UPPSALA
UNIVERSITET

Uppsala Center for Fiscal Studies
Department of Economics

Working Paper 2014:4

Distribution-Free Structural Estimation with Nonlinear Budget Sets

Che-Yuan Liang

Uppsala Center for Fiscal Studies
Department of Economics
P.O. Box 513
SE-751 20 Uppsala
Sweden
Fax: +46 18 471 14 78

Working paper 2014:4
April 2014

Distribution-Free Structural Estimation with Nonlinear Budget Sets

Che-Yuan Liang

Distribution-Free Structural Estimation with Nonlinear Budget Sets

Che-Yuan Liang[#]
April 14, 2014

Abstract: I develop a structural method for evaluating labor supply in nonlinear budget sets that does not require any distributional assumptions. The model only requires that preferences are convex on the budget frontier. It can be extended to account for features such as fixed costs of work and the stigma cost of welfare participation. It can also be adapted for estimation of earnings, hours of work, and functions that depend on the labor supply distribution, including tax revenue and cumulative distribution functions. The method is applied to estimate the effects of taxes on various labor supply outcomes in the U.S. and Sweden.

Keywords: nonlinear budget sets, structural models, distribution-free estimation, labor supply

JEL classification: D04, H24, J22

[#] Uppsala Center for Fiscal Studies, Department of Economics, Uppsala University, P.O. Box 513, SE-75120, Uppsala, Sweden; Telephone: +46737780051; E-mail: che-yuan.liang@nek.uu.se.

1. Introduction

Optimal design of tax systems requires knowledge about the behavioral effects of taxes. There are, however, many challenges in estimating the effects of taxes on, e.g., labor supply in the form of earnings or hours of work. A major difficulty is that real-world tax systems create nonlinear budget sets. The most popular approach to addressing this difficulty is linearizing the budget set around the observed labor supply outcome[1] and estimating the effect of this budget set characterized by its slope and intercept.[2] However, the slope and intercept are determined by the marginal tax rate, which is endogenous to the labor supply outcome. Furthermore, optimization errors in labor supply may cause the observed slope to differ from the slope the individual acted upon. In addition to presenting estimation difficulties, these issues also make it difficult to simulate the labor supply effects of a tax reform.

Structural approaches can account for many of the nonlinear budget set issues. The popular static discrete choice method treats the decision problem as a multiple choice problem and estimation is conducted using maximum likelihood methods.[3] It is possible to extend the standard model to account for features such as fixed costs of work, labor demand quantity restrictions, and the stigma cost of welfare participation. However, a number of distributional and functional form assumptions are typically needed to account for complications such as observed and unobserved preference heterogeneity, and optimization and measurement errors.[4] The potential bias arising from imposing these types of assumptions (if they do not hold) or from not accounting for these complications has rarely been investigated. The paper by Van Soest et al. (2002) is an exception: they show that the standard approach of not accounting for an additive optimization or measurement error in hours of work may induce a significant bias.[5]

In this paper, I develop a structural method that accounts for nonlinear budget sets and complications such as preference heterogeneity, and optimization and measurement errors, without requiring any distributional assumptions. As in the standard discrete choice method, I discretize the labor supply options. However, instead of estimating the utility function, I derive the labor supply function. This function can then be flexibly approximated using a polynomial in its arguments, and it can be estimated with least squares. Once having the budget sets, estimation is simple to implement unlike the standard method.

A labor supply option can be characterized by its labor supply level and the amount of net income this level provides which depends on the tax and transfer system. Each option is therefore two-dimensional. The labor supply function depends in general on the entire budget set and has a dimensionality of two times the number of options. This function would not be

---

[1] The linearized and the original budget sets yield the same optimum if data are generated by utility maximization with globally convex preferences and there are no optimization errors, as noted by Diewert (1971).

[2] The slope is the marginal net-of-tax rate when earnings are the outcome and the marginal net wage when hours of work are the outcome. In a linear (not linearized) budget set, the intercept would be unearned income.

[3] Earlier papers that developed this method include Dagsvik (1994), Van Soest (1995), Hoynes (1996), Keane and Mofitt (1998), and Van Soest et al. (2002). Numerous other papers apply the method including Aaberge et al. (1995), Aaberg et al. (1999), and Blundell and Shephard (2012).

[4] Although many of these assumptions can be relaxed, doing so typically complicates the already complicated estimation considerably.

[5] This is even though the standard approach allows for a Type I extreme value distributed error term in the utility function that often is interpreted as an optimization error.

feasible to estimate if no additional assumptions were made about preferences or the budget set. I show that imposing a weak joint restriction on preferences and the budget set reduces the dimensionality of the labor supply function. The assumption is that preferences on the budget frontier are convex in labor supply, implying that there is a uniquely preferred, but not necessarily practically available or chosen, option for every individual. This assumption allows for but does not limit behavior to utility maximization with convex preferences in convex budget sets.[6]

The resulting labor supply function is a sum of three-dimensional functions over the labor supply options where each option depends on the net income level, the net income derivative (with respect to labor supply), and the labor supply level on the budget frontier. Because the summation is over functions with the same structure, the dimensionality of the sum is also three, and it is feasible to approximate and estimate the sum using polynomials imposing only the equality restrictions implied by the model. Precision and computational requirements are therefore independent of the number of options used to approximate the budget set, in contrast with the standard discrete choice method. The approximating terms are transparent and have a geometrical interpretation: they represent functions of distances, slopes, and areas under the budget frontier. The labor supply function is an intuitive and parsimonious extension of a labor supply function on a linear budget set that would be two-dimensional and depend on the net income level and derivative (at some fixed labor supply level) on the budget frontier.

It is possible to extend the model by specifying a more complicated utility function, as in the standard discrete choice method, to account for features such as fixed costs of work, labor demand quantity restrictions, and the stigma cost of welfare participation.

The drawback of not making any distributional assumptions is that the distribution of error terms cannot be estimated. Consequently, the expected labor supply function does not provide any knowledge about the labor supply *distribution*. I show, however, that it is straightforward within my framework to derive and estimate a tax revenue function, and labor supply cumulative distribution and probability density functions (including a participation probability function), to assess the distributional effects of taxes.

My method is applicable for both earnings and hours of work, unlike the standard discrete choice method, which can only be applied to hours of work. Earnings are a broader measure of labor supply because earnings include the effort dimension of work, which was demonstrated to be potentially important by Feldstein (1995, 1999).[7] Focusing on the effects on earnings avoids the need for data on hours of work and gross wage rates, which often are not available when administrative data sources are used. I show how my framework can be

[6] The assumption allows for nonconvexities in the budget set that are less nonconvex than the preferences are convex.

[7] The taxable income literature typically uses the linearization approach. Differencing techniques are used to account for any unobserved heterogeneity correlated with the budget set, an issue that is largely neglected by structural methods. The endogeneity of the net-of-tax-rate issue is addressed using instrumental variables. See Gruber and Saez (2002) for the standard method and Saez et al. (2012) for an overview. The typical instrument is based on functions of lagged income, which may have its own effect on current income and therefore must be controlled for. Identification therefore rests on functional form assumptions, and Kopczuk (2005) and Giertz (2008) show that results are very sensitive to the functional form chosen. Weber (2014) also reports sensitivity to which set of lags to use for constructing the instruments. In the hours of work literature, Blomquist (1996) reports similar sensitivity to instrument choice in the linearization approach.

adapted for investigating hours of work without requiring access to a gross wage rate variable. This is advantageous because it circumvents issues related to cases when the gross wage rate cannot be observed for everyone (it is often not observed for nonworkers) or is imputed through dividing earnings with hours of work, which could potentially cause a serious division bias in the estimation (Eklöf and Sacklén, 1999).

The approach that I take of focusing on the labor supply functions rather than the utility function is similar to two other structural methods for estimating hours of work, the Hausman method (Burtless and Hausman, 1978; Hausman, 1985) and Blomquist and Newey's (2002) method. A major difference is that both of these methods begin from piecewise-linear convex budget sets and convex preferences. The Hausman method is similar to the standard method in making a number of distributional assumptions and employing maximum likelihood estimation. Blomquist and Newey show that the distributional assumptions in the Hausman method could bias the estimation.

The developed method shares a number of similarities with Blomquist and Newey's (2002) method, including not requiring any distributional assumptions and estimating labor supply functions that are functions of the entire budget set. The flexible polynomial approximation of the labor supply functions is another common feature.[8] My labor supply function is, however, more intuitive and parsimonious.[9] It also makes better use of the variation between budget sets by pooling the variation across different parts of the budget sets, resulting in less sensitivity to the functional form and outliers and better precision.[10] In addition to Blomquist and Newey's method also being more restrictive on the shapes of budget sets and preferences, extending their model to account for fixed costs of work, labor demand quantity restrictions, and the stigma cost of welfare participation does not seem straightforward.[11]

I include two empirical applications in the paper: estimation of labor supply effects in the U.S. and Sweden, using data from 2006, from the Panel Study of Income Dynamics (PSID) for the U.S. and Hushallens Ekonomi (HEK) for Sweden. I investigate the effects for households that filed jointly in the U.S. and for married or cohabiting men in Sweden. I report slope elasticities that correspond to uniform increases in tax rates across the budget set. For the U.S., I find an earnings slope elasticity of 0.27, an hours of work slope elasticity of 0.01, and a tax revenue slope elasticity of -2.43. Except for the tax revenue elasticity, precision is low. For Sweden, I find an earnings slope elasticity of 0.42 and a tax revenue slope elasticity of -0.17. Precision for the tax revenue elasticity is low. I also estimate the effects on the distribution of earnings for Sweden. The results show that tax decreases primarily decreased

[8] Van Soest et al. (2002) show that the utility function in the standard discrete choice method can also be flexibly approximated and estimated.

[9] For instance, my functions each contain 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.

[10] These advantages could be important for the implementability in many applications in which either the data set or the variation in budget sets is small, e.g., when only a single cross section is available. Applications of Blomquist and Newey's method thus far have always used several cross sections with different tax regimes to obtain more variation in budget sets (e.g., Kumar, 2007; Liang, 2012).

[11] Although their method was developed for analyzing expected hours of work, it is possible to extend their method for analyzing earnings, and it is likely also possible to extend it for analyzing functions that depend on the labor supply distribution. See Blomquist et al. (2014) for developments in these directions.

the share of individuals with one-eighth to one-half the average earnings and increased the share of individuals with average or above average earnings. All estimated income elasticities are close to zero for both the U.S. and Sweden.

The paper proceeds as follows: In the next section, I describe the basic framework. Section 3 presents some extensions of the model. Section 4 shows how to estimate the model. Sections 5 and 6 report results from the empirical applications. The final section concludes.

2. Basic framework

2.1 Basic setup

The individual choice problem at hand is two-dimensional. The individual chooses (c, y) subject to a constraint $c(y)$. There are J y-options indexed by j, and the options are spaced with a distance δ between each. Hence, $y_j = j\delta$, $j \in \{0,1,\dots,J\}$. Because of the constraint, choosing y_j uniquely determines $c_j = c(y_j)$. Letting $\delta \to 0$ and $J \to \infty$, we obtain the case that y is a continuous choice. The constraint function $c(y)$ can be described by $B = (y_0, c_0, \dots, y_J, c_J)$ containing $2J$ variables.[12] The setup allows but does not require $c(y)$ to be continuous.

In the static labor supply application, c is consumption or net income and y is a labor supply measure such as earnings or hours of work. $c(y)$ is then the budget constraint and B the budget set, and both describe the consumption possibilities on the budget frontier. The problem is complicated by the fact that $c(y)$ could be nonlinear. Real tax and transfer systems often, but not always, produce piecewise-linear continuous convex budget sets.

