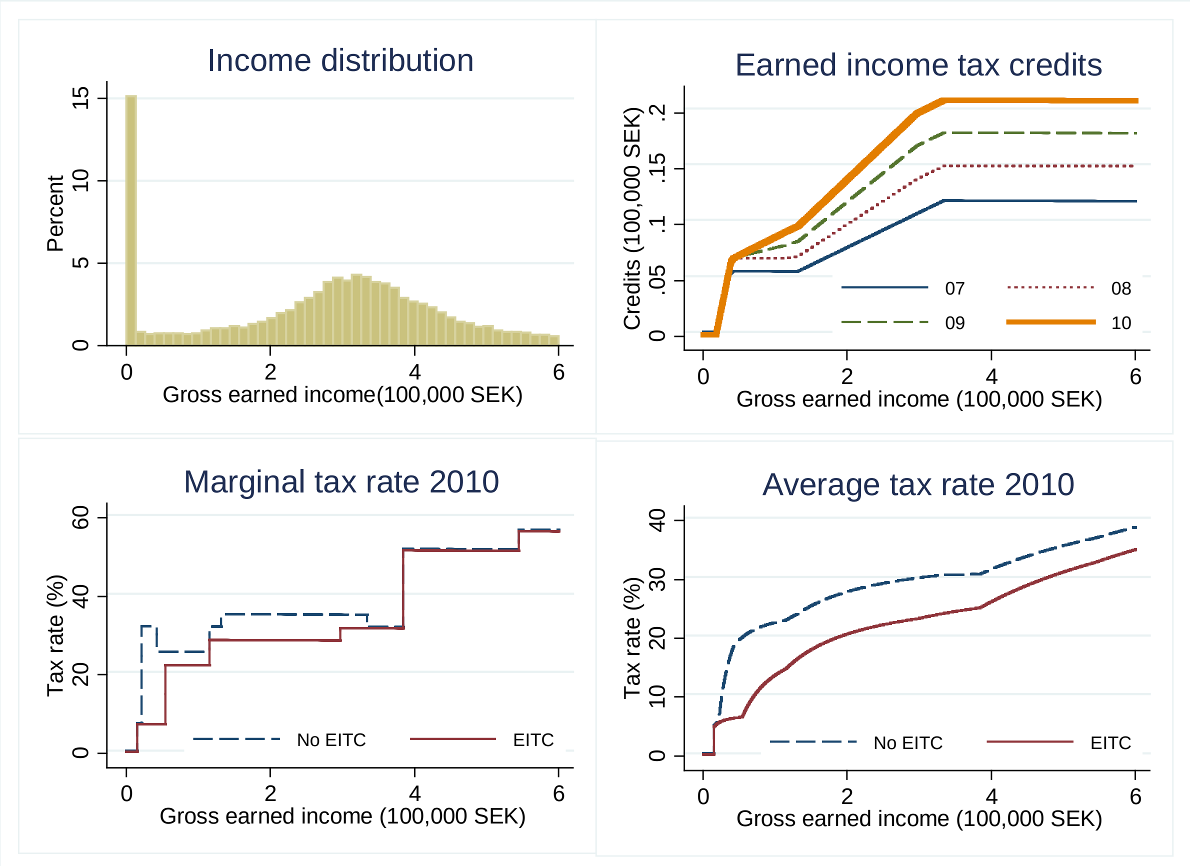

Figure 3. The effects of the EITC on the budget set

The sample distribution of yearly incomes during the reform period is plotted in the top-left figure in Figure 3. We see that nonparticipation in the sense of not having any earned income is approximately 15% for this period. The distribution is then almost bell-shaped and centered at approximately 320,000 SEK. We see from this distribution and the marginal tax rate schedule that most of the individuals in the bottom half of the distribution experience decreasing marginal tax rates. The exception is the nonparticipants.

The phase-in regions increase the marginal payoff of increased labor supply and create uncompensated net-of-tax rate effects at low and medium income levels. It is generally believed that positive substitution effects dominate the eventual negative income effects because most estimated uncompensated elasticities are positive.[31] Except for in parts of the low-income region, the marginal effects are not large (and are even smaller in percentage point terms if indirect taxes are accounted for). To avoid reverse uncompensated net-of-tax rate effects, the Swedish government constructed the program without a phase-out region. However, this design is expensive because it provides most individuals employed full time at medium and high income levels with a credit that is close to the maximum amount of credit. It also results in income effects that increase with earned income. These effects could have substantially adverse effects at the intensive margin. The reform effects on earned incomes are therefore ambiguous from a theoretical perspective.

[31] For Sweden, see, e.g., Blomquist and Hansson-Brusewitz (1990), Blomquist and Newey (2002), Hansson (2007), Blomquist and Selin (2010), and Liang (2012).

At the extensive margin, the EITC makes all participation options more attractive than nonparticipation. The reform effects on participation should therefore be nonnegative for single men and women.

For married men and women, within the secondary-earner framework used here, there are additional effects on unearned incomes as the spouses' net incomes increase.[32] These income effects may create negative incentives at both the intensive and extensive margins for couples, thereby making the expected directions of the reform effects ambiguous for them.

4.3 Variation used for identification

One way to evaluate the EITC would be to exploit the between-year variation in budget sets induced by the program. This can be achieved through a pre- and post-reform comparison of labor supply or in a more sophisticated manner by exploiting all four different stages of the program. Because of other time effects, such an approach would not be very convincing. Instead, we would like to exploit existing between-individual (or between-group) variation in reform effects. However, because of the universal nature of the program, it is not apparent that any such variation exists. Below, I demonstrate that the way in which the EITC interacts with the remainder of the pre-reform tax and transfer system creates such, previously unexploited, variation.

First, the EITC effects on budget sets vary across local governments because of differing local tax rates. This variation generally changes the credit and marginal tax rates proportionally at most gross earned income levels and results in different maximum potential credits among individuals in a given year. However, this variation, which was exploited by Edmark et al. (2012), is small because the variation in local tax rates is small and thickly clustered around the mean.

Second, because different tax credits are deducted against the same taxes and the credits are nonrefundable, they offset one another at different earned income levels. If there are more credits than taxes at a certain level of earned income, the remaining credits are postponed until the individual reaches a higher income level. The replacement between the EITC and other tax credits is one to one but only for local taxes. Interactions arise between the EITC and, e.g., the property tax and capital loss credits, which vary across individuals.

Third, the EITC cannot be deducted against taxes levied on taxable transfers. At a given level of earned income, the amount of taxable transfers affects the local tax on earned income against which the EITC can be deducted. As these transfers increase, marginal tax rates increase. This affects the income levels at which the EITC can offset local taxes.

