ANNUAL REVIEWS

Annual Review of Economics

The Econometrics of Nonlinear Budget Sets

Sören Blomquist,[1] Jerry A. Hausman,[2,3] and Whitney K. Newey[2,3]

[1]Uppsala Center for Fiscal Studies, Department of Economics, Uppsala University, Uppsala, Sweden

[2]Department of Economics, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA; email: wnewey@mit.edu

[3]National Bureau of Economic Research, Cambridge, Massachusetts, USA

Annu. Rev. Econ. 2023. 15:287–306

First published as a Review in Advance on April 11, 2023

The *Annual Review of Economics* is online at economics.annualreviews.org

https://doi.org/10.1146/annurev-economics-082222-065323

JEL codes: C14, C24, H31, H34, J22

Keywords

random utility model, budget set regression, bunching, instrumental variables

Abstract

This article surveys the development of nonparametric models and methods for estimation of choice models with nonlinear budget sets. The discussion focuses on the budget set regression, that is, the conditional expectation of a choice variable given the budget set. Utility maximization in a nonparametric model with general heterogeneity reduces the curse of dimensionality in this regression. Empirical results using this regression are different from maximum likelihood and give informative inference. The article also considers the information provided by kink probabilities for nonparametric utility with general heterogeneity. Instrumental variable estimation and the evidence it provides of heterogeneity in preferences are also discussed.

1. INTRODUCTION

There are important individual choices whereby prices vary with individual purchases and budget sets are nonlinear. A leading set of examples are labor supply and the choice of how much income to earn. With progressive taxation marginal, after-tax wages depend on the tax law and total income earned. These examples are very pertinent to economists, with tax policy being an important issue. Other labor supply examples also exist. Transfer payments by governments are often tied to higher marginal tax rates. Social security benefits, disability benefits, Aid to Families with Dependent Children (AFDC) payments, and food stamps all have high marginal tax rates associated with them that create nonlinearities in the budget sets. Electricity prices are another example of nonlinear budget sets. Many electric utilities engage in declining block pricing, wherein the marginal price of successive kilowatt hours of electricity decrease at a given number of points determined by total demand. Thus, the budget set is piecewise linear.

Figure 1 illustrates the case of a budget set that could arise with progressive income taxation. We consider a two-good case where taxable income is given on the horizontal axis and consumption of the composite good is given on the vertical axis. In this simplified example, the consumer is faced with four distinct tax rates, with high tax rate (low slope) at low income levels, when the intercept of nonlabor income includes government support, followed by a sharp reduction in tax rate and then increasing marginal rates (decreasing slopes). Here each consumer chooses their taxable income to maximize utility so that variation in taxable incomes results from heterogeneity of preferences and possibly measurement and/or optimization errors, even if there were just a single budget set. This is a random utility model where the word "random" refers to variation in preferences across individuals.

Other kinds of nonlinearities are also present in consumer choices. Intended utilization of a durable good will often depend on the cost per unit of output; for example, the price per mile driven depends on the fuel efficiency of the car being used. When consumers face the joint decision

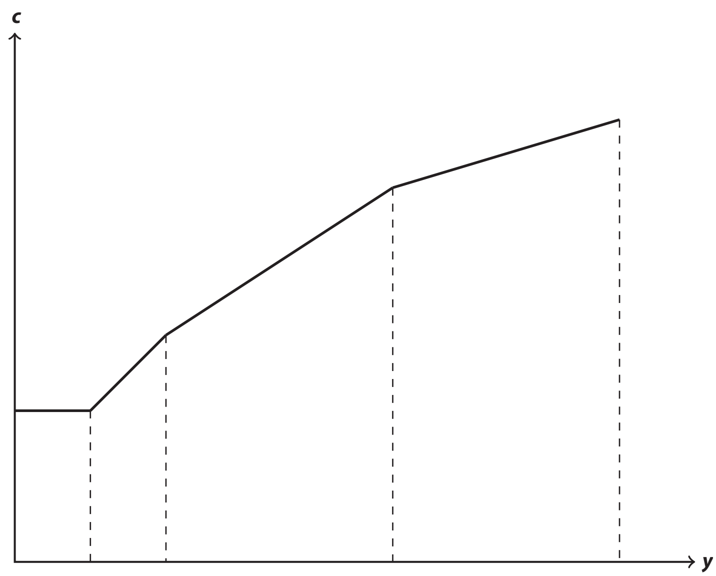

Figure 1

Nonlinear budget set. The figure illustrates a two-good case; taxable income is given on the horizontal axis and consumption of the composite good is given on the vertical axis.

of purchasing the durable along with its intended utilization, they face a nonlinear budget set that is further complicated because the purchase price of the durable good usually varies with its utilization efficiency as well as other features. Two-part tariffs also give nonlinear budget sets. Here a consumer can subscribe to a different type of service where a fixed payment is levied and the price of a unit of demand varies accordingly to the type of service chosen. Telephone service has this character in the United States, where a monthly fee is charged and a given price is charged per local call after an initial allowance of free calls is surpassed. Electricity demand may also have this characteristic when more than one type of service is offered with different rate schedules. The focus of this article is on choices that can be well represented as being over a single good with a nonlinear budget set. Taxable income and labor supply are examples of this type of choice that are very important and motivate this review.

The development and use of nonparametric models and methods are important innovations in econometrics. These methods allow one to avoid errors that can arise from functional form misspecification. There are limits to these methods associated with the so-called curse of dimensionality, which refers to the potential inaccuracy of nonparametric estimation of functions of many variables. Modeling and estimation of choice with nonlinear budget sets is a potentially important area for nonparametric approaches. The models are quite nonlinear, with results potentially depending on the distribution of preferences and measurement and/or optimization errors. Nonparametrics can help avoid errors resulting from incorrect specification of distributions and functional forms. Also, utility maximization can help avoid the curse of dimensionality, as we will discuss. This article focuses mainly on the development and use of nonparametrics for estimation with nonlinear budget sets. We consider the implications of utility maximization with nonparametric utilities and general heterogeneity across individuals—i.e., of a nonparametric random utility model (RUM) like those of McFadden & Richter (1991), Blomquist & Newey (2002), McFadden (2005), Blomquist et al. (2014, 2015), Blundell et al. (2014), Manski (2014), Hausman & Newey (2016), Kline & Tartari (2016), and Kitamura & Stoye (2018).