The individual maximizes the utility function

$$U = U(c, y, e), \tag{1}$$

where e is an error term representing individual heterogeneity in preferences for labor supply and consumption. The utility function is assumed to be monotonic, complete, continuous, and increasing in c and decreasing in y. I work with a one-dimensional e, but it is straightforward to allow e to be multi-dimensional. Additional error-term dimensions could, e.g., be used to represent unobserved job characteristics, as in the standard discrete choice method.[13] It is possible to let individual heterogeneity consist of two terms, observed and unobserved heterogeneity, z and e, and work with the utility function $U(c, y, z, e)$; the derivations in this paper would still be possible.

The individual's decision problem is now

$$\max_{c,y} U(c, y, e) \quad \text{s.t.} \quad c = c(y) \Leftrightarrow \tag{2}$$

$$\max_{y} u(y, e) = U(c(y), y, e). \tag{3}$$

[12] These variables vary between individuals but they are parameters from the point of view of the function $c(.)$ for a given individual.

[13] In this method, $U = U_1(c, y, e_{-1}) + \sum_j 1(y_j = j)\, e_j$ where e_{-1} represents individual heterogeneity and e_j represents unobserved job characteristics attached to labor supply option j. Whereas e_{-1} is typically assumed to be normally distributed, each e_j is assumed to be Type I extreme value distributed.

The optimization problem results in the desired labor supply choice

$$y^d(B,e) = y_j \quad \text{if } \mu_j(B,e) \Leftrightarrow u_j > u_k \text{ for all } k \neq j, \tag{4}$$

$$u_j = u\big(y_j, e\big) = U\big(c_j, y_j, e\big). \tag{5}$$

u_j is the utility function on the budget frontier.

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

$$y = y(y^d, \varepsilon). \tag{6}$$

ε can capture that individuals may not be able to fine-tune their labor supply choices, e.g., because of job availability issues, and often have to choose among a limited set of labor supply options.[14] A priori, this component can be large and important for many individuals.[15] This type of error term tacked on to the labor supply outcome follows the approach taken by the Hausman (1985) method.[16] Van Soest et al. (2002) remark that a suitably specified error term in the utility function (see footnote 13) can serve the same purpose. However, in their simulation exercise, they show that estimation is biased when there is an additive optimization error term in desired labor supply, despite their allowing for an error term in the utility function that intended to capture this error.[17] I allow ε to be multidimensional and contain, e.g., a measurement error in addition to the optimization error.

I assume that the error terms are independent of the budget set in Assumption 1.

Assumption 1. e and ε are independent of B.

When allowing for observed heterogeneity z, we only require independence conditional on z.

Letting $F(e, \varepsilon)$ be the distribution function,

$$E(y) = Y(B) = \sum\nolimits_j \iint y\big(y_j, \varepsilon\big) dF\left(e, \varepsilon | \mu_j(B,e)\right). \tag{7}$$

Structural methods focus on estimating the parameters of $Y(.)$ or $U(.)$. The standard way to proceed is to make functional form assumptions on $U(.)$ and $y(.)$, and distributional assumptions on $F(e, \varepsilon)$, and to use maximum likelihood estimation, in which case the parameters of $F(e, \varepsilon)$ can also be estimated. In the standard discrete choice method (e.g., Dagsvik, 1994; Van Soest, 1995; Hoynes, 1996; Keane and Mofitt, 1996; Van Soest et al., 2002), $c(y)$ is allowed to have any functional form. In the Hausman method (e.g., Burtless and Hausman, 1979; Hausman, 1985), $c(y)$ is restricted to be piecewise linear and weakly concave.

---

[14] If we know the exact options for the individuals, these restrictions can be addressed by appropriately modifying the budget set B.

[15] Structural models for piecewise-linear budget sets typically produce a disproportionate number of individuals who want to locate at kink points. Typically little or no bunching is observed in practice (Saez, 2010), which implies either small behavioral responses or large optimization errors.

[16] Hausman does this additively.

[17] It is possible that this bias arises because of the restrictive distributional assumption (type I extreme value distribution) on the optimization error in the utility function that is essential for implementability in the standard discrete choice method.

Another way to proceed is to not make any distributional assumptions and to directly and only estimate parameters of $Y(.)$, as in Blomquist and Newey (2002). This can be done by approximating $Y(B)$ by functions of B. The approximation can be done flexibly by using polynomials, and estimation can be carried out with least squares. The fundamental difficulty is that B is $2J$-dimensional. Accounting for the entire budget set leads to a curse-of-dimensionality problem, and estimation is not feasible when the number of options grows. Blomquist and Newey restrict the budget set to be piecewise linear and convex as in the Hausman (1985) method, and they impose a number of restrictions that are implied by utility maximization with convex preferences. Their procedure reduces $Y(B)$ to two terms, one having a dimensionality of two and the other a dimensionality of three, respectively, which is feasible to approximate and estimate.

The approach taken here is similar, but it does not restrict the budget set to be piecewise linear and convex. Note that the summation over labor supply options sums over integrals with the same functional form because the optimization problem at each option has the same structure. Summing over options imposing the equality restrictions therefore increases neither the dimensionality nor the number of terms to be approximated. The fundamental problem, however, is that the optimality of each option described by the condition $\mu_j(B, e)$ depends on the entire budget set. I reduce the dimensionality by imposing one restriction on this condition.

2.2 Key assumption

To reduce the dimensionality of the expected labor supply function in Eq. (7), I only make one joint assumption on preferences and the budget set, which is described in Assumption 2.

Assumption 2. Preferences on the budget frontier are convex in labor supply:

$$u_j > u_k \text{ and/or } u_j \geq u_l \text{ for all } y_k < y_j < y_l \qquad (8)$$ [18]

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) option 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.[19] 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. All proofs are presented in Appendix A1.

[18] For simplicity, I assume strict convexity from below. This implies that an individual chooses the smaller of two labor supply options if the options have the same utility. It would be straightforward to assume strict convexity from above instead or to allow weak convexity from above and below, and/or to have individuals randomizing between adjacent options with the same utility without any implications for estimation.

[19] 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 function of the problem in Eq. (1) being quasi-concave.

Lemma 1. Assumption 2 implies that an option is locally optimal if it is preferred to its two adjacent options. For interior labor supply options,

$$y^d = y_j \quad \text{if and only if } \mu_j'(c_j, c_{j-1}, c_{j+1}, y_j, e) \Leftrightarrow u_j > u_{j-1} \text{ and } u_j \geq u_{j+1}. \tag{9}$$

For the corner solutions y_0 and y_J, we only require $u_0 \geq u_1$ and $u_J > u_{J-1}$ respectively.

Lemma 2. Assumption 2 implies that no interior option can be less preferred than both of its two adjacent options:

$$(c_j, y_j): u_j \leq u_{j-1} \text{ and } u_j < u_{j+1} \text{ is empty for all } y_j, \tag{10}$$

Lemma 1 guarantees that there is only one local maximum that is also a global maximum. We can therefore check whether a labor supply option is the preferred one by checking whether it is locally optimal, which we do by comparing it with its neighboring options and ignoring all other options. Lemma 2 rules out local minima. Assumption 2 and the implied Lemmas 1 and 2 are restrictions on the convexity of the budget set in relation to the convexity of preferences. In Corollary 1, I further prove that Assumption 2 allows, but does not limit behavior to, utility maximization with convex preferences in convex budget sets.

Corollary 1. Assumption 2 allows for utility maximization with convex preferences in convex budget sets.

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, in which utility is increasing in the nonconvex region. It also does not specifically rule out cases in which individuals may want to locate in nonconvex regions. One such case is illustrated in the top-right figure in Figure 1. There, we see that we require that preferences are more convex than the budget set is nonconvex. 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.[20] This implies that there is a utility minimum in this area, which is not allowed by Lemma 2. In the bottom-right figure, this nonconvexity creates two globally optimal labor supply options.

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

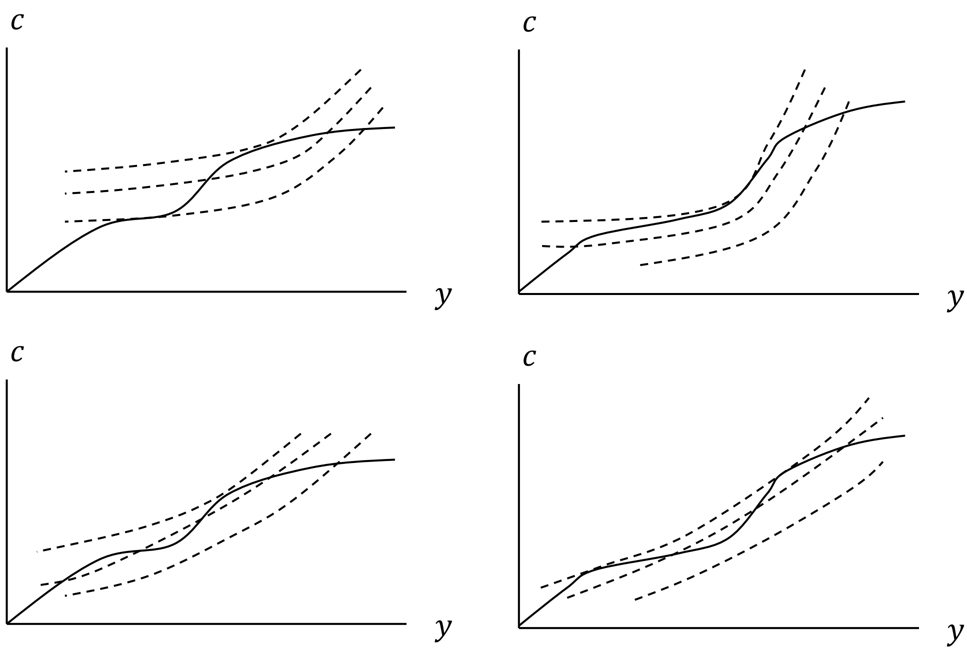

Figure 1. Nonconvexities allowed (top figures) and not allowed (bottom figures)

2.3 Expected labor supply function

Lemma 1 reduces the global optimization problem of which labor supply option to choose into many local problems of whether to choose a certain option. Each local problem only depends on a comparison between neighboring options and a condition $\mu'(.)$ that depends on the four budget set variables: c_j, c_{j-1}, c_{j+1}, and y_j. Straightforward summation over the interior options imposing the equality restrictions therefore gives a four-dimensional function. In Theorem 1, I show that exploiting the repetitive structure implied by Lemma 1 and a decomposition implied by Lemma 2 reduces the dimensionality to three and accounts for the end points at the same time.

Theorem 1. Given Assumptions 1 and 2 and the utility function in Eq. (1), expected labor supply is (by the law of total expectation)

$$
\begin{aligned}
&E(y) = \sum_{j=0}^{J-1} \pi\left(c_j, c_j', y_j\right), \\
&\pi\left(c_j, c_j', y_j\right) = \iint \left[y\left(y_j + \delta, \varepsilon\right) - y\left(y_j, \varepsilon\right)\right] dF\left(e, \varepsilon | u_{j+1} > u_j\right), \\
&u_{j+1} = u\left(c_{j+1} = c_j + c_j', y_j + \delta\right), u_j = u\left(c_j, y_j\right).
\end{aligned}
\tag{11}
$$

The labor supply expression in Eq. (11) has an intuitive interpretation. It is a sum over terms at all labor supply options, where the term at each option depends only on $\left(c_j, c_j', y_j\right)$, which are the net income level, the net income derivative, and the labor supply level on the budget frontier. The dimensionality is independent of the number of options used to approximate the budget frontier. When allowing for observed heterogeneity z, the labor supply function increases in dimensionality by the dimension of z. It is possible to restrict observed heterogeneity to entering through a single index γz, in which case the dimensionality would only increase by one. γ can be estimated using the procedure described by Ichimura (1993).