Another effect of taxable transfers operates through the basic deductions. At a given level of earned income, as taxable transfers increase, basic deductions increase, which decrease the EITC (see the formula in Table 2). The sum of this and the previously mentioned (sometimes counteracting) effects of taxable transfers on the EITC results in a complicated pattern of how the EITC modifies the tax schedule.

To illustrate some of the between-individual variation in reform effects on budget sets, I plot the reform effects on marginal tax rates (excluding indirect taxation and taxes that do not

[32] However, the secondary-earner framework ignores cross-effects of the budget set (and changes in the budget set) of the secondary earner on the labor supply of the primary earner.

interact with the EITC regarding marginal tax rates) for four different single individuals in the left figure in Figure 4. The effect for an individual who faces the average local tax rate of 31.6%, without taxable transfers and other tax credits, is represented by the thick solid line. This is the difference between the marginal tax rates without the EITC minus the marginal tax rates including the EITC in Figure 3. I then increase the local tax rate to the maximum rate of 34.2% and plot the reform effects with a dotted line. We can observe a small upward shift in reform effects.

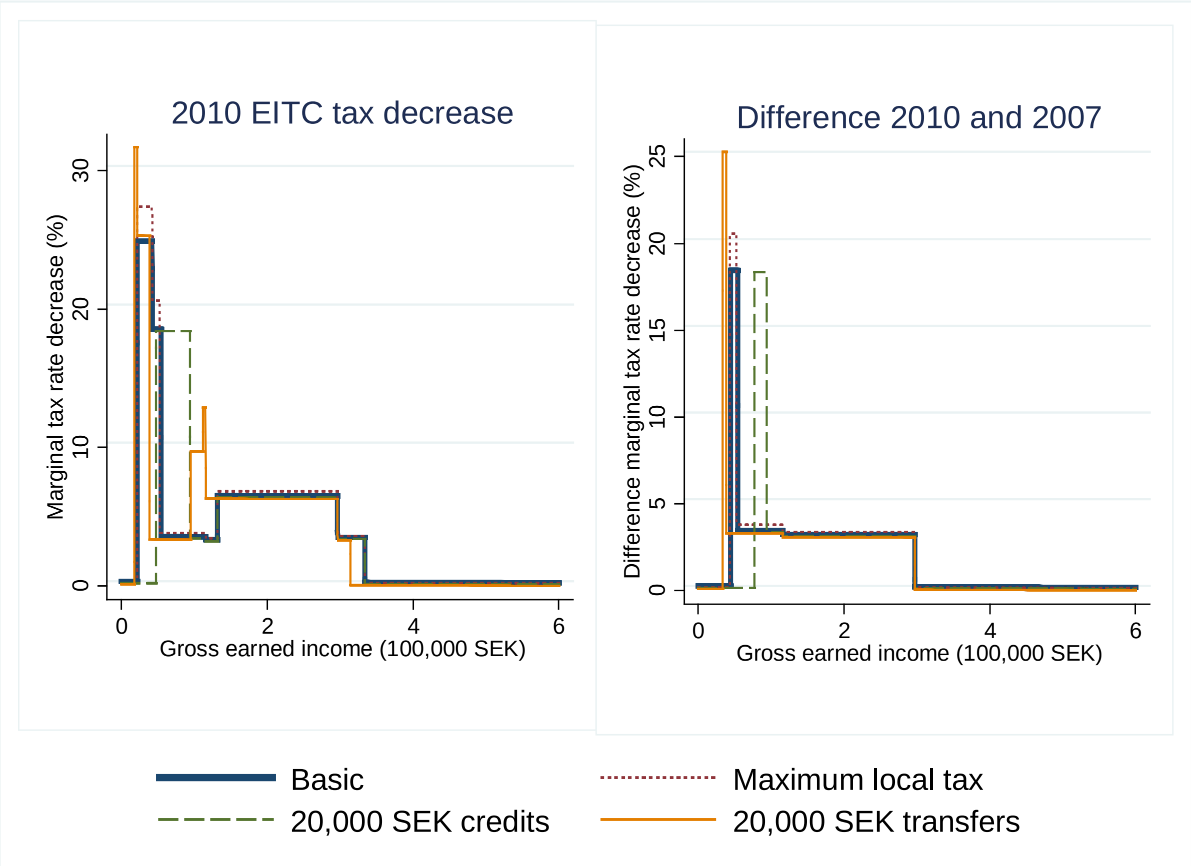

Figure 4. Changes in marginal tax rates due to the EITC

Next, I instead increase other tax credits to 20,000 SEK. This is a sizable but not uncommon amount because 20% of the observations in the selected sample had credits greater than this sum. The resulting reform effect line is represented with a dashed line. We observe how the reform effects are shifted towards higher income levels, although the shift is not simply horizontal.

Finally, I instead increase taxable transfers to 20,000 SEK, which is again a sizable but not uncommon amount because 10% of the observations in the selected sample had transfers greater than this sum. The resulting reform effects are plotted with a solid line. We observe that some segments of the reform effects now appear at lower income levels and other segments appear at higher income levels.

Because the EITC formula differed in each of the years between 2007 and 2010, the differences in reform effects across individuals also vary across years. This difference can be visualized by constructing graphs of the reform effects for each of the reform years and comparing the graphs for the different years. I perform this comparison for the same prototype individuals as before but using the 2007 EITC and construct the differences in reform effects between 2010 and 2007. These differences across reform years are plotted in the right figure in Figure 4. The positive curves are due to the 2010 EITC being more extensive, with additional tax reductions for all individuals. However, we observe that this extension was not uniform but rather led to additional differences in reform effects across individuals.

A major part of the between-individual variation consists of *when* (at which labor supply level) the credit becomes available. There are several reasons why this variation could lead to different labor supply responses (and therefore represent variation in reform effects that can be exploited). First, credits at lower income levels lead to income effects at higher income levels. How soon the taxes appear in the budget set therefore determines how large the income region suffering from income effects will be. Second, there is no reason to expect labor supply to be equally sensitive to taxes in different income regions. Additionally, the effect of the same credit may depend on the rest of the budget set unrelated to the EITC, for which the tax and transfer system creates substantial differences (e.g., the child allowance differs). All such effects produce variation in the changes in the shapes of budget sets and results in variation in the approximating terms, such as variation in the change of the area under the budget set ($\Delta \sum c$) of the structural model (see the discussion in Section 3.1 and Figure 2).