A particular focus of this article is the budget set regression that is the conditional expectation of an outcome variable that is either labor supply or taxable income, given the budget set, under RUM. This regression is the average over preferences and measurement and/or optimization errors of the outcome given the budget set. This regression can be used to identify policy effects that are changes in averages resulting from changes within the range of budget sets observed in the data. This regression is subject to a potentially severe curse of dimensionality because the conditioning variable—i.e., the budget set—is a high-dimensional object. Blomquist & Newey (2002) showed that with scalar, monotonic heterogeneity this regression is only a three-dimensional function, and Blomquist et al. (2014, 2015) showed that the same results hold for the RUM model with general heterogeneity when the budget set is piecewise linear and convex. To the best of our knowledge, that makes the tax policy estimates of Blomquist & Newey (2002) the first that are valid for the RUM. Using this implication of the RUM means that the curse of dimensionality can be avoided for the budget set regression. Also, some nonconvexities can be allowed without increasing the dimension too much, making the budget set regression potentially useful for estimating the effects of policy changes. Furthermore, it is possible to check the restrictions implied by the RUM using the budget set regression.

The endogeneity of budget sets is a potentially important specification concern for the budget set regression and for earlier parametric models of labor supply and taxable income. Blomquist et al. (2014, 2015) showed, and we discuss here, that control variables can be used to estimate budget set regressions with endogeneity in the RUM, as shown in earlier versions of the model by Hausman & Newey (2016) and Kitamura & Stoye (2018). Also, instrumental variable methods can be used to control for endogeneity in the linear instrumental variable (IV) models common in the taxable income literature (e.g., Gruber & Saez 2002, Blomquist & Selin 2010, Burns & Ziliak 2017, Kumar & Liang 2020). Although the choice models here are much more restrictive than the RUM, the ability to account for endogeneity is a very attractive property. In addition, models of taxable income must control for productivity growth in data from different time periods. Productivity growth is an important determinant of changes in taxable income over time. Blomquist et al. (2015) showed how productivity growth that varies over individuals according to their effort can be controlled for, as discussed here.

The RUM also helps clarify what can be learned from bunching, as proposed by Saez (2010), Chetty et al. (2011), and Kleven & Waseem (2013). Blomquist et al. (2015) showed that in the RUM the probability of a kink is a weighted average over the heterogeneity of a compensated tax effect for individuals located at the kink. This weighted average is not identified because it depends on the preferences of individuals located at the kink. McCallum & Seegert (2017) gave identification results when covariates are present and preferences have a Gaussian distribution. Blomquist & Newey (2017) showed that for a parametric, isoelastic utility function, the distribution of taxable income for one convex budget set provides no information about the elasticity and provided identification bounds. These results are presented by Blomquist et al. (2021). Bertanha et al. (2017) considered nonidentification results and gave bounds based on variability of the preference density. This article describes the RUM results of Blomquist et al. (2015).

Beginning with Feldstein (1995, 1999) and continuing with Gruber & Saez (2002), Blomquist & Selin (2010), and Burns & Ziliak (2017), the focus of the econometrics of nonlinear budget sets shifted to IV estimation of preference parameters, especially taxable income elasticities. A recent contribution is by Kumar & Liang (2020), who used IV to estimate different weighted averages of taxable income elasticities and thus provided convincing evidence of important preference heterogeneity. We discuss some of this work in this article. The evidence of preference heterogeneity in this work motivates the use of RUM, which allows for general heterogeneity and nonparametric utilities.

Section 2 describes some of the contributions of the pioneering work on the econometrics of nonlinear budget sets and the sensitivity to parametric assumptions by Blomquist & Newey (2002) and Kumar (2008). Section 3 gives the budget set regression under RUM and outlines how it can be used to estimate interesting policy effects. Section 4 considers endogeneity of the budget set and how the budget set regression may be used with endogeneity. Section 5 gives a RUM formula for kink probabilities and considers partial identification from kinks. Section 6 considers what can be learned from IV with partially nonparametric models. Section 7 offers some conclusions and points to open problems.

This article cites theorems for the RUM with nonlinear budget sets by Blomquist et al. (2015) as propositions to help provide a coherent account of econometrics for nonlinear budget sets. These results are not original to this article but are provided and discussed here for expositional purposes. For simplicity, we do not state the exact conditions that are sufficient for these results, which can be found in the original paper by Blomquist et al. (2015).

2. PARAMETRIC MODELS

Early econometrics for nonlinear budget sets was based on using parametric models of preferences and distributions to specify the distribution of labor supply given budget sets and other observable variables. Maximum likelihood was used to estimate the parameters of these models, and simulation methods based on the specified distributions were used to predict the effects of tax changes. This literature includes works by Burtless & Hausman (1978), Wales & Woodland (1979), Hausman (1981), and Blomquist (1983). Some of the models included multidimensional preference heterogeneity, including the random income effect specifications of Burtless & Hausman (1978), Hausman (1981), and Blomquist (1983). These papers are important precursors to RUM specifications in emphasizing the theoretical and empirical importance of multidimensional preference heterogeneity and measurement and/or optimization errors.

The earliest work was focused on budget frontiers that were piecewise linear and concave. Hausman (1979) gave a general solution procedure for individual choice in these models. This procedure was applied to labor supply by Hausman (1981) in a subsequent paper in which nonconvexities were also allowed. Later work on parametric models further developed approaches for nonconvex budget sets. Hausman (1985) formulated parametric models where indirect utility functions determined the choices between different budget segments, and the choice within a segment followed from utility maximization over the segment. MaCurdy et al. (1990) contributed to this literature by using smooth budget set approximations for estimation. Eklöf & Sacklen (2000) and Keane (2011) showed that differences between the estimates of Burtless & Hausman (1978) and MaCurdy et al. (1990) could be attributed to differences in variable definitions between the two papers. Blundell et al. (1998) used IV estimation combined with ignoring data near a kink and correcting for the resulting sample selection.