Consider the case of a linear (or linearized) budget set. The budget set B could then 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 earnings. $Y(B) = Y(c_0, c')$ would then be two-dimensional. It is straightforward to see how the derived labor supply function nests this special case as $c_j = c_j(c_0, c') = c_0 + c' y_j$ and $\pi(c_j, c_j', y_j) = \pi(c_0, c', y_j)$ in this case. A polynomial approximation results in $E(y) = \sum \pi(c_0, c', y_j) = m(c_0, c') \sum n(y_j) = Y(c_0, c')$. From the point of view of the full model, we see that the linearized budget set model of the nonlinear budget set model results in a misspecification because the constant consumption level c_0 and slope c' are entered instead of the varying consumption levels c_j and slopes c_j'. The nonlinear budget set extension is made straightforward by involving the net income level and derivative not only at one point but at all points. Furthermore, each point may contribute differently, which can be described by the interaction effects with the labor supply level of that point. This feature increases the dimensionality by one.

The derivation of the expected labor supply function here is related to but simpler than the one presented by Blomquist and Newey (2002). In addition to being less restrictive on preferences and the budget set, the derived 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 frontier. Because budget sets have different numbers of kinks and segments of different lengths that are located at different labor supply levels, the interpretation of the sum term is less intuitive in their function. Their expression also contains another two-dimensional term that depends on characteristics of the last segment.[21] The reduction in the number of terms here and the resulting "symmetrical" (in each labor supply option) expression is attributable to the imposition of more equality restrictions that is implied by utility maximization.

Although convex budget-constrained preferences are assumed, the structure of the labor supply expression can accommodate more general data-generating processes. The essential ingredient is that the contribution of a labor supply option in a budget set to the labor supply function only depends on the level and derivative characteristics of that option.[22] The convexity assumption, however, must hold in order for the labor supply estimates to be coherent with the behavioral utility maximization model. I do not, however, empirically impose the assumption by constraining the parameters in the empirical specification.

---

[21] Their labor supply expression, for a piecewise-linear budget set with J segments indexed by j, is characterized by (θ_j, r_j, k_j), $j = 1, \ldots, J$, where θ is the slope, y is the intercept with the consumption axis, and k is the kink point, and it has the form $E(Y) = \bar{Y}(\theta_J, r_J) + \sum_{j=1}^{J-1} [\varphi(\theta_j, r_j, k_{j+1}) - \varphi(\theta_{j+1}, r_{j+1}, k_{j+1})]$. Liang (2012) and Kumar (2012) demonstrate that an additional term $\rho(w_1, y_1)$ is needed to address censoring at zero labor supply (which I allow for here).

[22] 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.

3. Extensions

3.1 Augmented preferences

It is simple to use the developed framework to incorporate additional features of the labor supply decision. In this subsection, I show how to extend the model to account for fixed costs of work, labor demand quantity restrictions, and the stigma cost of welfare participation. These are features that can be addressed using the standard discrete choice method but not by Hausman's (1985) or Blomquist and Newey's (2002) methods.

In the data, one typically observes excess density at zero labor supply. One explanation could be that it is the outcome that individuals who desire less than zero labor supply (such as paying for more leisure, if this choice were available) would choose. Even if the distribution of desired labor supply is smooth, observed labor supply may bunch. Another explanation is that there are fixed costs of work such as commuting costs. The standard discrete choice method typically has to model fixed costs to fit the excess density observed in the data. We can allow for such costs here by introducing an indicator variable $y_{-0} = 1(y > 0)$ in the utility function that becomes

$$U^F(c, y, y_{-0}, e), \tag{12}$$

which is decreasing in y_{-0}. Furthermore, rewrite Eq. (5) as

$$u_j^F = u^F(y_j, e) = max\left[U^F(c_j, y_j, 1, e), U^F(c_j, 0, 0, e)\right]. \tag{13}$$

Fixed costs of work allow for a discrete jump in utility, and, clearly, convex budget-constrained preferences may not hold at zero labor supply. To check whether an interior option is optimal, we also must therefore always compare it with the zero labor supply option. Using Eq. (13), Assumption 2 can be appropriately modified, and it implies and comes to require that the fixed-cost-augmented preferences are convex on the budget frontier such that for each option, we choose the option $(c_j, y_j, 1)$ or $(c_0, y_0, 0)$ with the higher utility. Non-augmented preferences being convex imply that fixed-cost-augmented preferences are convex, and the modified assumption is therefore less restrictive. Through a similar derivation, such as for Theorem 1, we can derive the expected labor supply expression in Theorem 2.

Theorem 2. Given Assumptions 1 and 2 and the utility function in Eq. (12), which can account for fixed costs of work, expected labor supply is

$$\begin{aligned} E^F(y) &= \sum_{j=0}^{J-1} \pi\left(c_j, c_j', c_0, y_j\right), \\ \pi\left(c_j, c_j', c_0, y_j\right) &= \iint \left[y\left(y_j + \delta, \varepsilon\right) - y\left(y_j, \varepsilon\right)\right] dF\left(e, \varepsilon | u_{j+1}^F > u_j^F\right). \end{aligned} \tag{14}$$

We see that accounting for fixed costs of work increases the dimensionality of the expected labor supply function by one compared with the basic function in Eq. (11).

It is possible to introduce discrete labor demand quantity constraints in labor supply in a conceptually and methodologically similar manner as that for fixed costs of work by letting certain options carry a disutility. In the case of hours of work, take the case of full-time to be what is normally offered by employers. Other options may then be rarely offered and associated with an additional search cost. Describe full-time work by $y_{40} = 1(y = 40)$. The utility function may then be described by $U^{40} = U(c, y, y_{40}, e)$, which is increasing in y_{40}, where instead of giving other options a disutility, we give full-time work an extra utility. Derivation of the labor supply function may then proceed as it does with fixed costs, resulting in a labor supply function that depends on the additional argument c_{40}. Obviously, allowing for too many discrete constraints increases the dimensionality to a great extent, and it decreases the possibility of feasible estimation.

When the quantity constraints are continuous in y, the original utility function in Eq. (4) that allows a continuous effect of y on the utility suffices. Assumption 2 then comes to require that the quantity-constraints-augmented preferences are convex on the budget frontier. If the number of options is larger in the interior (such as for the commonly concavely distributed earnings distribution), the augmented preferences on the budget frontier would normally be more concave and, therefore, less restrictive.

Now, let us augment the utility function to incorporate the stigma cost of welfare participation. Describe the budget frontier without welfare participation $c(y)$, the budget frontier with welfare participation $c^S(y)$, and the stigma cost s. The utility function then becomes

$$U^S(c, y, s, e). \tag{15}$$

Furthermore, rewrite Eq. (5) as

$$u_j^S = u^S(y_j, e) = max[U^S(c_j, y_j, 0, e), U^S(c_j^S, y_j, 1, e)] \tag{16}$$

Assumption 2 is now a requirement that the welfare-participation-augmented preferences are convex on the welfare-participation-augmented budget frontier where for each option, we choose the option $(c_j, y_j, 0)$ or $(c_j^s, y_j, 1)$ with the higher utility. The expected labor supply function increases in dimensionality by two compared with the basic function in Eq. (11) and is presented in Theorem 3.

Theorem 3. Given Assumptions 1 and 2 and the utility function in Eq. (15) which can account for the stigma cost of welfare participation, expected labor supply is

$$\begin{aligned} E^S(y) &= \sum_{j=0}^{J-1} \pi(c_j, c_j', c_j^S, c_j'^S, y_j), \\ \pi(c_j, c_j', c_j^S, c_j'^S, y_j) &= \iint [y(y_j + \delta, \varepsilon) - y(y_j, \varepsilon)] dF(e, \varepsilon | u_{j+1}^S > u_j^S) \end{aligned} \tag{17}$$

3.2 Revenue and distribution functions

One drawback of focusing on expected labor supply is that this entity does not provide knowledge about the distributional effects of the budget set. Using methods that make distributional assumptions, one can typically estimate distributional parameters and obtain the distributional effects through simulations. In this subsection, I show how to derive estimable functions that depend on the distributional effects. The functions include the expected revenue function, and the cumulative distribution and probability density functions.

Observe that instead of integrating over y in Theorems 1 to 3, we could replace the integrand by functions of y. With regard to revenue from labor supply defined as the difference in revenue between a labor supply option and not working, we have:

$$r(y) = r(c, c_0, y, y_0,) = (y - y_0) - (c - c_0) \tag{18}$$

The expected revenue function increases in dimensionality by one compared with the expected labor supply function in Eq. (11) and is presented in Theorem 4.

Theorem 4. Given Assumptions 1 and 2, the utility function in Eq. (1), and the revenue function in Eq. (18), the expected revenue function is:

$$\begin{aligned} E(r) &= \sum_{j=0}^{J-1} \pi\left(c_j, c_j', c_0, y_j\right), \\ \pi\left(c_j, c_j', c_0, y_j\right) &= \iint \left[r\left(c_j + c_j', c_0, y_j + \delta, \varepsilon\right) - r\left(c_j, c_0, y_j, \varepsilon\right)\right] dF\left(e, \varepsilon | u_{j+1} > u_j\right) \end{aligned} \tag{19}$$

Now I derive the cumulative distribution and probability density functions (cdf and pdfs). The cumulative distribution function at k is defined as

$$y_k(y) = 1[y > k] \tag{20}$$

Labor supply participation is the special case in which $k = 0$. A probability density function can be defined as

$$y_{k,l}(y) = 1[y > k \text{ and } y \leq l] \tag{21}$$

The expectation functions are presented in Theorem 5.

Theorem 5. Given Assumptions 1 and 2 and the utility function in Eq. (1), $\Pr(y > k)$ and $\Pr(k < y \leq l)$ are:

$$\begin{aligned} Pr(y > k) &= \sum_{j=0}^{J-1} \pi\left(c_j, c_j', y_j\right), \\ \pi\left(c_j, c_j', y_j\right) &= \iint \Big[1\big[y\left(y_j + \delta, \varepsilon\right) > k\big] - 1\big[y\left(y_j, \varepsilon\right) > k\big]\Big] dF\left(e, \varepsilon | u_{j+1} > u_j\right), \end{aligned} \tag{22}$$

$$\begin{aligned} Pr(k < y < l) &= \sum_{j=0}^{J-1} \pi\left(c_j, c_j', y_j\right), \\ \pi\left(c_j, c_j', y_j\right) &= \iint \Big[1\big[k < y\left(y_j + \delta, \varepsilon\right) \leq l\big] - 1\big[k < y\left(y_j, \varepsilon\right) \leq l\big]\Big] dF\left(e, \varepsilon | u_{j+1} > u_j\right). \end{aligned} \tag{23}$$

It is also possible to recover the pdfs from the estimates of the cdf over the entire earnings range. Furthermore, the quantile function is the inverse cumulative distribution function, and the distribution of quantile effects can therefore also be recovered from the cdf effects.