For married individuals, there is a large indirect source of variation in the spouses' gross earned incomes. Spouses with different gross earned incomes receive different credits, which leads to different reform effects on *the individuals'* unearned incomes. This variation is plotted in Figure 5. The variation for married men is shown in the top figures, and the variation for married women is shown in the bottom figures. The variation for all reform years is shown in the left figures, and the variation for 2010 is shown in the right figures. We observe that there is continuous variation between 0 and 20,000 SEK for both married men and women, with a higher density at higher changes for married women, which reflects the higher gross earned incomes of married men. For 2010, we observe that the distribution has a higher density at higher incomes compared with the graph for all reform years, which reflects that the EITC was expanded each reform year.

The variation in the spouse's gross income also affects some transfers and tax credits of the individuals, e.g., the housing allowance, because the income base sometimes relies on family income. These transfers in turn affect the reform effects. There is therefore substantially more between-individual variation in reform effects for married individuals than for singles.

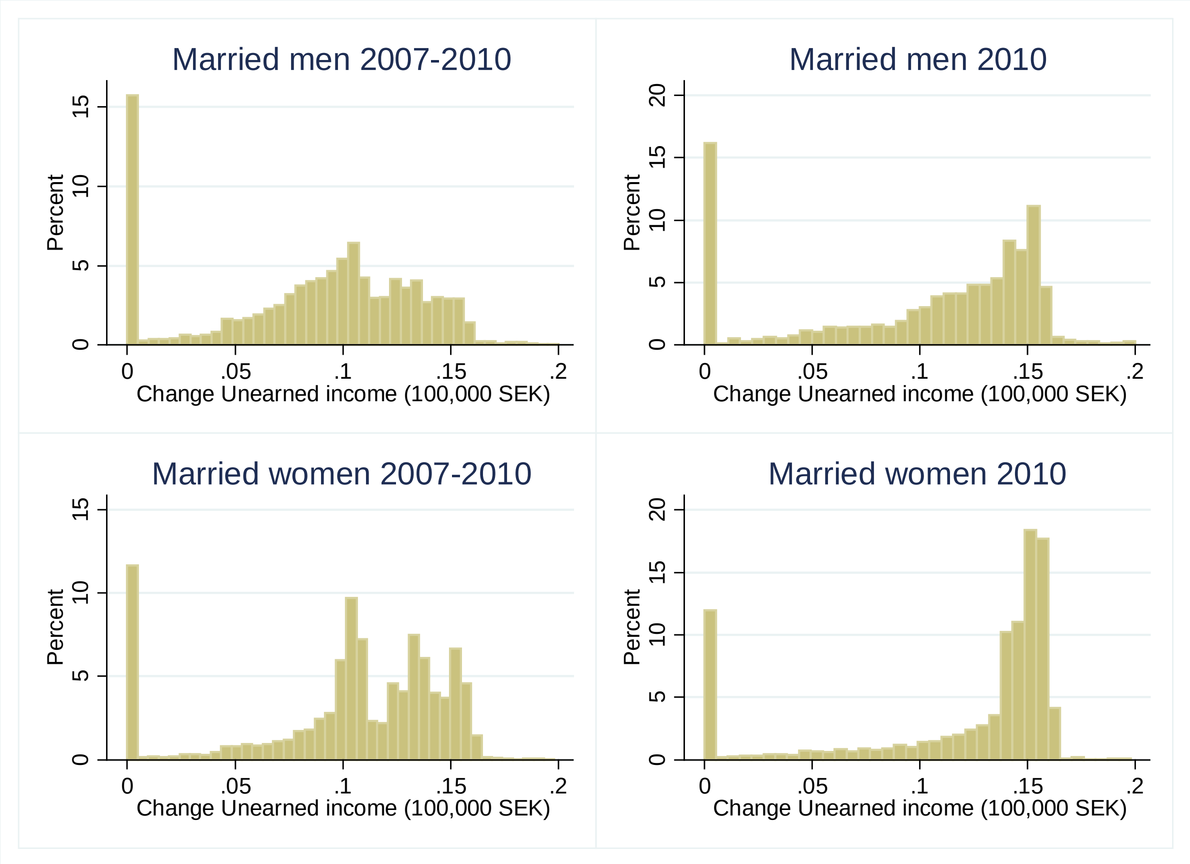

Figure 5. Changes in unearned incomes due to the EITC

5. Data

5.1 HEK and FASIT

I use data from HEK (Hushållens Ekonomi), which is a combined registry and survey data set provided by Statistics Sweden. The data set contains repeated cross sections of approximately 17,000 randomly sampled individuals from the population and members of their households for each year. The response rate is approximately 70%. The registry component contains income, tax, and demographic information used by the authorities for taxation purposes and contains the same variables as in the more well-known LINDA (Longitudinal Individual Data) data set (Edin and Fredriksson, 2000). The survey component primarily contains the housing variables required to construct several housing-related budget set variables such as the housing allowance; these variables are important components of the budget sets.[33] I use data from 2003-2010, which includes four years with different EITC programs and four pre-program years with only minor changes in the tax code.

To construct the individual budget sets, I use a micro simulation model, FASIT, developed by Statistics Sweden. In principle, FASIT captures all features of the Swedish tax and

[33] There are registry data regarding housing-allowance related variables for individuals that receive housing allowances. However, tax authorities do not collect such data for individuals that do not receive housing allowances. I need such (survey) data to compute tax and transfers at *hypothetical* earned incomes at which non-receivers *would* receive housing allowance, because I need the entire budget sets.

transfer system that are relevant for individuals. Some aspects of the pension system are missing in the software, but I exclude retired individuals from the sample. FASIT is used by, e.g., the Swedish Ministry of Finance, to simulate the mechanical effects of different tax policies, including potential future policies. Because structural differencing exploits variation in most parts of the budget set, it is important to accurately and completely characterize the budget sets, which FASIT makes possible. Single cross sections of this model have been previously employed by Flood et al. (2007), Aaberge and Flood (2008), and Ericson et al. (2009). Blomquist et al. (2012) and this paper are, however, the first to employ several years of FASIT and variation in budget sets over time in the empirical analysis.

I construct the budget sets by iteratively letting FASIT calculate net family incomes by varying individuals' gross earned incomes. I set unearned income to be the net income the family receives (would receive) if the individual does not have any gross earned income and the spouse has its actual income. Additionally, I include the implicit income from residence-owned housing in unearned income. I also correct for indirect taxation. Payroll taxes are generated by FASIT, whereas I make a simple rudimentary correction for consumption taxes using the quotient of aggregate consumption tax revenues divided by aggregate private consumption for each year. These additional corrections are similar to those in Blomquist et al. (2001) and Blomquist et al. (2012).