Dagsvik (1994), Aaberge et al. (1995), van Soest (1995), and Keane & Moffitt (1998) developed parametric models in which hours are restricted to a finite set of points. These models constitute discrete choice models that are estimated by maximum likelihood. Variants of these models are widely used in policy analysis. Blundell & Shephard (2012) and Beffy et al. (2019) provide more recent studies using this approach along with innovative preference specifications. An advantage of these types of models is that they can easily handle complicated, sometimes nonconvex, budget constraints. A disadvantage is that they do not allow for measurement errors in hours of work. A potentially serious consequence of this is illustrated by the results presented by van Soest et al. (2002), who perform Monte Carlo simulations for two values of the standard deviation of an additive measurement error in hours of work. For the larger value of the standard deviation, the change in hours of work of a hypothetical tax reform is overestimated by 140%.

There are several potential sources of specification error in parametric models, as pointed out by Hausman (1985). These include (*a*) deviation between desired and actual hours due to unexpected events, (*b*) measurement error in the labor supply variable, (*c*) unobserved variation in preferences by the specification of disturbances and their distributions, and (*d*) specification error to the extent that the labor supply function or utility function for desired hours is incorrect. Because the models and estimators are quite nonlinear, the empirical conclusions could be very sensitive to any or all of these specification errors. This sensitivity motivates nonparametric models and methods like those that are the focus of this article.

Sensitivity to specification was found by Blomquist & Newey (2002) in estimating labor supply for Swedish data. That paper compared parametric estimates like those of Burtless & Hausman (1978) and Blomquist (1983) with nonparametric budget set regression estimates that, as shown by Blomquist et al. (2014, 2015), are correct for the RUM and allow measurement and/or optimization errors in desired or actual hours of work. Nonparametric estimates of wage elasticities and the effect of the Swedish tax reform from 1980 to 1991 were found to be much lower than parametric estimates, with differences that were statistically significant. Specifically, the parametric estimate of the uncompensated wage elasticity was 0.123, while, for example, the nonparametric estimate for one rich, Gaussian power series specification was 0.0819, with standard error 0.0200. Based on Hausman (1978), the standard error for the nonparametric estimate is an upper bound for the standard error of the difference, so the difference in elasticities, $0.123 - 0.0636 = 0.0596$, has a standard error bounded above by 0.0200, and hence the difference is statistically significant at conventional levels. Similarly, the parametric estimate of the effect of tax reform is a 5.46%

increase in the average labor supply, while the nonparametric estimate for the same specification as the elasticity is 3.31% with standard error 0.795, also giving a statistically significant difference. The sensitivity to specification found in this application motivates interest in nonparametric methods such as those discussed in the rest of this article.

Some differences between nonparametric and parametric estimates are also present for US data. Kumar (2008) estimated wage and income elasticities of 0.14 and −0.04, respectively, from the cross-sections of observations in the Panel Study of Income Dynamics (PSID) before and after the tax reform of 1986. Hausman (1981) estimated wage and income elasticities of 0.00 and −0.17, respectively, from earlier years of the PSID. These elasticity estimates are quite different, although the compensated wage elasticity estimates should be much closer. We are not aware of any studies that compare the nonparametric and parametric estimates using the same US data.

Manski (2014) and Kline & Tartari (2016) also gave estimates of effects of interest for labor supply based on RUM. Manski (2014) found it was difficult to obtain useful estimates of some policy effects using choice among discrete labor supply alternative without allowance for measurement and/or optimization errors. The empirical results of Blomquist & Newey (2002) cited above demonstrate that useful estimates of other interesting policy effects can be obtained for a RUM with measurement and/or optimization errors. Kline & Tartari (2016) used a more general nonparametric model of choice to partially identify interesting counterfactual effects of an experiment. Our focus is on nonparametrically estimating interesting effects of changing the budget set from more common observational data.

3. THE RANDOM UTILITY MODEL FOR TAXABLE INCOME

We consider preferences defined over after-tax income c (value of consumption) and before-tax income y (cost of effort). As pointed out by Feldstein (1995, 1999), this specification allows for individuals to have margins of choice other than hours worked, such as effort, that help determine y. For this reason, the choice of taxable income has been widely considered in the public finance literature for some time. We follow this trend and focus on RUM for taxable income choice in most of the rest of this article. The results also apply to labor supply if we consider y as corresponding to hours worked, though the form of the budget set is different in that case.

We assume that before-tax income and after-tax income are related by

$$c = y - T(y) = A(y),$$

where $T(y)$ is the amount of taxes paid when income is y and $A(y)$ is the after-tax income. The utility function of an individual will be

$$U(c, y, \eta),$$

where η is a possibly multidimensional vector representing individual preferences. We will assume throughout that for each η the utility function $U(c, y, \eta)$ is increasing in c, decreasing in y, and strictly quasi-concave. Strict quasi-concavity is equivalent to $U(R + \rho y, y, \eta)$ having a unique maximum for all positive R and ρ, that is, to utility-maximizing choice of taxable income choice being unique for any linear (affine) $A(y)$. Strict quasi-concavity is a common assumption in the applied literature.

For an individual with preferences η, we denote the choice of y that maximizes utility $U(A(y), y, \eta)$ by

$$y(A, \eta),$$

where we assume that the maximizing value exists and is unique with probability 1 in the distribution of η. In general, the choice $y(A, \eta)$ may depend on the whole after-tax function A. **Figure 1** illustrates this choice for a particular after-tax function $A(y)$. An individual with preferences η will choose the point on the budget set where their utility is highest. Different individuals may have different η and so may choose different taxable incomes. The distribution of taxable income for a given budget set comes from variation in preferences. Heterogeneity of preferences is necessary in order to have a distribution of taxable income when there is a single $A(y)$. If preferences were homogeneous, we would have one point on a single budget constraint.