3.3 Hours of work function

It is possible to use the basic framework to estimate hours of work functions by letting the outcome variable be hours of work, now denoted by h. The budget frontier would then be determined not only by the tax system but also by the gross wage rate w according to $c_j = c(h_j w)$. The gross wage rates therefore provides additional variation in the budget set

$$B = (c_0, h_0, \ldots, c_J, h_J). \tag{24}$$

This is the standard approach to modeling hours of work. An advantage is that useful variation in the gross wage rate helps identification considerably when there is little variation in the tax system. However, it requires that the gross wage rate is exogenous and observed. If it is not observed for some workers, e.g., nonworking individuals, we could predict the wage rate, although the prediction error would need to be addressed.[23] Often, the analyst only has access to earnings and hours of work and can only obtain a measure of the gross wage rate as $w = y/h$ where y is used to denote earnings here. In these cases, the prediction error often generates a large division bias in the estimation as shown by Sacklén and Eklöf (1999).

There is an alternative approach to model hours of work within the framework here that does not require the gross wage rate to be observed or predicted. Let the utility function be

$$U(c, y, w, e) = U(c, h(y, w), e). \tag{25}$$

Retain $c_j = c(y_j)$ and the description of the budget set $B = (c_0, y_0, \ldots, c_J, y_J)$. Now let the gross wage rate be a function of a gross wage rate prediction w^p and a prediction error ε^w, i.e., $w = w(w^p, \varepsilon^w)$. In the most extreme form, we can set the prediction to a constant. Now, incorporate ε^w as an element in the error terms e and ε in the utility and observed hours of work functions, respectively. The expected hours of work function is almost identical to the expected earnings function and is presented in Theorem 6.

Theorem 6. Given Assumptions 1 and 2 and the utility function in Eq. (25), the expected hours of work function is:

$$\begin{aligned} E(h) &= \sum_{j=0}^{J-1} \pi(c_j, c_j', y_j), \\ \pi(c_j, c_j', c_0, y_j) &= \iint [h(y_j + \delta, \varepsilon) - h(y_j, \varepsilon)] dF(e, \varepsilon | u_{j+1} > u_j) \end{aligned} \tag{26}$$

A crucial assumption in Theorem 6 (as with most other methods) is that the gross wage rate prediction error is exogenous. This can be relaxed by introducing observed heterogeneity z and letting the prediction depend on z, i.e., $w^p = w^p(z)$ and $w = w(z, \varepsilon_w)$. As in the case when observed heterogeneity is included for earnings, this would increase the dimensionality of the labor supply function by the dimensionality of z. However, we do not actually need to estimate the gross wage rate function.

[23] Unobserved wage rates due to censoring at zero hours of work are a common obstacle when trying to estimate female labor supply functions. Kumar (2013) provides a recent application that fully accounts for this complication by imputing wage rates.

4. Estimation

4.1 Polynomial approximation

The terms in the labor supply functions in Theorems 1 to 6 can be flexibly approximated using polynomials and estimated by least squares. Polynomials are convenient because the equality constraints implied by the model are easily incorporated into the estimation by excluding interaction terms (Stone, 1985) and by imposing the equality of coefficients (Porter, 1998). I show how to approximate and estimate the basic expected labor supply function in Eq. (11). Approximate it as

$$\sum_{j=0}^{J-1} \pi\left(c_j, c_j', y_j\right) \sim x^P = (x_1, \dots, x_P),\ \ x_p = \sum_{j=0}^{J-1} c_j^{p_c(p)} c_j'^{p_{c'}(p)} y_j^{p_y(p)}. \tag{27}$$

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 a measure of the average slope in the budget set, and it 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 the number of these possibilities that 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 weights to net incomes and net income derivatives at different labor supply levels. The second-order terms characterize the budget set beyond its average level and starting and ending points by providing measures of *where* net incomes appear in the budget set. The same holds true for terms of even higher 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 cannot be included in the estimation. The results are, however, quite insensitive to the exact set of higher-order terms included.

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 the different regions of the budget set.

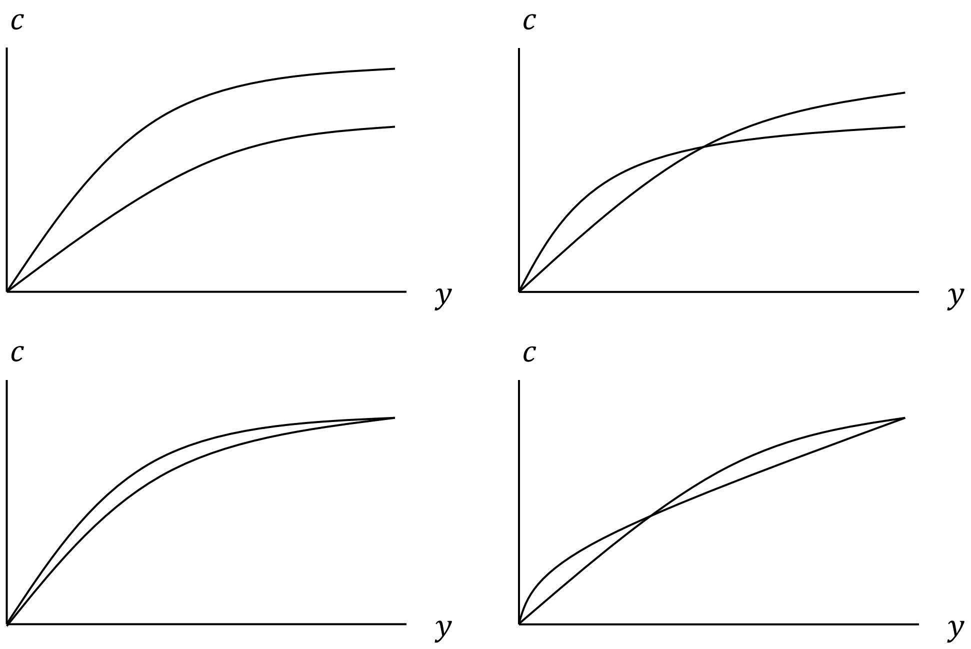

Figure 2. Budget sets with different approximating terms

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

Blomquist and Newey (2002) also approximate their expected labor supply expression using polynomials.[24] Because the labor supply expression here is more parsimonious, fewer approximating terms are needed for the expression. The difference arises because each term here makes use of characteristics of the entire budget set and pools the variation.[25] 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, which are common problems of higher-order terms, less likely.[26] These advantages may not matter in some applications but could be decisive for precision in other applications, such as in applications with cross-sectional data.

4.2 Selection of approximating terms

It is possible to select the approximating terms more or less flexibly or even nonparametrically. This can be done using cross-validation as in Blomquist and Newey (2002) or using Lasso as in Blomquist et al. (2014). I use the cross-validation criterion

[24] They also demonstrate that their sum term (see the second term of the labor supply expression in footnote 21) 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 also applicable here, and it thereby results in a convergence rate that is independent of the number of points used to approximate the budget frontier.

[25] 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 set is therefore needed for identification, unlike here.

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

$$CV = 1 - \frac{SSE}{\sum_i (y - \hat{y}_i)^2}, \quad SSE = \sum_i \left(y - \hat{y}_i^{-i}\right)^2. \tag{28}$$

SSE is the sum of squares of the forecast errors where the forecast of observation i, $\hat{y}_i^{-i}$, is the predicted value using all but the i-th observation. Dividing SSE by the sample sum of squares of y makes the cross-validation value invariant to the scale of y. Maximizing the cross-validation value minimizes the asymptotic mean-square error, and the bias decreases to zero at the same rate as the standard error according to Andrews (1991). More terms are needed for asymptotic inference so that the bias decreases to zero more quickly than the standard error. I choose specification by sequentially adding the lowest-order term that has not yet been added and that performs well regarding the cross-validation criterion compared with the other terms. I continue until the cross-validation value begins to decrease substantially.

4.3 Labor supply elasticities

The parameters of the labor supply functions describe the labor supply effects of various aspects of the budget set. Individually, however, they are completely uninformative, but they can be used to construct measures of the responsiveness of labor supply to budget set changes. In linear or linearized budget sets, the responsiveness can be measured by the labor supply effect of varying the slope or the intercept. In nonlinear budget sets, there is not a single slope or intercept. I follow the approach taken by Blomquist et al. (2014), who measure the responsiveness to slope changes by tilting the slope of each segment in the piecewise-linear budget set. They show that tilting the slope by the same absolute amount keeps the intercept of each segment fixed.[27] The linear budget set slope elasticity is a special case that tilts the only slope keeping the only intercept fixed. From a policy perspective, the tilt corresponds to the rate decrease of a proportional tax such as a local proportional tax.[28]

Adapted to the discretized budget set, tilting the slope corresponds to increasing the net income at each labor supply level with an amount that is proportional to the labor supply level, i.e., increasing c_j to $c_j + y_j a$, and $B = (c_0, y_0, \ldots, c_J, y_J)$ to $B^* = (c_0^* = c_0 + y_0 a, y_0, \ldots, c_J^* = c_J + y_J a, y_J)$. Letting $f(B^*)$ be the labor supply function, the marginal slope effect can be constructed as

$$\frac{df(B^*)}{da} = \sum_{j=0}^{J-1} \frac{df(B^*)}{dc_j^*} y_j. \tag{29}$$

A measure of slope elasticity can be obtained by normalizing the marginal effect by a suitable labor supply level and slope. I normalize by the average labor supply level and average slope (over points and individuals) in the data set.

---

[27] It also keeps the kink points fixed. In contrast, a proportional increase in slope, which is often used in the literature, does not keep the intercepts fixed. In piecewise-linear budget sets, it is possible to increase slopes proportionally for every segment in a way that keeps intercepts fixed by shifting kink points proportionally downwards. In that case, however, moving these kink points will have additional effects on labor supply in addition to the slope effect.

[28] It does not, however, correspond to decreases in payroll tax or an increase in the gross wage rate because these also involve changes in intercepts or kinkpoints.

A measure of income elasticity can be constructed by increasing the net income level at all labor supply levels by the same amount, i.e., increasing c_j to $c_j + b$ and $B = (c_0, y_0, \dots, c_J, y_J)$ to $B^{**} = (c_0^{**} = c_0 + b, y_0, \dots, c_J^{**} = c_J + b, y_J)$. The marginal income effect is then

$$\frac{df(B^{**})}{db} = \sum_{j=0}^{J-1} \frac{df(B^{**})}{dc_j^{**}}. \tag{30}$$

I normalize to obtain an elasticity using the average labor supply and average unearned income.

5. Application: The U.S.

5.1 Data

In the first empirical application, I estimate the effects of taxes on earnings, hours of work, and tax revenues for the U.S. using data from the Panel Study of Income Dynamics (PSID). To avoid complications related to how to address productivity growth over time (as discussed by Blomquist et al., 2014), I use a single cross section from 2006. The PSID is a survey data set and contains many demographic variables that are important to include as control variables because variation in budget sets within years is largely determined by demographic variables. In the taxable income literature based on linearizing the budget set, it is important to have accurate information on income variables to avoid measurement errors in the linearized budget set (which depends on the income variables). Therefore, administrative data are often preferred despite the lack of extensive demographic variables in these data. Because I account for the entire budget set, such measurement error problems do not arise here. Instead, the measurement errors in the income variables appear in the labor supply outcomes and are accounted for in the adopted framework.[29]

In the estimation, I limit the sample to households that filed jointly (most households did this) and had heads of household between 21 and 60 years old. This is economically the most important group. Households with income levels above 200,000 USD, which comprise approximately 5% of the sample, are excluded. This leaves 2,775 observations.