It is common in the labor supply literature to select a highly restrictive sample. One reason for the selection is that it is difficult to satisfactorily characterize the budget sets of certain individuals. Another reason is that the two-dimensional, consumption-labor supply space may not capture the essential dimensions of choice for certain individuals, e.g., students. A third reason is that the effects of taxes may differ across groups. The selection made here is very similar to those made in previous labor supply papers. I limit my sample to individuals who are between 21 and 60 years old. I also exclude those that receive sick benefits, parental benefits, income from self-employment, or student financial aid that is greater than half of the average monthly gross earned income. For sampled individuals who receive incomes from those sources, I consider these incomes to be part of unearned income. I also drop individuals with earned or unearned income greater than 1 million SEK. These individuals make up approximately 1% of the sample once all other restrictions have been imposed. The selection leaves 100,977 out of 237,384 observations. Besides decreasing standard errors, the results are quite insensitive to most of these sample restrictions and very insensitive to smaller variations in sample restrictions.

For the remainder of the sample, I conduct the analysis separately for married or cohabiting men, married or cohabiting women, and singles. I pool single men and women because precision is very low since there is substantially less variation in reform effects on budget sets, and because there are much fewer observations.

The data set contains a rich set of demographic variables. Because taxes and transfers depend on demographics, as discussed in the institutional background section, and demographic characteristics may have own effects on labor supply, it is important to control for them, as discussed in the empirical strategy section. As demographic control variables, $\boldsymbol{z}$, I include group dummies based on age (8 groups), number of children below six years old (3 groups), education level (7 groups), occupation (8 groups), and county of residence (21 groups), and whether the individual is born in Sweden or not. For singles, I also include a control for gender.

5.2 Sample statistics

The means and standard deviations of some variables in the regressions are reported in Table 2. Each row first reports the statistics for one variable for the entire sample period, then for the pre-reform period 2003-2006, and finally for the post-reform period 2007-2010. The statistics for gross earned income (with pre-payroll tax), and the probability of having positive income and income above 100,000 SEK, are reported in rows 1-3. Marital status, gender, age, and the number of children below six years old are reported in rows 4-7. For the budget regressors, I focus on the most intuitive sum of net income variables and only up to quadratic terms. Most other terms are highly correlated with these terms.[34] I report sum of net incomes $\sum c^{Pre}$ and sum of net income squared $\sum c^{Pre^2}$ for the EITC exclusive budget sets, and the EITC effects on those variables, $\Delta \sum c$, and $\Delta \sum c^2$.

Table 2. Variable means and standard deviations by time period

	Entire sample		Pre-reform		Post-reform	
Variable	Mean	Std. dev.	Mean	Std. dev.	Mean	Std. dev.
y	3.354	2.041	3.350	1.949	3.358	2.124
$D(y > 0)$	0.888	0.315	0.905	0.293	0.872	0.334
$D(y > 100)$	0.834	0.372	0.854	0.353	0.815	0.388
$married$	0.696	0.460	0.720	0.449	0.673	0.469
man	0.518	0.500	0.522	0.500	0.514	0.500
age	42.622	11.224	42.642	11.131	42.602	11.311
$children6$	0.147	0.430	0.141	0.421	0.152	0.439
$\sum c^{Pre}$	3.095	1.479	3.095	1.408	3.096	1.544
$\sum c^{Pre^2}$	12.334	12.725	12.117	11.745	12.541	13.589
$\Delta \sum c$	0.074	0.085	0	0	0.144	0.061
$\Delta \sum c^2$	0.558	0.803	0	0	1.089	0.826

Note: y, $\sum c^{Pre}$, $\sum c^{Pre^2}$, $\Delta \sum c$, and $\Delta \sum c^2$ are in 100,000 SEK.

Monetary variables are expressed at the 2010 price level, adjusted for inflation using the CPI, and in 100,000 SEK. The budget set variables are created using numerical integration over 100 points with a distance of 10,000 SEK in gross earned income between the points and normalized by the number of terms in the sum to provide a measure of the average net income in the budget set.

We observe that average earned income rose by 0.2% from 3.350 to 3.358. The two participation rates fell by 3-4 percentage points (0.905 to 0.872 and 0.854 to 0.815 respectively) between the pre- and post-reform periods.[35] 69.6% of the sample consists of married couples, and 51.8% of the sample is men. The average age is 42.6 years, and the average number of children below six years old is 0.147. The marriage rate fell during this period.

[34] Other terms are, however, included in the regressions, although some are dropped due to perfect or near perfect collinearity. The main fourth-order specification contains, e.g., around 30 budget regressors.

[35] The income and participation trends are partly due to the sample selection that drops individuals with income above 1 million SEK and individuals receiving a large amount of transfers. The size of these groups of individuals changes over time. The results are, however, quite insensitive to these selection criteria.

For the budget set variables, the EITC exclusive average net income was 309,500 SEK, with a standard error of half that amount. The amounts were similar between the pre- and post reform periods reflecting that there were no other tax reforms with major effects on the budget sets besides the EITC during this period. The sum of the net income square terms increased, however, between the periods reflecting that net incomes in the budget set appear at higher gross earned incomes. This is mainly due to adjustments in starting points of different tax brackets.[36]

Turning to the EITC effects on the budget sets, there were, by construction, no effects in the pre-reform period. In the post-reform period, EITC increased average net income by 14,400 SEK. We are, however, more interested in how the effect varies between individuals. The standard errors provide such a measure, and we see that they are 6,100 SEK or 42.4% of the average net income. The sum of square terms reveals a similar picture. There is therefore plenty of variation in EITC effects on budget sets in the post-reform period.

All the variation in EITC effects in Table 2 is not useful because some of the variation in reform variables arises between different groups and between different post-reform years. Histograms of $\Delta \sum c$ are plotted for couples and singles separately in Figure 6. The top figures illustrate the distribution for couples and the bottom figures illustrate the distribution for singles. The left figures pool all post-reform years, whereas the right figures show the distributions for 2010. We observe that the EITC effect varies between 5,000 SEK and 30,000 SEK for couples and between 3,000 SEK and 15,000 SEK for singles, confirming the picture that reform effects vary substantially between individuals.

The figures for 2010 show substantial variation also within a post-reform year. Because the credits were increased over time, the impacts in 2010 are in the high end of the distribution for all post-reform years. Although we control for time effects exploiting only variation within years, it is still useful that the reform effect range differs between post-reform years, because this increases the support of the reform variable. The figures also confirm that there is much more variation for couples than for singles.

To examine the determinants of the reform variables, I report a correlation matrix between the variables in Table 3 for 2010 (the matrix is cut so no correlation is repeated to save space). We observe that the reform variables are correlated with the program-exclusive budget set variables and the demographic variables. This is because the program effects were not randomly assigned. The correlations between the reform effect on average net income and the three labor supply measures are positive indicating that getting the credit early is associated with higher labor supply.