An important example is the isoelastic utility function considered by Saez (2010),

$$U(c, y, \eta) = c - \frac{\eta}{1 + \frac{1}{\beta}} \left(\frac{y}{\eta}\right)^{1+1/\beta}, \ \eta > 0, \ \beta > 0, \qquad 1.$$

where η is a scalar. Maximizing this utility function subject to a linear budget constraint $A(y) = \rho y + R$ with slope (net of tax rate) ρ and intercept (nonlabor income) R gives the taxable income function $y(\rho, \eta) = \rho^{\beta}\eta$. The taxable income elasticity $\partial \ln y(\rho, \eta)/\partial \ln \rho = \beta$ is constant for this specification and there is no income effect of changing R. The variable η represents unobserved individual heterogeneity in preferences. We note that $y(\rho, \eta)$ is decreasing in β (by $\rho < 1$) and increasing in η and ρ. Allowing β to vary with individuals (i.e., β is included in η) is a specification with multidimensional heterogeneity. The Burtless & Hausman (1978) specification is also included as a special case where income and level effects can vary separately and η is two-dimensional. The RUM includes such specification as well as those with more general functional forms with any or all parameters varying over individuals. In the fully nonparametric specification, we allow η to be of unknown dimension. We do need to restrict η and $U(c, y, \eta)$ so that probability statements can be made, but these are technical side conditions that do not affect our interpretation of η as representing general heterogeneity and are given by Blomquist et al. (2015).

The model here is a RUM for $U(c, y, \eta)$ that is strictly quasi-concave, with strict quasi-concavity being equivalent to a single-valued choice of taxable income for any linear $A(y)$. Some of the results we discuss also impose smoothness on the taxable income function. Single-valued, smooth specifications are often used in applications, and so we adopt this assumption here. In particular, smoothness has often proven useful in applications of nonparametric models, as it will be here.

The taxable income for a linear budget set will be

$$y(\rho, R, \eta) = \arg\max_{y} U(\rho y + R, y, \eta), \ y \geq 0, \ \rho \geq 0, R \geq 0.$$

This taxable income for linear taxes has an important role in the identification and estimation results to follow. Define $F(y|\rho, R) = \Pr(y(\rho, R, \eta) \leq y)$, where the probability statement is with respect to the distribution of η. McFadden (2005) derived restrictions on $F(y|\rho, R)$ that are necessary and sufficient for a RUM. With choice over two dimensions (c and y) there is a simple, alternative characterization of the RUM. The characterization is that the cumulative distribution function (CDF) satisfies a Slutsky-like condition, referred to henceforth as the Slutsky condition. The following result holds under additional technical conditions not given here and is presented by Blomquist et al. (2015). Let $F_\rho(y|\rho, R) = \partial F(y|\rho, R)/\partial \rho$ and $F_R(y|\rho, R) = \partial F(y|\rho, R)/\partial R$ when these partial derivatives exist.

Proposition 1 (theorem 1 of Blomquist et al. 2015). For $F(y|\rho, R)$ that is continuously differentiable in ρ and R, we have

$$F_\rho(y|\rho, R) - yF_R(y|\rho, R) \leq 0. \qquad 2.$$

Also, if (*a*) for all $\rho, R > 0$, $F(y|\rho, R)$ is continuously differentiable in y, ρ, R; (*b*) the support of $F(y|\rho, R)$ is $[y_\ell, y_u]$, $\partial F(y|\rho, R)/\partial y > 0$ on (y_ℓ, y_u); and (*c*) Equation 2 is satisfied, then there is a RUM with CDF of taxable income conditional on (ρ, R) being equal to $F(y|\rho, R)$.

In the conclusion of this result, $F_\rho(y|\rho, R) - yF_R(y|\rho, R)$ is a compensated distribution effect of a change in the slope ρ. In this sense, for two goods and single-valued smooth demands, the revealed stochastic preference conditions are that this compensated distribution effect is nonpositive. The compensated effect on taxable income will be positive because utility is decreasing in taxable income (i.e., effort and/or hours of work), leading to a negative compensated effect on the CDF (because the distribution of taxable income is shifted to the right).

This result will be useful in Section 5 and is of interest in its own right. Dette et al. (2011) showed that each quantile of $y(\rho, R, \eta)$ satisfies the Slutsky condition for demand functions under similar conditions. Hausman & Newey (2016) observed that when a quantile function satisfies the Slutsky condition there is always a demand model with that quantile function, and so they gave a quantile version of Proposition 1. Proposition 1 is essentially the result of Hausman & Newey (2016) combined with the inverse function theorem. The fact that scalar heterogeneity is sufficient for the RUM model with two goods may explain why Blomquist & Newey (2002) obtained the dimension reduction provided by the RUM using only scalar heterogeneity.

For taxable income, productivity growth can be incorporated by specifying that the utility function in time period t is $U(c, y/\phi(t), \eta)$, where $\phi(t)$ is a relative productivity in the t^{th} time period. As shown by Blomquist et al. (2014, 2015), this specification allows for effects of productivity growth that vary with individual preferences by effort level. This specification of productivity growth leads to adjustments in the budget set regression described in Section 4.

4. BUDGET SET REGRESSIONS

The budget set regression is the expected value of taxable income Y conditional on an after-tax function $A(y)$. This regression can be used for estimating the average effect on taxable income of changes in policy effects when the regression is stable across different policy regimes. It was used to evaluate the effect of the Swedish tax reform by Blomquist & Newey (2002). In this section we describe this regression when the budget frontier is piecewise linear and continuous.

In applications, most tax systems have a finite number of rates that change at certain income values. In such cases the after-tax function $A(y)$ is piecewise linear. A piecewise linear $A(y)$ with J segments, indexed by j, can be described by a vector $(\rho_1, \ldots, \rho_J, R_1, \ldots, R_J, \ell_1, \ldots, \ell_{J-1})$ of net-of-tax rates ρ_j (slopes), virtual incomes R_j (intercepts), and kinks or notches ℓ_j, with $\ell_0 = 0$ and $\ell_J = \infty$. We can represent a piecewise linear $A(y)$ as

$$A(y) = \sum_{j=1}^{J} 1(\ell_{j-1} \le y < \ell_j)(R_j + \rho_j y),$$

where we assume that tax rates can change at $y = \ell_j$, $(j = 0, \ldots J)$.