Because the marginal tax rates apply to the joint income of the household, each household faces a single budget set. I treat the joint labor supply choice as a single household decision, i.e., using a unitary household model, as in most other papers that use American household data (e.g., Gruber and Saez, 2002). I use NBER-TAXSIM to construct the budget set of each household. Taxes and transfers are calculated for each 1,000 USD in earnings from 0 to 200,000 USD, holding unearned income constant. Unearned income is set to be the net income generated by the gross income difference between adjusted gross income and earnings, and it consists of net capital income and transfers received when having zero earnings. Consumption taxes are not adjusted for (in contrast with the later application on Swedish data).

[29] This is in contrast to methods that focus on estimating the utility function.

I use the following control variables: number of children (3 groups), age (8 groups), education level (12 groups), occupation (11 groups), industry sector (11 groups), and ethnicity (2 groups).

In Appendix B, sample statistics for a number of outcome, budget set, and demographic variables are first reported in Table B1. The distribution of gross earnings and the marginal tax rate schedule for a representative household (residing in Colorado) are then presented in Figure B1. On average, households earned 74,754 USD and worked 3,474 hours. They paid 23% of their gross earnings in taxes, generating average tax revenues of 17,308 USD. The budget sets had, on average, a net income level of 75,919 USD, a net-of-tax rate of 0.70, and an unearned income level of 2,359 USD. Approximately 3% of the households had no earnings. Typically, first-dollar marginal tax rates were approximately 0 and top marginal tax rates were approximately 0.45.

5.2 Results

The estimation results for household earnings are reported in Table 1. I use the model that allows for fixed costs of work in Eq. (14). The empirical specification is described in Eq. (27). Regressors are sequentially added[30] as we proceed to specifications further down in the table and are selected using the cross-validation criterion CV in Eq. (28). Slope and income elasticities are reported as the marginal effects in Eqs. (29) and (30) normalized by the sample average gross earnings, budget set slope, and unearned income.

Table 1. Elasticity estimates for earnings

Budget set regressors	CV	Slope elasticity	Income elasticity
$\sum c, \sum c'$	0.2612	0.49	-0.01
		(0.25)	(0.00)
$\sum c, \sum c', c_0^2, \sum c'^2$	0.2637	0.27	-0.03
		(0.26)	(0.00)
$\sum c, \sum c', c_0^2, \sum c'^2, c_0 \sum c, c_0 \sum c'$	0.2650	0.37	-0.02
		(0.26)	(0.01)
$\sum c, \sum c', c_0^2, \sum c'^2, c_0 \sum c, c_0 \sum c', \sum c'y, c_0$	0.2654	0.27	-0.02
		(0.27)	(0.01)
$\sum c, \sum c', c_0^2, \sum c'^2, c_0 \sum c, c_0 \sum c', \sum c'y, c_0, \sum cc', \sum cy$	0.2641	0.30	-0.02
		(0.29)	(0.01)

Notes: Standard errors are reported in parentheses. All specifications include control variables for number of children, age, education level, occupation, industry sector, and ethnicity. Elasticities are evaluated at the sample average gross earnings level of 74,754 USD, the average slope of 0.70, and the average unearned income level of 2,359 USD.

The cross-validation value first increases when additional regressors are added, until the fourth reported and preferred specification, when the value is 0.2654. After this specification, the value decreases. The estimated slope elasticity is 0.49 in the first specification and then decreases towards 0.30; it is 0.27 in the preferred specification. The standard errors are, however, large. They are approximately the size of the estimated slope elasticities in most

[30] They are added one at a time in the estimation but two at a time in the presented table to save space.

specifications, and the slope elasticities are therefore statistically insignificant.[31] The estimated income elasticity is approximately zero in all specifications. The income elasticities are all statistically insignificant, but they are quite precisely estimated.

The estimation results for household earnings, hours of work, and tax revenues are reported in Table 2. Only estimates from the preferred specification (chosen by the cross-validation criterion) are reported to save space. Estimates without control variables are also reported to illustrate the importance of their inclusion.

Table 2. Elasticity estimates for earnings, hours of work, and tax revenues

Outcome	Budget set regressors	Slope elasticity	Income elasticity
Earnings	$\sum c, \sum c', c_0^2, \sum c'^2, c_0 \sum c, c_0 \sum c', \sum c'y, c_0$	0.27	-0.02
		(0.27)	(0.01)
	No controls	1.01	-0.01
		(0.30)	(0.01)
Hours of work	$\sum c, \sum c', \sum c'^2, \sum cy, \sum c^2, c_0, \sum c'y$	0.01	-0.02
		(0.20)	(0.01)
	No controls	0.51	-0.01
		(0.20)	(0.00)
Tax revenues	$\sum c, \sum c', \sum c'^2, c_0^2, \sum c^2$	-2.43	
		(0.39)	
	No controls	-1.27	
		(0.43)	

Notes: Standard errors are reported in parentheses. Except for the No controls specifications, the control variables included are number of children, age, education level, occupation, industry sector, and ethnicity. Elasticities are evaluated at the sample average gross earnings level of 74,754 USD, the average slope of 0.70, and the average unearned income level of 2,359 USD. The mechanical slope elasticity of tax revenues without behavioral effects is -3.02.[32]

The first specification merely reproduces the results for earnings from the preferred specification in Table 1. For hours of work, the estimated slope elasticity is 0.01 and statistically insignificant. The income elasticity is -0.02 and also statistically insignificant. For tax revenues, the estimated slope elasticity is -2.43 and statistically significant. A uniform increase in marginal net-of-tax rates obtained by decreasing tax rates therefore decreases tax revenues significantly. The estimate can be compared with the loss in tax revenues from a tax decrease when there are no behavioral effects, which would result in a slope elasticity of -3.02. The estimated elasticity is smaller because the tax decrease increases labor supply, generating tax revenues that counteract the mechanical loss. If tax rates were at the Laffer maximum, the slope elasticity would be zero because the behavioral effects would fully counteract the mechanical effect. Finally, we note that not including control variables leads to larger labor supply slope elasticities and therefore a smaller tax revenue elasticity.

[31] Because of the low precision, I do not report results for the estimation of the cdf and pdfs, which are all statistically insignificant. I do, however, do this in the next section for Sweden, for which I have a larger sample and therefore higher precision.

[32] Because increasing unearned income is not directly related to tax rates, the income elasticity of tax revenues is not interesting and therefore not reported.

The earnings and hours of work slope elasticities that I obtain are low compared with the estimates in the previous literature.[33] However, the exact elasticity concept and the point of evaluation often differ between studies. Blomquist et al. (2014) show that the nonlinear budget set slope elasticity that is used here is lower than a linear budget set elasticity.[34]

6. Application: Sweden

6.1 Data

In the second empirical application, I estimate the effects of taxes on earnings and tax revenues for Sweden using data from Hushallens Ekonomi (HEK) provided by Statistics Sweden. Again, I use a single cross section from 2006. The data set contains administrative data on many economic variables as well as demographic variables. Additionally, the data set has been complemented with survey data on a number of economic variables related to the budget set, such as housing allowance variables, which are important components of the budget set that are often ignored.[35] The data set allows precise construction of the budget set. It also contains more observations than that for the U.S., resulting in better precision and enabling me to produce informative estimation results for the cdf and pdfs. However, there is no information on hours of work. In addition to having high-quality data, Sweden is interesting because it has among the highest tax rates in the world. Some researchers think that Sweden may be close to the tax revenue Laffer curve maximum, where tax rate decreases are close to revenue neutral.

Unlike in the U.S., all individuals are taxed individually in Sweden. The sample is limited to married or cohabiting men between 21 and 60 years old. Individuals who received medical leave benefits, parental leave benefits, income from self-employment, or student financial aid above half of the average monthly gross earnings, which were approximately 17,700 SEK,[36] are excluded because of the difficulty of modeling their budget sets.[37] Individuals with incomes above 1 million SEK, who comprise approximately 5% of the sample, are excluded. Out of 6,468 married or cohabiting men between 21 and 60 years old, 4,028 observations remain after the sample restrictions.

[33] Gruber and Saez (2002) report a slope elasticity of 0.26 for the 1990s for taxable income and 0.13 for broad income, whereas Weber (2014) reports a slope elasticity of 0.86 for taxable income and 0.48 for broad income. Earnings are narrower than both taxable income and broad income. For hours of work, Ziliak and Kniesner (1999) report a slope elasticity of 0.15, and Kumar (2007) reports a slope elasticity of 0.14.

[34] This is based on estimating the linear budget set elasticity in a manner that accounts for real budget sets that are nonlinear, as in Kumar (2007). In contrast, the hours of work elasticity based on a percentage increase in gross wage rates often used in the literature (e.g., Aaberge et al., 1999) should be lower than the elasticity concept adopted here (if we believe that the uncompensated slope elasticity is positive) because a gross wage rate change also shifts the kink points downward.

[35] For instance, the administrative data only give information on the housing allowances individuals actually received, but not on how much they would have received if they were at another earnings level, which is needed for constructing the entire budget set.

[36] The SEK/USD exchange rate varied between 6.8 and 8.0 SEK/USD in 2006.

[37] It is difficult to satisfactorily include these sources of income in the budget set. Eliminating everyone with a positive income from these sources would, however, result in a large loss of observations. I describe below how income from these sources is treated in the budget set in the selected sample.

I adopt a secondary earner model in which the husband makes his choice conditional on the wife's choice, as in most other papers that use Swedish data (e.g., Blomquist and Newey, 2002). I use FASIT to construct the budget set of each individual in the sample. FASIT is a tax simulation program similar to NBER-TAXSIM that was developed by Statistics Sweden. It simulates in principle all features of the Swedish tax and transfer system that are relevant for individuals, and it is used by the Swedish Ministry of Finance to simulate the mechanical effects of a variety of tax policies, including potential future policies. Taxes and transfers are calculated for each 10,000 SEK in earnings from 0 to 1 million SEK, holding unearned income constant. Unearned income is set to be net household income when the husband has zero earnings, which includes the net income of the spouse, capital income, and transfers received at these earnings levels. However, I make two modifications. First, medical leave compensation and capital gains are set to zero, because these components are difficult to predict for individuals at the beginning of the year when they are planning how much to work during the year. Second, implicit income from residence-owned housing is included. Additionally, I include a simple rudimentary correction for consumption taxes using the quotient of aggregate consumption tax revenues divided by aggregate private consumption. These additional corrections are similar to those in much of the previous literature (e.g., Blomquist and Newey, 2002).

I include the following control variables: number of children below six years old (2 groups),[38] age (8 groups), education level (7 groups), occupation (8 groups), and county of residence (21 groups).

In Appendix B, sample statistics for a number of outcome, budget set, and demographic variables are first reported in Table B2. The distribution of gross earnings and the marginal tax rate schedules with and without[39] consumption taxes for a representative individual are then presented in Figure B2. On average, the husbands earned 425,069 SEK. They paid 63% of gross earnings in taxes, generating average tax revenues of 265,615 SEK, of which 207,696 SEK came from taxes other than consumption taxes. The budget sets had, on average, a net income level of 353,550 SEK, a net-of-tax rate of 0.31, and an unearned income level of 179,786 SEK. Approximately 5% of the individuals had no earnings. First-dollar marginal tax rates were approximately 0.45 and top marginal tax rates were approximately 0.75. The tax revenues and tax rates were considerably higher in Sweden than in the U.S., even when using the comparable figures without consumption taxes. Part of the higher tax revenue share of income is, however, because the husbands faced a higher marginal tax rate than did the wives (because the husbands earned more), from whom tax revenues were included in the American numbers.