[36] In particular, the starting point of the top tax bracket increased. Such adjustments may affect the budget sets even without other tax reforms. Even lack of adjustments, inflation has “bracket creep” effects. I exclude any such effects and only include EITC reform effects on the budget sets in $\Delta \boldsymbol{x}$.

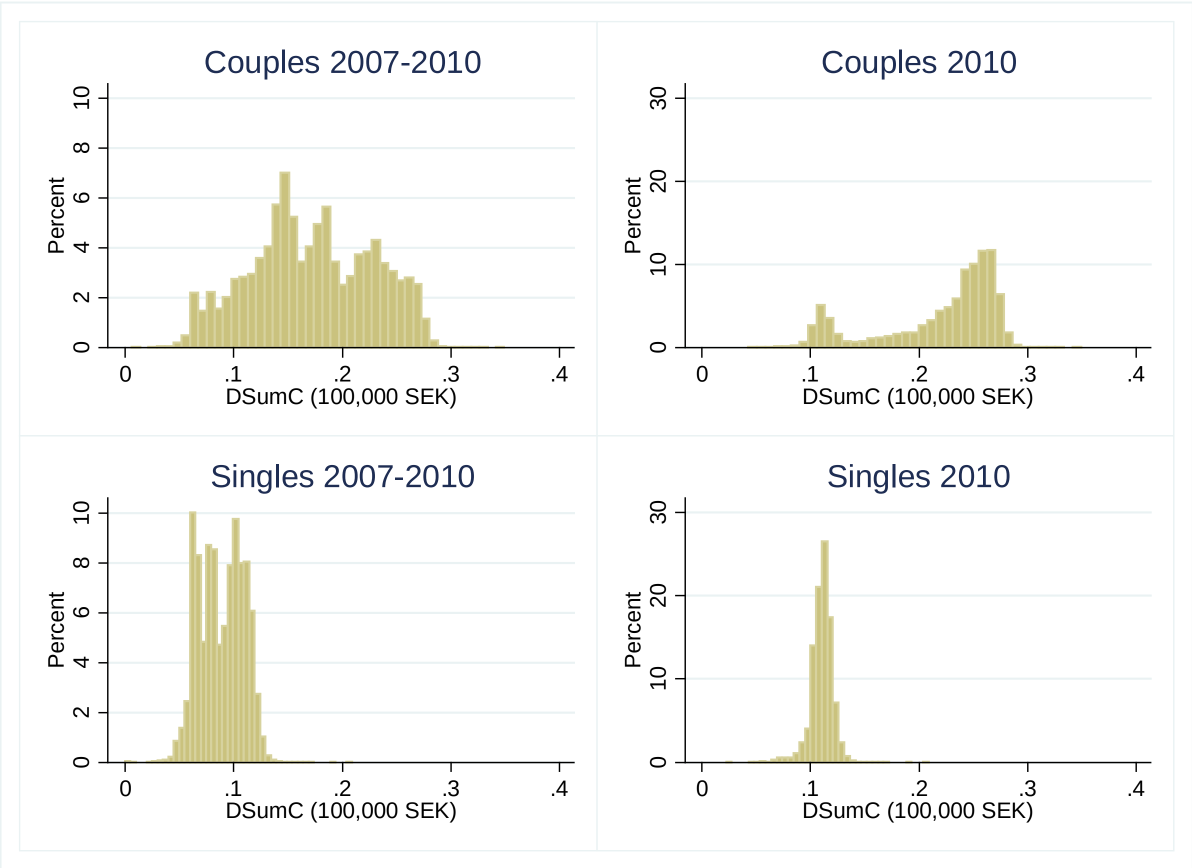

Figure 6. Variation in EITC impact on the budget set

Table 3. Correlation matrix

	$\Delta\sum c$	$\Delta\sum c^2$	$\sum c^{Pre}$	$\sum c^{Pre\,2}$	$married$	man	age	$children6$	y	$D(y>0)$
$\Delta\sum c^2$	0.87									
$\sum c^{Pre}$	0.71	0.94								
$\sum c^{Pre\,2}$	0.55	0.86	0.94							
$married$	0.76	0.68	0.64	0.47						
man	-0.11	-0.14	-0.12	-0.11	-0.02					
age	0.28	0.30	0.31	0.24	0.34	-0.06				
$children6$	0.06	0.08	0.11	0.08	0.18	0.03	-0.21			
y	0.22	0.26	0.26	0.23	0.21	0.16	0.23	0.01		
$D(y>0)$	0.17	0.15	0.12	0.10	0.12	0.03	0.04	0.01	0.60	
$D(y>100)$	0.19	0.17	0.15	0.12	0.15	0.02	0.11	0.01	0.73	0.79

6. Results

6.1 Reform effect estimates

Estimates of the effects of the full EITC program in 2010 are reported in Table 4. The results are reported for gross earned income in the first section and for the probability of having positive income and income greater than 100,000 SEK (less than a third of the average earned income) in the second and third sections, respectively. The earned income estimates are reform estimates in % relative to the observed earned incomes in 2010. The participation estimates are expressed in percentage points. The results are reported separately for married men, married women, and singles. Each estimate is the result of a separate regression according to Equations (11) and (12). Time dummies, a fourth-order polynomial in the reform-exclusive budget regressors $\boldsymbol{x}^{Pre}$, the full set of demographic control variables for age, number of children less than six years old, education level, occupation group, county of residence, country of birth, and sex are included in all reported specifications.

Table 4. Estimates of reform effects as polynomial order is varied

		First order	Second order	Third order	Fourth order
y	Married men	-1.378	-2.871**	-2.535*	-2.498*
		(1.403)	(1.430)	(1.447)	(1.447)
	Married women	3.017*	2.293	2.455	2.137
		(1.583)	(1.599)	(1.622)	(1.631)
	Singles	4.268	3.937	3.792	3.424
		(4.226)	(4.227)	(4.228)	(4.229)
$D(y > 0)$	Married men	1.872***	1.807***	1.676***	1.687***
		(0.318)	(0.324)	(0.328)	(0.328)
	Married women	3.291***	3.111***	2.954***	2.907***
		(0.378)	(0.381)	(0.387)	(0.389)
	Singles	1.735	1.646	1.603	1.620
		(1.145)	(1.145)	(1.146)	(1.146)
$D(y > 100)$	Married men	0.821*	0.750	0.898*	0.910*
		(0.481)	(0.491)	(0.497)	(0.497)
	Married women	0.547	0.342	0.514	0.588
		(0.588)	(0.594)	(0.602)	(0.605)
	Singles	2.793	2.754	2.726	2.630
		(1.698)	(1.698)	(1.699)	(1.699)

Notes: Each estimate is an estimate from a separate regression. For earned income, estimated reform effects are in % of the observed earned incomes. For participation, estimated reform effects are in percentage points. Standard errors are constructed using the delta method. * significant at 10%, ** significant at 5%, * significant at 10%.