For a continuous $A(y)$, the endpoints of the budget segments are determined by the intersection of neighboring linear segments, where $\ell_j = (R_{j+1} - R_j)/(\rho_j - \rho_{j+1})$, $(1 \le j \le J-1)$. Such an $A(y)$ is given in **Figure 1**. It is completely parameterized by the $2J + 1$ vector $\mathbf{B} = (J, \rho_1, R_1, \ldots, \rho_J, R_J,)'$. We focus first on increasing marginal tax rates and give generalizations that allow for marginal tax rates to decrease. To describe the budget set regression, recall that $F(y|\rho, R)$ is the CDF of taxable income for a linear budget set. Define

$$\bar{y}(\rho, R, F) = \int yF(\mathrm{d}y|\rho, R), \qquad 3.$$

$$\nu(\rho, R, \ell, F) = \int 1(y < \ell)(y - \ell)F(\mathrm{d}y|\rho, R), \ \lambda(\rho, R, \ell, F) = \int 1(y > \ell)(y - \ell)F(\mathrm{d}y|\rho, R).$$

The first object is the conditional mean of taxable income for a linear budget set with slope ρ and intercept R. The other objects are integrals over values of Y below or above a possible kink point ℓ. Note that all of these objects are known functions of the conditional CDF $F(y|\rho, R)$ of taxable income for a linear budget set. Thus, these objects depend only on $F(y|\rho, R)$.

The budget set regression is given in the following result. Let $\mu(\mathbf{B}) = E[Y|\mathbf{B}]$.

Proposition 2 (theorem 5 of Blomquist et al. 2015). If $\int |y(\rho, R, \eta)| G(\mathrm{d}\eta) < \infty$ for all ρ, $R > 0$, $\rho_1 > \cdots > \rho_J$, and $\mathbf{B}$ is independent of η, then we have

$$\mu(\mathbf{B}) = \bar{y}(\rho_J, R_J, F) + \sum_{j=1}^{J-1} [\nu(\rho_j, R_j, \ell_j, F) - \nu(\rho_{j+1}, R_{j+1}, \ell_j, F)] \qquad 4.$$

$$= \bar{y}(\rho_1, R_1, F) + \sum_{j=1}^{J-1} \left[\lambda(\rho_{j+1}, R_{j+1}, \ell_j, F) - \lambda(\rho_j, R_j, \ell_j, F)\right].$$

This result shows that the budget set regression depends only on a single three-dimensional function, the CDF of taxable income $F(y|\rho, R)$ for a linear budget set. Based on this feature, we give a three-dimensional series approximation in this section, which is a much smaller dimension than the $2J + 1$ dimensions of $\mathbf{B}$. In this way, utility maximization reduces the dimension of the budget set regression.

Proposition 2 generalizes theorem 2.1 of Blomquist & Newey (2002) by allowing general heterogeneity, whereas Blomquist & Newey (2002) assumed scalar η. Proposition 2 shows that the decomposition of the budget set regression into the sum of a two-dimensional and a three-dimensional function holds without the restriction of scalar heterogeneity. Consequently, the empirical conclusions drawn by Blomquist & Newey (2002) about the average labor supply effect of the 1980 to 1991 Swedish tax reform are valid under nonparametric utility and general heterogeneity, as stated in Section 2. To the best of our knowledge that makes the tax policy estimates of Blomquist & Newey (2002) the first that are valid with nonparametric preferences and general heterogeneity.

Proposition 2 differs from Blomquist & Newey's (2002) theorem in that the budget regression depends only on the single three-dimensional function $F(y|\rho, R)$. One can think of this as a cross-function restriction between the two-dimensional function $\bar{y}$ and the three-dimensional function v, with each function depending on $F(y|\rho, R)$ in a known way. This more parsimonious specification makes the budget set regression easier to estimate than in Blomquist & Newey's (2002) model.

Measurement and/or optimization errors can be allowed for by specifying that $Y^* = Y(A, \eta)$ is desired taxable income (or hours) and that actual taxable income Y departs from desired taxable income in such a way that

$$E[Y|Y^*, \mathbf{B}] = Y^* = Y(A, \eta).$$

Then, by iterated expectations and Proposition 2, we have

$$E[Y|\mathbf{B}] = E[Y^*|\mathbf{B}] = \mu(\mathbf{B}).$$

Allowing for such a departure of actual taxable income from desired taxable income is consistent with the fact that in most data sets there is no bunching at kink points of budget sets, as proposed by Burtless & Hausman (1978) and Hausman (1981). This specification also allows for random measurement error in taxable income (or hours), an important robustness feature of the taxable income regression. It is true that this specification would not generally arise from optimization errors as modeled by Chetty (2012), but optimization errors need not have that form. The type of errors allowed here is plausible in allowing random departures of taxable income (or hours) from utility maximization. Such random errors could be more plausible than errors obtained by solving a complicated optimization problem involving limits on utility departures from rationality.

It is common practice to measure behavioral effects in terms of elasticities. Often the object of interest is the elasticity with respect to the net of tax rate for a linear budget constraint. A problem with nonlinear budget constraints is that this elasticity may not be identified. The elasticity for a linear budget constraint would often be thought of as corresponding to $\bar{y}(\rho, R)$. As discussed by Blomquist et al. (2015), stringent conditions are required to identify this function. The authors show that the set of net of tax rates where this function is identified can be very small, even empty. Furthermore, since everyone generally faces a nonlinear budget set, and policy changes are not likely to eliminate this nonlinearity, effects from changes in nonlinear budgets are important objects of interest.