[38] Because families with small children are allowed to take parental leave days, number of small children affects labor supply more than total number of children does.

[39] The marginal tax rate schedule without consumption taxes is presented to allow a comparison with the U.S. schedule in Figure B1.

6.2 Results

The estimation results for the earnings of and tax revenues from married or cohabiting men are reported in Table 3. Only estimates from the preferred specification (chosen by the cross-validation criterion) are reported to save space. The estimated slope elasticity of earnings is 0.42,[40] and standard errors are one-third of the point estimate (0.14). The estimated income elasticity of earnings is 0.05, and standard errors are one-fifth of the point estimate (0.01). The estimated slope elasticity of tax revenues is -0.17. This is roughly one-third of the mechanical slope elasticity without behavioral effects, which is -0.49. Although the earnings slope elasticity is more precisely estimated for Sweden than for the U.S., the tax revenue elasticity is relatively less precisely estimated. The standard errors are too large to rule out behavioral effects or to determine if the behavioral effects completely counteract the mechanical effect and that tax rates therefore are at the tax revenue Laffer curve maximum.

Table 3. Elasticity estimates for earnings and tax revenues

Outcome	Budget set regressors	Slope elasticity	Income elasticity
Earnings	$\sum c, \sum c', c_0 \sum c', \sum cy$	0.42 (0.14)	0.05 (0.01)
Tax revenues	$\sum c, \sum c', \sum c'^2, c_0^2, \sum c^2$	-0.17 (0.24)	

Notes: Standard errors are reported in parentheses. All specifications include control variables for number of children below six years old, age, education level, occupation, and county of residence. Elasticities are evaluated at the average earnings level of 425,069 SEK, the average tax revenue level of 265,615 SEK, the average slope of 0.31, and the average unearned income level of 179,786 SEK. The mechanical slope elasticity of tax revenues without behavioral effects is -0.49.

I also report the estimation results for the cdf and pdfs of slope effects for earnings in Figure 3. Rather than reporting elasticities at each earnings level, I report marginal percentage point effects on earnings of a percent change in average slope to ensure comparability across earnings levels. The functions are estimated in earnings intervals of 50,000 SEK. Smaller intervals lead to a substantial loss of precision for the pdfs.

The estimated cdf shows that decreasing the tax rates uniformly across the budget set increases the probability of having an income above 100,000 SEK. The increase is largest for earnings from above 200,000 to above 500,000 SEK. In this region, the estimated cdf effects are also statistically significant. In terms of pdfs, the probability of having earnings between 50,000 and 200,000 SEK decreases and the probability of having earnings between 350,000 and 800,000 SEK increases. The estimated participation effect is, however, close to zero and statistically insignificant, possibly because of the high participation rate of the husbands. Uniform tax changes therefore primarily prompt individuals with low earnings in the part-time job region to increase their earnings (by working more or exerting more effort), resulting in more individuals beginning to earn close to or more than average earnings. This result is, however, not necessarily achieved by the part-time workers beginning to obtain full-time earnings. It is more likely that the entire distribution above 50,000 SEK marginally shifts upward.

[40] This number can be compared with the 0.3 to 0.8 obtained by Hansson (2007).

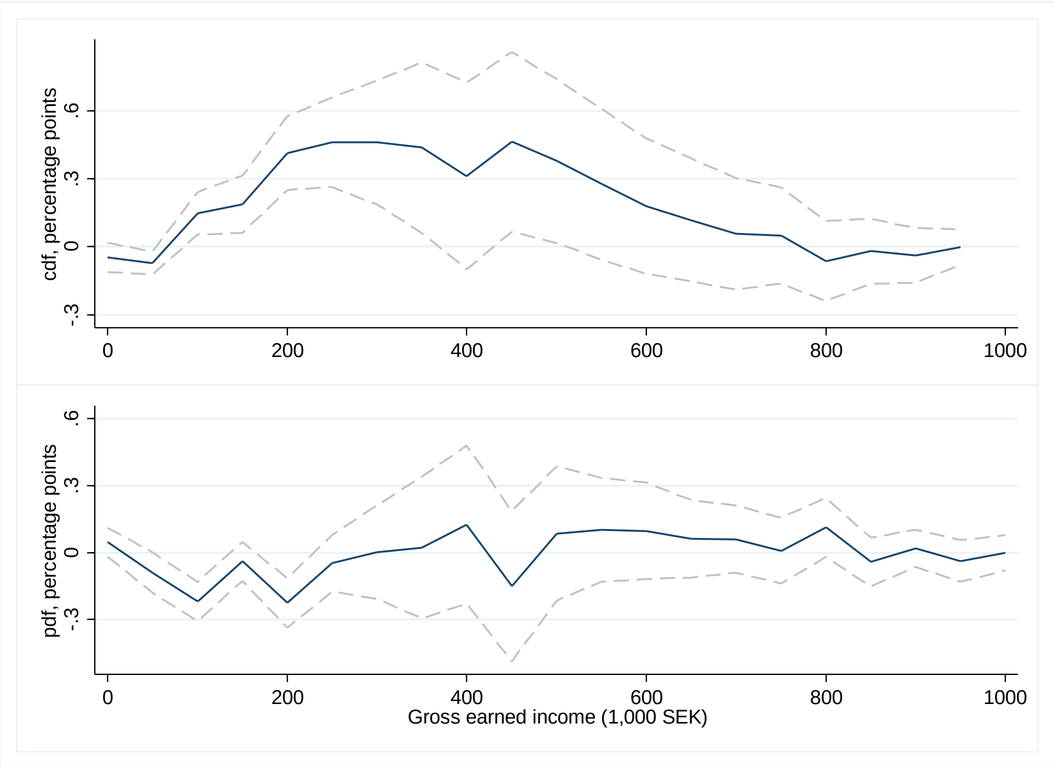

Figure 3. Estimated cdf and pdfs of slope effects for earnings

Notes: Marginal slope effects in percentage points of a percent change in slopes are reported in intervals of 50,000 SEK. The solid lines are the point estimates and the dashed lines are 95% confidence intervals. The cdf at x is $Pr(y > x)$, and the pdf at x is $\Pr(x - 50 < y \leq x)$ except at $x = 0$ where $\Pr(x = 0)$ is reported.

7. Conclusions

I have developed a structural method for evaluating labor supply in nonlinear budget sets that does not require any distributional assumptions. As in the standard discrete choice method, I discretized the labor supply options. However, instead of estimating the utility function, I imposed the assumption that preferences are convex on the budget frontier and derived a three-dimensional expected labor supply function. This function can be approximated flexibly using a polynomial in its arguments, and it can be estimated with least squares. I also showed how to extend the standard model to account for features such as fixed costs of work, labor demand quantity restrictions, and the stigma cost of welfare participation. Furthermore, the framework is applicable to estimating earnings, hours of work, and functions that depend on the labor supply distribution, including tax revenue and cumulative distribution functions.

I then applied the method to estimate the effects of taxes on various labor supply outcomes in the U.S. and Sweden in 2006. For the U.S., I used data from the PSID and investigated households that filed jointly. I found an earnings slope elasticity of 0.27, an hours of work slope elasticity of 0.01, and a tax revenue slope elasticity of -2.43. For Sweden, I used data from HEK and investigated married or cohabiting men. I found an earnings slope elasticity of 0.42 and a tax revenue slope elasticity of -0.17. Furthermore, the results showed that tax decreases primarily decrease the share of individuals with one-eighth to one-half of average earnings and increase the share of individuals with average or above average earnings.

I have primarily discussed how to use the developed framework for analyzing the labor supply effects of taxes. The framework is, however, also useful for other consumer choice applications with nonlinear prices that create nonlinear budget sets, such as health insurance, pension contributions, and electricity plans. There is also more methodological work to be done. In particular, addressing productivity growth would allow for using repeated cross sections or panel data that could provide more observations and a wider range of budget sets from different tax regimes. Furthermore, if there is sufficient variation in budget set changes between years, differencing or fixed-effect techniques could be used to address unobserved individual or group specific heterogeneity.

Acknowledgements

I thank Sören Blomquist, Mikael Elinder, Lennart Flood, Anil Kumar, Whitney Newey, Håkan Selin, and the seminar participants at the Uppsala Center for Fiscal Studies, Department of Economics, Uppsala University, the Workshop on Public Economics and Public Policy in Copenhagen, the 28th Annual Conference of the European Economic Association in Gothenburg, and the 2014 American Economic Association Meeting 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.

Appendix

Appendix A. Proofs

Proof of Lemma 1. Assume $u_j = \max_y u(y,e)$. Then, obviously $u_j > u_{j-1}$ and $u_j \geq u_{j+1}$, proving the implication. Assume $u_j > u_{j-1}$ and $u_j \geq u_{j+1}$. Eq. (8) implies $u_{j-1} > u_{j-2}$ and $u_{j+1} \geq u_{j+2}$ and, more generally, $u_{j-a} > u_{j-a-1}$ and $u_{j+a} \geq u_{j+a+1}$ for every a. Therefore, $u_j > u_{j-b}$ and $u_j \geq u_{j+b}$ for every b proving the reverse implication. Note that, $y_{j-1}(y_j) = y_j - \delta$ and $y_{j+1}(y_j) = y_j + \delta$ where δ is a constant reducing the dimensionality of $\mu_j(.)$. ■

Proof of Lemma 2. Suppose not; then there is a $u_j \leq u_{j-1}$ and $u_j < u_{j+1}$, and then there is a $k = j-1$ and an $l = j+1$ such that Eq. (8) does not hold. ■

Proof of Corollary 1. Suppose not, i.e., that we have utility maximization with convex preferences in convex budget sets but that Eq. (8) does not hold. Then, there is a $u_j < \min(u_k, u_l)$ such that $y_k < y_j = a y_k + (1-a) y_l < y_l$, $0 < a < 1$. Convexity of preferences implies that $U(c_a = a c_k + (1-a) c_l, y_j) > \min(u_k, u_l)$. Convexity of the budget set implies that $c_j > c_a$ and $u_j > U(c_a, y_j)$. However, then $u_j > \min(u_k = u_l)$, which is a contradiction.[41] ■

Proof of Theorem 1. Lemma 2 implies

$$\{(c_j, y_j): u_j > u_{j-1} \text{ and } u_j \geq u_{j+1}\} = \{(c_j, y_j): u_j > u_{j-1}\} - \{(c_j, y_j): u_{j+1} \geq u_j\}. \qquad \text{(A1)}$$

The law of total expectation gives

$$E(y) =^{L1} \sum_{j=0}^{J-1} \iint y(y_j, \varepsilon) dF(e, \varepsilon | u_j > u_{j-1} \text{ and } u_j \geq u_{j+1})$$

$$+ \iint y(y_0, \varepsilon) dF(e, \varepsilon | u_0 \geq u_1) + \iint y(y_J, \varepsilon) dF(e, \varepsilon | u_J > u_{J-1})$$

$$=^{\text{Eq.(A1)}} \sum_{j=1}^{J-1} \iint y(y_j, \varepsilon) dF(e, \varepsilon | u_j > u_{j-1}) - \sum_{j=1}^{J-1} \iint y(y_j, \varepsilon) dF(e, \varepsilon | u_{j+1} > u_j)$$

$$- \iint y(y_0, \varepsilon) dF(e, \varepsilon | u_1 > u_0) + \iint y(y_J, \varepsilon) dF(e, \varepsilon | u_J > u_{J-1}).$$

I used Lemma 1 ($L1$) for the first equality and Eq. (A1), an implication of Lemma 2, for the second equality. Collecting the terms gives Eq. (11). ■