To check the sensitivity to the functional form of the labor supply function, the polynomial order for the reform budget regressors $\Delta\boldsymbol{x}$ is varied horizontally. We observe that the results occasionally depend on the polynomial order. All estimates are fairly stable after adding second order terms. The standard errors increase with polynomial order.

The earned income reform effect estimate for married men increases in magnitude when the order is greater than two compared with the linear specification. The estimated reform effect is negative and in the region of -2.5%. It is statistically significant at the 5% level in the second-order specification but only statistically significant at the 10% level in the third- and fourth-order specifications. Predicted earned incomes would hence have been approximately 2.5% greater than the observed earned incomes for married men without the EITC in 2010. On average, the 2.5% decrease corresponds to a 0.5 SEK decrease of gross earned income for each SEK of credit, which is a sizeable decrease. Because more than 60% of gross earned income consists of taxes (including indirect taxes), for each SEK of credit that the government provides, it loses additionally 0.3 SEK in revenue because of behavioral effects. However, from the individual’s perspective, net earned income increases by almost 80% of the value of the credit. Only approximately 20% of the credit are hence counteracted by behavioral effects and “consumed” as leisure.

For married women and singles, the estimated reform effects are statistically insignificant. The standard errors for married women are of the order that we can rule out effects of greater than 3.2% (at the 5% level). For singles, the precision is very low: we can only rule out effects of greater than 8.4%. The statistically insignificant estimates are therefore not very informative.

Turning to the participation effects, we see that they are all positive. For married women and men, the estimated reform effects of the probability of having positive income are positive and statistically significant at the 1% level. The effect is approximately 1.7 percentage points for married men and 2.9 percentage points for married women. For singles, the estimated effect is statistically insignificant. However, the standard errors are again much greater than for couples, but we can only rule out effects of greater than 2.3 percentage points.

However, half of the participation effect for married men disappears when setting the participation threshold to 100,000 SEK (the reform effect estimates are approximately 0.8 percentage points). The estimates are now only statistically significant at the 10% level for 3 out of 4 estimates. For married women, the effects are almost gone (approximately 0.5 percentage points) and statistically insignificant. Most of the participation reform effects are therefore at low incomes. This result is not surprising because the largest marginal tax effects are at low incomes (see Figure 3). For singles, the estimated effect is again statistically insignificant. Although it is greater than for the effect on the probability of having positive income, the standard errors are even greater and we can only rule out effects of greater than 3.4 percentage points.

Taken together, the three sets of estimates suggest that the reform had positive participation effects for couples at low incomes but negative labor supply effects for married working men. However, the precision is too low to provide any reliable reform effect estimate for the earned income of singles. The theoretical predictions on the direction of the effects are ambiguous for couples. The results suggest that substitution effects dominate for married non-working individuals, whereas income effects dominate for married working men.

I explore the reform effects in the reform period year by year in Table 5. The reform year is varied horizontally. Otherwise, the organization is the same as in Table 4. I report the estimates from the fourth-order specification in Table 4. The patterns across reform years are similar. For all reform years, I find statistically significant positive participation effects at zero income for couples and at 100,000 SEK for married men and negative overall effects for married men. The estimated reform effects typically increase in magnitude over time between the reform years as the program expanded.

There is one other study that provides a reform effect estimate of the Swedish EITC for the broad groups analyzed here: Ericson et al. (2009). They simulate the effects of the 2009 program using discrete-choice methods and conduct the estimation on a sample of individuals from 2006. They find a positive effect on hours of work of 1.5% for the three groups as a whole. No subgroup analysis or standard errors are provided. They also find a positive effect on the probability of having positive hours of work of 1.1%.

Table 5. Estimates of reform effects year by year

		2007	2008	2009	2010
y	Married men	-1.492*	-2.277**	-1.503**	-2.498*
		(0.817)	(1.065)	(1.236)	(1.447)
	Married women	1.210	1.198	2.285	2.137
		(0.958)	(1.210)	(1.436)	(1.631)
	Singles	1.293	1.510	2.793	3.424
		(2.443)	(3.001)	(3.788)	(4.229)
$D(y > 0)$	Married men	0.949***	1.132***	1.452***	1.687***
		(0.188)	(0.242)	(0.280)	(0.328)
	Married women	1.646***	1.934***	2.660***	2.907***
		(0.223)	(0.285)	(0.340)	(0.389)
	Singles	0.813	0.948	1.389	1.620
		(0.666)	(0.810)	(1.007)	(1.146)
$D(y > 100)$	Married men	0.523*	0.678*	0.800*	0.910*
		(0.284)	(0.366)	(0.424)	(0.497)
	Married women	0.319	0.345	0.661	0.588
		(0.348)	(0.443)	(0.529)	(0.605)
	Singles	1.398	1.697	2.254	2.630
		(0.987)	(1.201)	(1.493)	(1.699)

Notes: Each row of estimates contains estimates from a separate regression. For earned income, estimated reform effects are in % of the observed earned incomes. For participation, estimated reform effects are in percentage points. Standard errors are constructed using the delta method. * significant at 10%, ** significant at 5%, * significant at 10%.

I have also experimented with additional specifications and sample selection criteria. These include dropping some demographic control variables, using other functional forms for the demographic variables, dropping one year of data at a time (such as 2008, when the financial crisis occurred), including observations with more extreme earned or unearned incomes, and including observations with greater sick benefits, parental benefits, income from self-employment, and student financial aid. These exercises produce similar results but could increase the standard errors substantially.

6.2 Placebo estimates

I now conduct placebo tests as described in Subsection 3.3. I do this for the reform effect estimates of each of the reform years. As described, I first predict a one-dimensional synthetic reform intensity variable and place this in each of the pre-reform years. The point estimates are presented in Figure 7. The top, middle, and bottom rows of figures show the estimates for earned income, the probability of having positive income, and the probability of having income greater than 100,000 SEK, respectively. The left, middle, and right columns of the figures show the estimates for married men, married women, and singles, respectively. Different symbols are used for the synthetic reform effect estimates of different reform years. Each post-reform year contains one synthetic reform effect estimate. Each pre-reform year contains four placebo reform effect estimates, one for each post-reform year reform effect estimate.

First, we observe that the reform effect estimates for the reform years using the one-dimensional synthetic reform variable are all in the same direction and similar in magnitude to the estimates using the full reform effect vectors. This result reassures us that the synthetic reform variables do capture some of the reform effects.