The budget set regression is $\mu(\mathbf{B}) = \mu(J, \rho_1, \ldots, \rho_J, R_1, \ldots, R_J)$. We consider the effect on expected taxable income of an increase in all slopes of the budget set segments. When $\mu(\mathbf{B})$ is a differentiable function of the slopes, this object is

$$\theta_{\rho 0} = E\left[\frac{\partial \mu(J, \rho_1 + a, \ldots, \rho_J + a, R_1, \ldots, R_J)}{\partial a}\right] = E\left[\sum_{j=1}^{J} \frac{\partial \mu(J, \rho_1, \ldots, \rho_J, R_1, \ldots, R_J)}{\partial \rho_j}\right]$$

by the chain rule. It is plausible that this object could be identified from variation in the overall slopes in the data. Such variation could arise from changes in overall tax rates. Similarly, one could consider a nonlabor income effect that is the average effect of increasing each R_j:

$$\theta_{R0} = \sum_{j=1}^{J} \frac{\partial \mu(J, \rho_1, \ldots, \rho_J, R_1, \ldots, R_J)}{\partial R_j}.$$

The $\theta_{\rho 0}$ is the average effect of rotating the budget frontier and θ_{R0} is the average effect of shifting the budget frontier vertically, both while holding fixed the kink points. For policy purposes, $\theta_{\rho 0}$ is like a change in a proportional tax rate, and θ_{R0} is like a change in unearned income. Identification of $\theta_{\rho 0}$ and θ_{R0} only requires variation in the overall slopes and intercepts of the budget constraint across individuals and time periods. This is a common source of variation in nonlinear budget sets due to variations in overall degree of taxation and in nonlabor income, so it is plausible that the effects of such changes should be identified.

One can also specify an elasticity object corresponding to $\theta_{\rho 0}$ by multiplying θ_{ρ} by a constant $\tilde{\rho}$ that summarizes the vector of net-of-tax rates in a single number and then dividing by $\mu(\mathbf{B})$ to obtain

$$e_{\rho} = \tilde{\rho}\theta_{\rho 0}/\mu(\mathbf{B}).$$

The construction of $\tilde{\rho}$ can be done in many different ways. We use the averages of the net-of-tax rates and virtual incomes for the segments where individuals are actually located. The elasticity e_{ρ} is an aggregate elasticity, which is a policy-relevant effect, as argued by Saez et al. (2012). For nonlabor income, we stick with θ_{R0} as an income effect because there is more ambiguity in how to construct an income elasticity.

An important feature of Proposition 2 is that the budget set satisfies more restrictions than those identified by Blomquist & Newey (2002), as noted following Proposition 2. As a result, the budget set regression can be approximated with fewer regressors. This improvement can be carried out by specifying a linear in parameters approximation to the conditional CDF and then applying the known transformations in Proposition 2 to obtain an approximation for the budget set regression. To do so, let M be a positive integer and $F_1(y), \ldots, F_M(y)$ be CDFs chosen by the researcher. For example, these CDFs could be chosen to give more weight to different intervals of taxable income values. They are used here to approximate the counterfactual distribution of taxable income given a linear budget set, but they could be based on the observed taxable income.

In particular, $F_1(y)$ could be the empirical CDF of taxable income in the data, and $F_m(y)$ could be empirical CDFs for upper and lower parts of the data. The use of data in this way will not bias the budget set regression because all that is required to estimate the regression is a sufficiently rich approximation.

Also let $x := (\rho, R)$ and let $r_1(x), \ldots, r_K(x)$ denote approximating functions such as splines or power series. For example, a linear approximation could specify

$$(r_1(x), \ldots, r_K(x))' = (1, \rho, R).$$

Let β_{mk}, $(m = 2, \ldots, M;\ k = 1, \ldots, K)$, be coefficients of a series approximation to be specified below, and let $w_m(x, \beta) = \sum_{k=1}^{K} \beta_{mk} r_k(x)$. An approximation to the conditional CDF can be constructed as

$$\begin{aligned} F(y|x, \beta, K, M) &= F_1(y) + \sum_{m=2}^{M} w_m(x, \beta)[F_m(y) - F_1(y)] \qquad 5. \\ &= \sum_{m=1}^{M} w_m(x, \beta) F_m(y),\ w_1(x, \beta) = 1 - \sum_{m=2}^{M} w_m(x, \beta). \end{aligned}$$

This could be thought of as a mixture approximation to the conditional CDF with weights $w_m(x, \beta), (m = 1, \ldots, M)$, that are linear in parameters. The normalization that the weights sum to 1 is imposed by the formula for $w_1(x, \beta)$. This leads to a CDF approximation having the important property that for all β, we have

$$\lim_{y \longrightarrow \infty} F(y|x, \beta, K, M) = 1.$$

The conditional probability density function (pdf) could also be restricted to be nonnegative by requiring that $w_m(x, \beta) \geq 0$ on a grid of values for x. For computational simplicity we do not require that this nonnegativity restriction hold. We are primarily interested in approximating the budget set regression, and the CDF approximation in Equation 5 gives a linear in parameters approximation to the budget set that can be estimated by linear regressions. Imposing nonnegative weights would complicate the estimation of the budget set regression.

We obtain an approximation to the budget set regression by plugging the CDF approximation into the respective formulas for $\bar{y}(\rho, R)$ and $\nu(\rho, R, \ell)$ and then plugging these into the formula for conditional mean in Proposition 2. Let

$$\bar{y}_m = \int y F_m(\mathrm{d}y),\ \nu_m(\ell) = \int 1(y < \ell)(y - \ell) F_m(\mathrm{d}y),\ (m = 1, \ldots, M).$$

Substituting the CDF approximation in the expression for the conditional mean $\mu(\mathbf{B})$ in Proposition 2 gives

$$\begin{aligned} \mu(\mathbf{B}, \beta, K, M) &= \bar{y}_1 + \sum_{m=2}^{M} w_m(x_J, \beta)(\bar{y}_m - \bar{y}_1) \\ &\quad + \sum_{m=2}^{M} \sum_{j=1}^{J-1} [w_m(x_j, \beta) - w_m(x_{j+1}, \beta)][\nu_m(\ell_j) - \nu_1(\ell_j)] \\ &= \bar{y}_1 + \sum_{m=2}^{M} \sum_{k=1}^{K} \beta_{mk} p_{mk}(\mathbf{B}), \end{aligned}$$

$$p_{mk}(\mathbf{B}) = r_k(x_J)(\bar{y}_m - \bar{y}_1) + \sum_{j=1}^{J-1} [r_k(x_j) - r_k(x_{j+1})][\nu_m(\ell_j) - \nu_1(\ell_j)].$$

This has a linear regression form where the regressor with coefficient β_{mk} is $p_{mk}(\mathbf{B})$. Each such regressor is a linear combination of the k^{th} approximating function evaluated at the last segment $r_k(x_J)$ and the differences $r_k(x_j) - r_k(x_{j+1}), (j = 1, \dots, J-1)$, for adjacent segments. The dimension reduction is achieved by the regressors being linear combinations of differences across adjacent segments.