Proof of Theorem 2. Replacing u_j in Eq. (5) by u_j^F in Eq. (13), Assumption 2 implies a modified Lemma 1 for interior solutions:

[41] For simplicity, I assume strict convexity. The budget-constrained preferences are strictly convex if either preferences or the budget set are strictly convex.

$$y^d = y_j \quad \text{if and only if}$$
$$\mu_j'^F(c_j, c_{j-1}, c_{j+1}, y_j, c_0, e) \Leftrightarrow u_j^F > u_{j-1}^F, u_j^F \geq u_{j+1}^F, \text{and } u_j > u_0. \tag{A2}$$

Otherwise, $y^d = 0$. Note that, y_{-0} and y_0 are fixed between individuals at any given j and therefore drop out as arguments from $\mu_j'^F(.)$. A comparison of Eq. (A2) and Eq. (9) shows that the dimensionality of $\mu'^F(.)$ increases by the argument c_0. It is now straightforward to follow the derivation steps in Theorem 2 to derive Eq. (14). ■

Proof of Theorem 3. Lemma 1 becomes the following for an interior option:

$$y^d = y_j \text{ if and only if}$$
$$\mu_j'^S(c_j, c_{j-1}, c_{j+1}, c_j^S, c_{j-1}^S, c_{j+1}^S, y_j, e) \Leftrightarrow u_j^S > u_{j-1}^S \text{ and } u_j^S > u_{j+1}^S. \tag{A3}$$

It is now straightforward to follow the derivation steps in Theorem 1 to derive Eq. (17). ■

Proof of Theorem 4. Lemma 1 becomes the following for an interior option:

$$r^d = r_j = r(c_j, c_0, y_j) \quad \text{if and only if } \mu_j'(c_j, c_{j-1}, c_{j+1}, y_j, e). \tag{A4}$$

It is now straightforward to follow the derivation steps in Theorem 1 to derive Eq. (19). ■

Proof of Theorem 5. Lemma 1 becomes the following for an interior option:

$$y_k^d = 1[y^d = y_j > k] \quad \text{if and only if } \mu_j'(c_j, c_{j-1}, c_{j+1}, y_j, e). \tag{A5}$$

$$y_{k,l}^d = 1[k < y^d = y_j \leq l] \quad \text{if and only if } \mu_j'(c_j, c_{j-1}, c_{j+1}, y_j, e). \tag{A6}$$

It is now straightforward to follow the derivation steps in Theorem 1 to derive Eqs. (22) and (23). ■

Proof of Theorem 6. Lemma 1 then becomes the following for an interior option:

$$h^d = h^d(y^d, \varepsilon^w) = \frac{y_j}{w(\varepsilon^w)} \quad \text{if and only if } \mu_j'(c_j, c_{j-1}, c_{j+1}, y_j, e), \tag{A7}$$

Then, $h = h(h^d, \varepsilon) = h(y^d, \varepsilon)$. It is now straightforward to follow the derivation steps in Theorem 1 to derive Eq. (26). ■

Appendix B. Sample statistics

Table B1. Sample statistics for the U.S.

Variable	Mean	Std. dev.	Min	Max
Earnings	74,754	41,847	0	200,000
Hours of work	3,474	1,257	0	8,514
Tax revenues	17,308	13,578	-4,499	72,355
$\sum c$	75,919	5,327	56,499	119,750
$\sum c'$	0.70	0.03	0.56	0.85
c_0	2,359	6,639	-37,905	44,183
Age	41.69	10.32	21.00	60.00
Children	1.18	1.19	0	7

Notes: Earnings, tax revenues, $\sum c$, $\sum c'$, and c_0 are in USD, and age is in years. $\sum c$ and $\sum c'$ are normalized by 1/number of points (200 points). Tax revenues do not include consumption taxes.

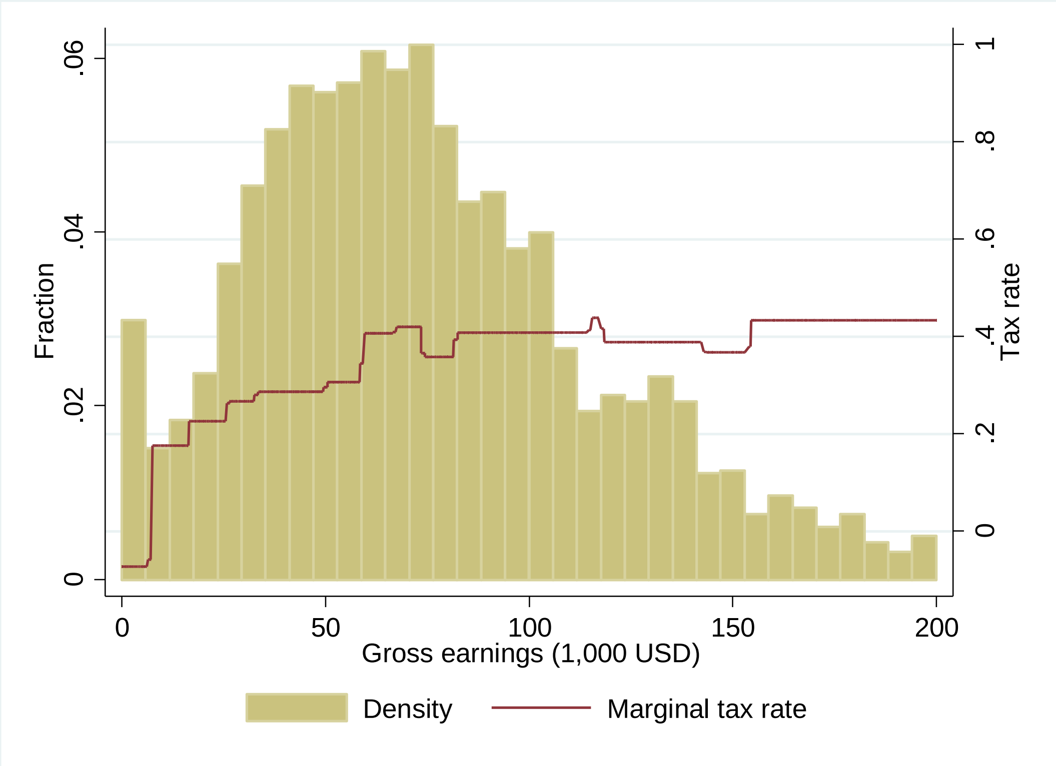

Figure B1. Distribution of gross earnings and marginal tax rate schedule for a representative individual in the U.S. in 2006

Table B2. Sample statistics for Sweden

Variable	Mean	Std. dev.	Min	Max
Earnings	425,069	185,145	0	997,990
Tax revenues	265,615	125,671	0	729,534
Tax revenues no VAT	207,696	104,567	0	644,136
$\sum c$	352,550	157,556	-169,094	5,538,507
$\sum c'$	0.31	0.02	0.13	0.42
c_0	179,786	157,117	-397,578	5,356,334
Age	43.67	10.01	21	60
Children < 6 years	0.32	0.64	0	4

Notes: Earnings, tax revenues, tax revenues No VAT, $\sum c$, $\sum c'$, and c_0 are in SEK, and age is in years. $\sum c$ and $\sum c'$ are normalized by 1/number of points (100 points). Tax revenues include consumption taxes and tax revenues no VAT do not include consumption taxes.

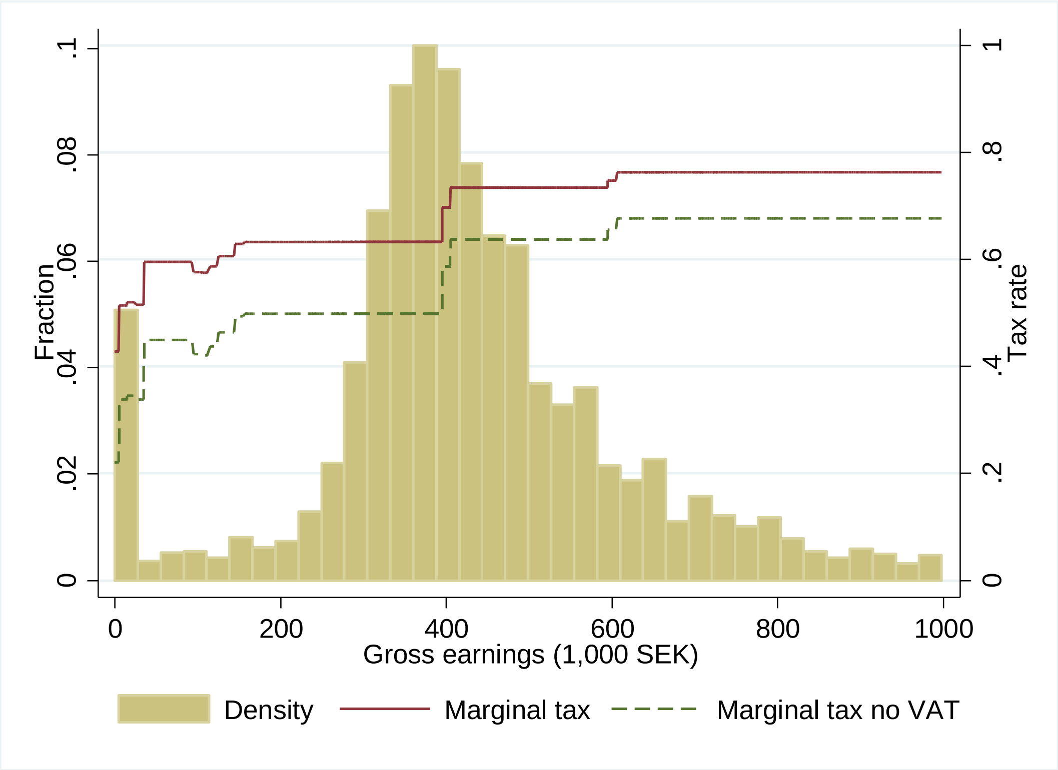

Figure B2. Distribution of gross earnings and marginal tax rate schedule for a representative individual in Sweden in 2006

Notes: Marginal tax includes consumption taxes and marginal tax no VAT does not include consumption taxes.

References

Aaberge, R., Colombino, U., and Strom, S., 1999. Labor supply in Italy: an empirical analysis of joint household decisions, with taxes and quantity constraints. Journal of Applied Econometrics 14, 403-422.

Aaberge, R., Dagsvik, J., Strom, S., 1995. Labor supply responses and welfare effects of tax reforms, Scandinavian Journal of Economics 97, 635-659.

Andrews, D., 1991. Asymptotic optimality of generalized C_L, cross-validation, and generalized cross-validation in regressions with heteroskedastic errors. Journal of Econometrics 47, 359-377.

Blomquist, S., 1996. Estimation methods for male labor supply functions: how to take account of nonlinear taxes. Journal of Econometrics 70, 383-405.

Blomquist, S., Kumar, A., Liang, C., Newey, W., 2014. Individual heterogeneity, nonlinear budget sets, and taxable income. Uppsala Center for Fiscal Studies Working Paper 2014:1.

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

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.

Dagsvik, J., 1994. Discrete and continuous choice, max-stable processes and independence from irrelevant attributes. Econometrica 62, 1179-1205.

Diewert, W., 1971. Choice on labour markets and the theory of the allocation of time, Manuscript, Research Branch, Department of Manpower and Immigration, Ottawa.

Eklöf, M., Sacklén, H., 1999. The Hausman-MaCurdy controversy: why do the results differ across studies. Journal of Human Resources 35, 204-220.

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.