For the earned income of married men, we see that all the placebo estimates are closer to zero and much less than any of the reform year estimates. They also have mixed signs. Furthermore, none of the placebo estimates are statistically significant (at the 10% level), whereas all the reform year estimates are negatively statistically significant (at the 10% level for 2007 and at the 1% level for 2008-2010).

For the probability of having income above zero, the results are similar, but with positive reform year effects. The placebo estimates are closer to zero than the reform year estimates except for 2007 and have mixed signs. The reform year estimates are statistically significant (at the 1% level) for 2008-2010 (but not for 2007), whereas none of the placebo estimates are statistically significant.

For the probability of having income greater than 100,000 SEK, the reform-effect estimates were small and only significant at the 10% level. The placebo results are also less clear. The placebo estimates in 2006 is of the same magnitude as in the reform years. 3 out of 4 reform year estimates are statistically significant (at the 10% level in 2007, at the 5% level in 2009, and at the 1% level in 2010), but all the placebo estimates in 2006 are also statistically significant (at the 5% or 10% levels), as is 1 out of the 12 other placebo estimates.

For the probability of having positive income for married women, all placebo estimates are less than the positive reform year estimates, although the placebo estimates in 2003 are of similar magnitude as the reform year estimates in 2007 and 2008 but with the opposite sign. All reform year estimates are statistically significant (at the 5% level for 2007 and at the 1% level for 2008-2010). However, 7 out of 16 placebo estimates are also statistically significant at the 10% level, out of which 4 are from 2003.

Some significant placebo estimates are expected when making so many tests. However, for the probability of having income greater 100,000 SEK for married men in 2006, and for the probability of having positive income for married women in 2003, it seems that individuals who received different reform intensities during the reform years may have differed from each other in ways with respect to the outcome variable that I have not fully accounted for. However, dropping these problematic years in these two cases in the estimation does not change the overall results.

The patterns for earned income and the probability of having income greater than 100,000 SEK for married women and all three labor supply measures for singles are different. The reform year estimates are mostly statistically insignificant and have different signs between reform years. The placebo estimates sometimes have the same magnitude and sign as the reform year estimates. They are also sometimes statistically significant.

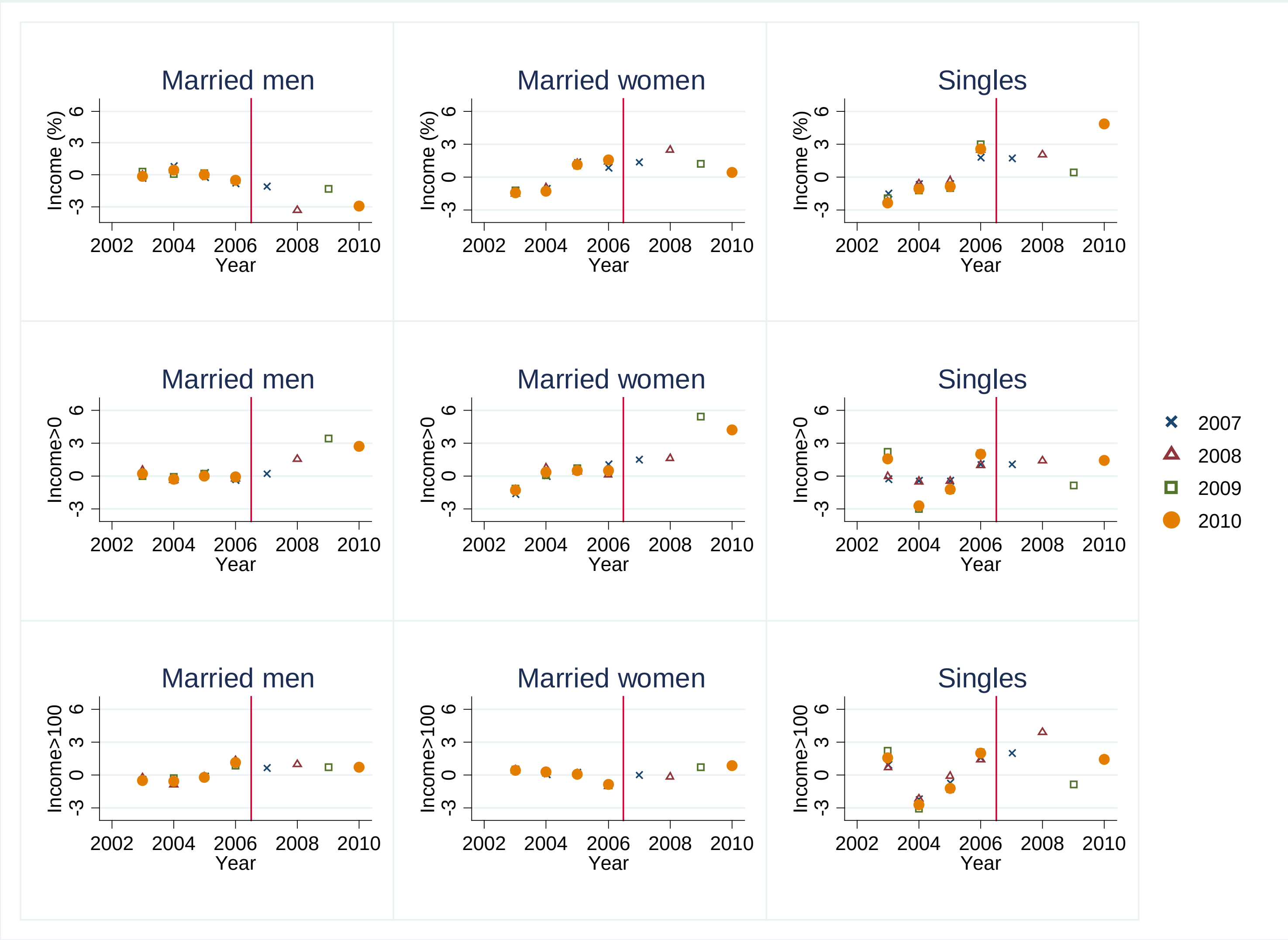

Figure 7. Placebo estimates of reform effects

Note: Participation in the six bottom graphs is in percentage points.

7. Conclusions

I have developed a reform evaluation method for behavior (in particular, labor supply) in nonlinear budget sets that is both structural and quasi-experimental. The model only requires preferences to be convex on the budget frontier, which can accommodate utility maximization with convex preferences on convex budget sets and some types of nonconvexities. The resulting simple three-dimensional labor supply function represents an intuitive and parsimonious extension of the standard two-dimensional labor supply function for linear budget sets.

The labor supply function is flexibly approximated using polynomials. These terms are transparent and include, e.g., the area under the budget frontier. I then developed a differencing method that only exploits between-individual variation in reform effects. This method was complemented with a method to conduct placebo tests when several pre-reform years are available.