For example, with three segments and a linear approximation where $(r_1(x), \dots, r_K(x)) = (1, \rho, R)$, there are three regressors for each m given by

$$\bar{y}_m - \bar{y}_1,\ \rho_3(\bar{y}_m - \bar{y}_1) + (\rho_1 - \rho_2)[\nu_m(\ell_1) - \nu_1(\ell_1)] + (\rho_2 - \rho_3)[\nu_m(\ell_2) - \nu_1(\ell_2)],$$

$$R_3(\bar{y}_m - \bar{y}_1) + (R_1 - R_2)[\nu_m(\ell_1) - \nu_1(\ell_1)] + (R_2 - R_3)[\nu_m(\ell_2) - \nu_1(\ell_2)].$$

Also, the regressors for each m will include a constant and it is only necessary to include one constant term in the regression, so the total number of regressors in this example is $2(M-1) + 1 = 2M - 1$.

These regressors will vary through variation in the budget sets across individuals. The regressors described here are specific to individual budget sets and so vary with those budget sets. These regressors may vary across individuals by the number of segments J and by slopes and intercepts ρ_j, R_j, and hence also by ℓ_j and $v_m(\ell_j)$, $(m = 2, \dots, M)$. For each individual there will be $(M-1)(K-1)+1$ regressors, with the form of those regressors depending on the individual's budget set.

This specification of the regressors must be modified to account for productivity growth in taxable income. As discussed in Section 3, we allow for productivity growth by assuming that in period t the individuals are maximizing utility over $\tilde{Y} = Y/\phi(t)$ for a utility function of the form $U(c, \tilde{y}, \eta)$, where η is drawn from the same distribution in each time period and for every individual. Consequently, each individual in time period t will behave as if each of their ρ_j was $\tilde{\rho}_j = \phi(t)\rho_j$ with corresponding kink points $\tilde{\ell}_j(t) = (R_j - R_{j+1})/(\tilde{\rho}_{j+1} - \tilde{\rho}_j) = \phi(t)^{-1}\ell_j$. Moreover, the budget set regression is that for $\tilde{Y}$, which must be multiplied by $\phi(t)$ to obtain the regression for Y. Consequently, with productivity growth the budget set regression depends on t and for $\tilde{x}(t) = (\phi(t)\rho, R)$ is given by

$$\begin{aligned}
\mu(\mathbf{B}, \beta, K, M, t) = \phi(t)\Bigg\{&\bar{y}_1 + \sum_{m=2}^{M} w_m(\tilde{x}_J(t), \beta)(\bar{y}_m - \bar{y}_1) \\
&+ \sum_{m=2}^{M}\sum_{j=1}^{J-1}[w_m(\tilde{x}_j(t), \beta) - w_m(\tilde{x}_{j+1}(t), \beta)][\nu_m(\tilde{\ell}_j(t)) - \nu_1(\tilde{\ell}_j(t))]\Bigg\} \\
= \phi(t)\bar{y}_1 &+ \sum_{m=2}^{M}\sum_{k=1}^{K}\beta_{mk}p_{mk}(\mathbf{B}, t),
\end{aligned}$$

$$p_{mk}(\mathbf{B}, t) = \phi(t)\left\{r_k(\tilde{x}_J(t))(\bar{y}_m - \bar{y}_1) + \sum_{j=1}^{J-1}[r_k(\tilde{x}_j(t)) - r_k(\tilde{x}_{j+1}(t))][\nu_m(\tilde{\ell}_j(t)) - \nu_1(\tilde{\ell}_j(t))]\right\}.$$

Here the regressor $p_{mk}(\mathbf{B}, t)$ depends on both the budget set variables $\mathbf{B}$ and the time period t. It is important to note that none of these regressors is constant over time, even when $r_k(x) = 1$ for some k. Because the conditional mean of $\tilde{Y} = Y/\phi(t)$ has been multiplied by $\phi(t)$, any regressors for $\tilde{Y}$ that are constant have been multiplied by $\phi(t)$.

The regressors here, plus additional control variables and covariates, can be combined using the automatic debiased machine learner of Chernozhukov et al. (2022) to estimate policy effects like the tax and nonlabor income effects discussed previously in this section. Lasso can be used to estimate the regression of taxable income on the budget set variables, controls, and covariates. An average derivative or difference quotient "plug-in" estimator can be constructed as the average derivative or difference quotient for this and with a bias correction term added, with cross-fitting for the plug-in and for the bias correction term (see Chernozhukov et al. 2022, section 3).

The budget set regression we have described can be theoretically justified by showing that the specification provides an approximation rate to the true regression that is uniform in individual budget sets. For simplicity we do this for specific choices of distributions $F_m(y)$ and approximating functions $r_k(x)$. We expect that this approximation result would also apply to other distributions and approximating functions. We consider $F_m(y)$ to be integrals of b-splines that are positive and normalized to integrate to 1 and $r_b(x)$ that are also splines. We also require that y and x be contained in bounded sets $\mathcal{Y}$ and $\mathcal{X}$ and that the conditional pdf of taxable income for a linear budget set $f(y|x)$ be smooth. For simplicity we state the result assuming there is no productivity growth, with the understanding that the conclusion also applies with productivity growth.