Giertz, S., 2008. Panel data techniques and the elasticity of taxable income. Congressional Budget Office Working Paper 2008-20.

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.

Ichimura, H., 1993. Semiparametric least squares (SLS) and weighted SLS estimation of single-index models. Journal of Econometrics 58, 71-120.

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

Kopczuk, W., 2005. Tax bases, tax rates and the elasticity of reported income. Journal of Public Economics 89, 2093-2119.

Kumar, A., 2007. Labor supply, deadweight loss and tax reform act of 1986: a nonparametric evaluation using panel data. Journal of Public Economics 92, 236-253.

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

Kumar, A., 2013. Lifecycle-consistent female labor supply with nonlinear taxes: evidence from unobserved effects panel data models with censoring, selection and endogeneity. Review of Economics of the Household, doi: 10.1007/s11150-013-9217-6.

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.

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.

Saez, E., Slemrod, J., Giertz, S., 2012. The elasticity of taxable income with respect to marginal tax rates: a critical review. Journal of Economic Literature 50, 3-50.

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., 2014. Toward obtaining a consistent estimate of the elasticity of taxable income using difference-in-differences. Journal of Public Economics, forthcoming.

Ziliak, J, Kniesner, P., 1999. Estimating life cycle labor supply tax effects. Journal of Political Economy 107, 326–359.

WORKING PAPERS
Uppsala Center for Fiscal Studies
Editor: Håkan Selin

2009:1 Sören Blomquist and Håkan Selin, Hourly Wage Rate and Taxable Labor Income Responsiveness to Changes in Marginal Tax Rates. 31 pp.

2009:2 Luca Micheletto, Optimal nonlinear redistributive taxation and public good provision in an economy with Veblen effects. 26 pp.

2009:3 Håkan Selin, The Rise in Female Employment and the Role of Tax Incentives. An Empirical Analysis of the Swedish Individual Tax Reform of 1971. 38 pp.

2009:4 Håkan Selin, Marginal tax rates and tax-favoured pension savings of the self-employed Evidence from Sweden. 32 pp.

2009:5 Tobias Lindhe and Jan Södersten, Dividend taxation, share repurchases and the equity trap. 27 pp.

2009:6 Che-Yan Liang, Nonparametric Structural Estimation of Labor Supply in the Presence of Censoring. 48 pp.

2009:7 Sören Blomquist, Vidar Christiansen and Luca Micheletto, Public Provision of Private Goods and Nondistortionary Marginal Tax Rates: Some further Results. 42 pp.

2009:8 Laurent Simula and Alain Trannoy, Optimal Income Tax under the Threat of Migration by Top-Income Earners. 26 pp.

2009:9 Laurent Simula and Alain Trannoy, Shall We Keep Highly Skilled at Home? The Optimal Income Tax Perspective. 26 pp.

2009:10 Michael Neugart and Henry Ohlsson, Economic incentives and the timing of births: Evidence from the German parental benefit reform 2007, 21 pp.

2009:11 Laurent Simula, Optimal Nonlinear Income Tax and Nonlinear Pricing: Optimality Conditions and Comparative Static Properties, 25 pp.

2009:12 Ali Sina Onder and Herwig Schlunk, Elderly Migration, State Taxes, and What They Reveal, 26 pp.

2009:13 Ohlsson, Henry, The legacy of the Swedish gift and inheritance tax, 1884-2004, 26 pp.

2009:14 Onder, Ali Sina, Capital Tax Competition When Monetary Competition is Present, 29 pp.

2010:1 Sören Blomquist and Laurent Simula, Marginal Deadweight Loss when the Income Tax is Nonlinear. 21 pp.

2010:2 Marcus Eliason and Henry Ohlsson, Timing of death and the repeal of the Swedish inheritance tax. 29 pp.

2010:3 Mikael Elinder, Oscar Erixson and Henry Ohlsson, The Effect of Inheritance Receipt on Labor and Capital Income: Evidence from Swedish Panel Data. 28 pp.

2010:4 Jan Södersten and Tobias Lindhe, The Norwegian Shareholder Tax Reconsidered. 21 pp.

2010:5 Anna Persson and Ulrika Vikman, Dynamic effects of mandatory activation of welfare participants. 37 pp.

2010:6 Ulrika Vikman, Does Providing Childcare to Unemployed Affect Unemployment Duration? 43 pp.

2010:7 Per Engström, Bling Bling Taxation and the Fiscal Virtues of Hip Hop. 12 pp.

2010:8 Niclas Berggren and Mikael Elinder, Is tolerance good or bad for growth? 34 pp.

2010:9 Heléne Lundqvist, Matz Dahlberg and Eva Mörk, Stimulating Local Public Employment: Do General Grants Work? 35pp.

2010:10 Marcus Jacob and Martin Jacob, Taxation, Dividends, and Share Repurchases: Taking Evidence Global. 42 pp.

2010:11 Thomas Aronsson and Sören Blomquist, The Standard Deviation of Life-Length, Retirement Incentives, and Optimal Pension Design. 24 pp.

2010:12 Spencer Bastani, Sören Blomquist and Luca Micheletto, The Welfare Gains of Age Related Optimal Income Taxation. 48 pp.

2010:13 Adrian Adermon and Che-Yuan Liang, Piracy, Music, and Movies: A Natural Experiment. 23 pp.

2010:14 Spencer Bastani, Sören Blomquist and Luca Micheletto, Public Provision of Private Goods, Tagging and Optimal Income Taxation with Heterogeneity in Needs. 47 pp.

2010:15 Mattias Nordin, Do voters vote in line with their policy preferences? The role of information. 33 pp.

2011:1 Matz Dahlberg, Karin Edmark and Heléne Lundqvist, Ethnic Diversity and Preferences for Redistribution. 43 pp.

2011:2 Tax Policy and Employment: How Does the Swedish System Fare? Jukka Pirttilä and Håkan Selin. 73 pp.

2011:3 Alessandra Casarico, Luca Micheletto and Alessandro Sommacal, Intergenerational transmission of skills during childhood and optimal public policy. 48 pp.

2011:4 Jesper Roine and Daniel Waldenström, On the Role of Capital Gains in Swedish Income Inequality. 32 pp.

2011:5 Matz Dahlberg, Eva Mörk and Pilar Sorribas-Navarro, Do Politicians' Preferences Matter for Voters' Voting Decisions?

2011:6 Sören Blomquist, Vidar Christiansen and Luca Micheletto, Public provision of private goods, self-selection and income tax avoidance. 27 pp.

2011:7 Bas Jacobs, The Marginal Cost of Public Funds is One. 36 pp.

2011:8 Håkan Selin, What happens to the husband's retirement decision when the wife's retirement incentives change? 37 pp.

2011:9 Martin Jacob, Tax Regimes and Capital Gains Realizations. 42 pp.

2011:10 Katarina Nordblom and Jovan Zamac, Endogenous Norm Formation Over the Life Cycle – The Case of Tax Evasion. 30 pp.

2011:11 Per Engström, Katarina Nordblom, Henry Ohlsson and Annika Persson, Loss evasion and tax aversion. 43 pp.

2011:12 Spencer Bastani and Håkan Selin, Bunching and Non-Bunching at Kink Points of the Swedish Tax Schedule. 39 pp.

2011:13 Svetlana Pashchenko and Ponpoje Porapakkarm, Welfare costs of reclassification risk in the health insurance market. 41 pp.

2011:14 Hans A Holter, Accounting for Cross-Country Differences in Intergenerational Earnings Persistence: The Impact of Taxation and Public Education Expenditure. 56 pp.

2012:1 Stefan Hochguertel and Henry Ohlsson, Who is at the top? Wealth mobility over the life cycle. 52 pp.

2012:2 Karin Edmark, Che-Yuan Liang, Eva Mörk and Håkan Selin, Evaluation of the Swedish earned income tax credit. 39 pp.

2012:3 Martin Jacob and Jan Södersten, Mitigating shareholder taxation in small open economies? 15 pp.

2012:6 Spencer Bastani, Gender-Based and Couple-Based Taxation. 55 pp.

2012:7 Indraneel Chakraborty, Hans A. Holter and Serhiy Stepanchuk, Marriage Stability, Taxation and Aggregate Labor Supply in the U.S. vs. Europe. 64 pp.

2012:8 Niklas Bengtsson, Bertil Holmlund and Daniel Waldenström, Lifetime versus Annual Tax Progressivity: Sweden, 1968–2009. 56 pp.

2012:9 Firouz Gahvari and Luca Micheletto, The Friedman rule in a model with nonlinear taxation and income misreporting. 36 pp.

2012:10 Che-Yuan Liang and Mattias Nordin, The Internet, News Consumption, and Political Attitudes. 29 pp.

2012:11 Gunnar Du Rietz, Magnus Henrekson and Daniel Waldenström, The Swedish Inheritance and Gift Taxation, 1885–2004. 50 pp.

2012:12 Thomas Aronsson, Luca Micheletto and Tomas Sjögren, Public Goods in a Voluntary Federal Union: Implications of a Participation Constraint. 7 pp.

2013:1 Markus Jäntti, Jukka Pirttilä and Håkan Selin, Estimating labour supply elasticities based on cross-country micro data: A bridge between micro and macro estimates? 35pp.

2013:2 Matz Dahlberg, Karin Edmark and Heléne Lundqvist, Ethnic Diversity and Preferences for Redistribution: Reply. 23 pp.

2013:3 Ali Sina Önder and Marko Terviö, Is Economics a House Divided? Analysis of Citation Networks. 20 pp.

2013:4 Che-Yuan Liang, Optimal Inequality behind the Veil of Ignorance. 26 pp.

2013:5 Spencer Bastani, Sören Blomquist and Jukka Pirttilä, How Should Commodities Be Taxed? A Counterargument to the Recommendation in the Mirrlees Review. 26 pp.

2013:6 Tobias Laun and Johanna Wallenius, Social Insurance and Retirement: A Cross-Country Perspective. 37 pp.

2013:7 Johannes Hagen, A History of the Swedish Pension System. 140 pp.

2013:8 Etienne Lehmann, Laurent Simula and Alain Trannoy, Tax Me If You Can! Optimal Nonlinear Income Tax between Competing Governments. 29 pp

2013:9 Tobias Lindhe and Jan Södersten, Distortive Effects of Dividend Taxation. 21 pp.

2013:10 Jonas Poulsen, After Apartheid: The Effects of ANC Power. 42 pp.

2013:11 Spencer Bastani, Using the Discrete Model to Derive Optimal Income Tax Rates. 14pp.

2013:12 Lovisa Persson, Consumption smoothing in a balanced budget regime. 39 pp.

2013:13 Karin Edmark och Roger Gordon, Taxes and the Choice of Organizational Form by Entrepreneurs in Sweden. 48 pp.

2013:14 Linuz Aggeborn, Voter Turnout and the Size of Government. 48pp.

2013:15 Michihito Ando, How Much Should We Trust Regression-Kink-Design Estimates? 69 pp.

2014:1 Sören Blomquist, Anil Kumar, Che-Yuan Liang and Whitney K. Newey, Individual Heterogeneity, Nonlinear Budget Sets, and Taxable Income. 37 pp.

2014:2 Oscar Erixson and Henry Ohlsson, Estate division: Equal sharing as choice, social norm, and legal requirement. 45 pp.

2014:3 Oscar Erixson, Health responses to a wealth shock: Evidence from a Swedish tax reform. 44 pp.

2014:4 Che-Yuan Liang, Distribution-Free Structural Estimation with Nonlinear Budget Sets. 31 pp.