In the empirical application, I applied the method to evaluate the earned income effects of the Swedish earned income tax credit introduced between 2007 and 2010. This reform was difficult to evaluate using quasi-experimental methods because the program is universal. However, the formula interacts with the rest of the tax and transfer system, thereby providing complicated variation in reform effects on budget sets between individuals. This variation could be harnessed and exploited by the developed method, which illustrates its usefulness.

I found that the reform had positive participation effects for couples. The effect on having positive earned income is 1.7 percentage points for married men and 2.9 percentage points for married women. However, half of the “newly supplied” jobs for married men and most of the jobs for married women provide incomes less than 100,000 SEK. I also found negative earned income effects for married *working* men that lead to overall negative earned income effects of 2.5% for married men. The reform effects on earned incomes for married women and the reform effects for singles were not statistically significant but were imprecisely estimated.

For married men, the results imply an additional 30% revenue loss from them for the government beyond the mechanical revenue loss of the credits. The program’s lack of a phase-out region is, hence, costly not only because it provides tax cuts that are non-decreasing in income but also because it has adverse intensive margin income effects. However, the estimated effects are short-run effects. The long-run effects may be different if adjustment takes time and because of general equilibrium effects.

References

Aaberge, R., Flood, L. (2008). Evaluation of an in-work tax credit reform in Sweden: effects on labor supply and welfare participation of single mothers, IZA Discussion Paper 3736.

Blomquist, S., Hansson-Brusewitz, U. (1990). The effect of taxes on male and female labor supply in Sweden. Journal of Human Resources 25, 317-357.

Blomquist, S., Eklöf, M., Newey, W. (2001). Tax reform evaluation using non-parametric methods: Sweden 1980-1991. Journal of Public Economics 79, 543-568.

Blomquist, S., Newey, W. (2002). Nonparametric estimation with nonlinear budget sets. Econometrica 70, 2455-2480.

Blomquist, S., Selin, H. (2010). Hourly wage rate and taxable labor income responsiveness to changes in marginal tax rates. Journal of Public Economics 94, 878-889.

Blomquist, S., Kumar, A., Liang, C., Newey, W. (2012). Nonparametric estimation of taxable income function. Manuscript.

Blundell, R., Duncan, A., Meghir, C. (1998). Estimating labor supply responses using tax reforms. Econometrica 66, 827-861.

Blundell, R., Brewer, M., Shephard, A. (2005). Evaluating the labour market impact of working families' tax credit using difference-in-differences. Manuscript.

Blundell, R., Shephard, A. (2012). Employment, hours of work and the optimal design of earned income tax credits. Review of Economic Studies 72, 481-510.

Burtless, G., Hausman, J. (1978). The effect of taxation on labor supply: evaluating the Gary negative income tax experiment. Journal of Political Economy 86, 1103-1130.

Chetty, R., Saez, E. (2013). Teaching the tax code: earnings responses to an experiment with EITC recipients. American Economic Journal: Applied Economics 5, 1-31.

Chetty, R., Friedman, J., Saez, E. (2013). Using differences in knowledge across neighborhoods to uncover the impacts of the EITC on earnings. Forthcoming in American Economic Review.

Edin, P-A., Fredriksson, P. (2000). LINDA – Longitudinal INdividual DAta for Sweden. Working paper 2000:19, Uppsala University.

Edmark, K., Mörk, E., Liang, C-Y., Selin, H. (2012). Evaluation of the Swedish earned income tax credit. Working paper 2012:3, Uppsala University.

Eissa, N. (1995). Taxation and labor supply of married women: the tax reform act of 1986 as a natural experiment. National Bureau of Economic Research Working Paper 5023.

Eissa, N., Hoynes, H. (2004). Taxes and the labor market participation of married couples: the earned income tax credit. Journal of Public Economics 88, 1931-58.

Eissa, N., Liebman, J. (1996). Labor supply response to the earned income tax credit. Quarterly Journal of Economics 111, 605-37.

Ericson, P., Wahlberg, R., Flood, L. (2009). SWEtaxben: a Swedish tax/benefit micro simulation model and an evaluation of a Swedish tax reform, IZA Discussion Paper 4106.

Feldstein, M. (1995). The effect of marginal tax rates on taxable income: a panel study of the 1986 tax reform act. Journal of Political Economy 103, 551-572.

Feldstein, M. (1999). Tax avoidance and the deadweight loss of the income tax. Review of Economics and Statistics 81, 674-80.

Flood, L., Wahlberg, R, Pylkkänen, E. (2007). From welfare to work: evaluating a tax and benefit reform targeted at single mothers in Sweden. Labour 21, 443-471.

Gruber, J., Saez, E. (2002). The elasticity of taxable income: evidence and implications. Journal of Public Economics 84, 1-32.

Hansson, Å. (2007). Taxpayers responsiveness to tax rate changes and implications for the cost of taxation. International Tax and Public Finance 14, 563-82.

Hausman, J. (1985). The econometrics of nonlinear budget sets. Econometrica 53, 1255-1282.

Hoynes, H. (1996). Welfare transfers in two-parent families: labor supply and welfare participation under AFDC-UP. Econometrica 64, 295-332.

Keane, M., and Mofitt, R. (1998). A structural model of multiple welfare program participation and labor supply. International Economic Review 39, 553-589.

Kumar, A. (2012). Nonparametric estimation of the impact of taxes on female labor supply. Journal of Applied Econometrics 27, 415-439.

Liang, C-Y. (2012). Nonparametric structural estimation of labor supply in the presence of censoring. Journal of Public Economics 96, 89-103.

Newey, W. (1997). Convergence rates and asymptotic normality for series estimators. Journal of Econometrics 79, 147-168.

Meyer, B., Rosenbaum, D. (2001). Welfare, the earned income tax credit, and the labour supply of single mothers, Quarterly Journal of Economics 116, 1063-1114.

Porter, J. (1998). Nonparametric estimation with panel data. Working Paper, Department of Economics, Harvard.

Saez, E. (2010). Do taxpayers bunch at kink points?. American Economic Journal: Economic Policy 2, 180-212.

Stone, C. (1985). Additive regression and other nonparametric models. Annals of Statistics 13, 689-705.

Van Soest, A. (1995). Structural models of family labor supply: a discrete choice approach. Journal of Human Resources 30, 63-88.

Van Soest, A., Das, M., Gong, X. (2002). A structural labour supply model with flexible preferences. Journal of Econometrics 107, 345-374.

Weber, C. (2013). Toward obtaining a consistent estimate of the elasticity of taxable income using difference-in-differences. Manuscript.