Proposition 3 (theorem 6 of Blomquist et al. 2015). If $\mathcal{Y}$ and $\mathcal{X}$ are compact, $f(y|x)$ is zero outside $\mathcal{Y} \times \mathcal{X}$ and is continuously differentiable to order s on $\mathcal{Y} \times \mathcal{X}$, and $\mathrm{d}F_m(y)/\mathrm{d}y$, $(m = 1, \ldots, M)$ and $r_k(x)$, $(k = 1, \ldots, K)$ consist of tensor product b-splines of order s on $\mathcal{Y} \times \mathcal{X}$, then there exist a constant C and β_{km} such that for all piecewise linear, concave budget frontiers with $x_j \in \mathcal{X}$, $(j = 1, \ldots, J)$, we have

$$\left| \mu(\mathbf{B}) - \bar{y}_1 - \sum_{m=2}^{M} \sum_{k=1}^{K} \beta_{mk} p_{mk}(\mathbf{B}) \right| \leq CM^{1-s} K^{(1-s)/2}.$$

This result shows that the series approximation described here does approximate the expected value of taxable income for piecewise linear budget sets with increasing marginal tax rates. We emphasize that the approximation rate is uniform in the number J of budget segments and other aspects of the budget set under the conditions given here. Thus, budget sets may vary over individuals in quite general ways while maintaining low bias from using the budget set regressions given here. The rate of approximation corresponds to a multivariate b-spline approximation to a function and its derivative, where approximating the derivative is useful for making the rate uniform in the number of budget segments.

We can also make allowance for nonconcave budget frontiers, similar to what Blomquist & Newey (2002) do. For example, suppose that the budget frontier is nonconcave over only two segments. As shown by Blomquist et al. (2015), the distribution over those segments will depend only on the slope and intercept of those two segments. Because the budget set regression is a sum of integrals over different segments, it would take the form

$$E[Y|\mathbf{B}] = \bar{y}(\rho_J, R_J, F) + \sum_{j=1}^{J-1} [\nu(\rho_j, R_j, \ell_j, F) - \nu(\rho_{j+1}, R_{j+1}, \ell_j, F)] + \varsigma(R_{\tilde{j}-1}, \rho_{\tilde{j}-1}, R_{\tilde{j}}, \rho_{\tilde{j}}),$$

where $\tilde{j}$ and $\tilde{j} - 1$ index the segments where the nonconcavities occur. The ς term represents the deviation of the mean from what it would be if the budget frontier were concave. It can be accounted for in the approximation by separately including series terms that depend just on $R_{\tilde{j}-1}$, $\rho_{\tilde{j}-1}$, $R_{\tilde{j}}$, and $\rho_{\tilde{j}}$. If the nonconcavities are small or few people have taxable income where they occur, then ς will be small and including terms to account for ς will lead to little improvement. The integration across individuals to obtain the expected value reduces the importance of nonconcavities.

5. BUDGET SET ENDOGENEITY AND PRODUCTIVITY GROWTH

Two important specification issues for the budget set regression are endogeneity and productivity growth. Endogeneity of budget sets is a potentially important problem to overcome. Nonlabor income may be endogenous due to the inclusion of spouses' income or in other ways. Also, tax rates vary across locations, creating endogeneity due to correlation between preferences for location and the choice of taxable income.

Control variables can be used to make the budget set regression robust to endogeneity in the RUM. A control variable is ξ such that $\mathbf{B}$ and η are independent conditional on ξ and the conditional support of ξ given that $\mathbf{B}$ equals the marginal support of ξ. In that case it follows, as shown by Blundell & Powell (2006) and Hausman & Newey (2016), that

$$\int E[Y|\mathbf{B}, \xi]F_\xi(\mathrm{d}\xi) = \mu(\mathbf{B}),$$

where $F_\xi(\xi)$ is the marginal CDF of ξ. This integral can be estimated by letting the functions $r_b(x, \xi)$ depend on ξ and letting the regressor corresponding to coefficient β_{mk} be

$$p_{mk}(\mathbf{B}, \xi, t) = \phi(t)\left\{r_k(\tilde{x}_J(t), \xi)(\bar{y}_m - \bar{y}_1) + \sum_{j=1}^{J-1}[r_k(\tilde{x}_j(t), \xi) - r_k(\tilde{x}_{j+1}(t), \xi)][\nu_m(\tilde{\ell}_j(t)) - \nu_1(\tilde{\ell}_j(t))]\right\}.$$

A budget set regression with endogeneity controls is then obtained from regressing taxable income on these budget set regressors to obtain coefficients β_{mk}, $(m = 1, \ldots, M; k = 1, \ldots, K)$. Integrating over the control variable then gives an approximation to the average taxable income function obtained by integrating over the marginal distribution of preferences, that is,

$$\mu(\mathbf{B}) \approx \phi(t)\bar{y}_1 + \sum_{m=2}^{M}\sum_{k=1}^{K}\beta_{mk}\bar{p}_{mk}(\mathbf{B}, t),\ \bar{p}_{mk}(\mathbf{B}, t) = \int p_{mk}(\mathbf{B}, \xi, t)F_\xi(\mathrm{d}\xi).$$

Although conditions for the existence of a control variable are quite strong (see Blundell & Matzkin 2014), this approach does provide a way to allow for some forms of endogeneity.

One can use this approach to control for endogeneity of nonlabor income by including a control variable that is a residual from the regression of nonlabor income on an instrument and covariates. Blomquist et al. (2015) did so using an instrument represented by benefits at zero spouse income. To obtain precise estimates it may be difficult to fully interact this control variable with the budget set regressors, suggesting the possible importance for practice of more parsimonious specifications. Similarly, to control for endogeneity of location one could estimate a separate budget set regression for each locality. Generally one will not have enough data to make this approach practical, so one could just include location dummies to control in part for endogenous location.

There may be other sources of endogeneity that are not amenable to using control variables. In US data, one possible problem arises from the choice of whether to own or rent a home, which has large effects on budget sets. If that choice is unrelated to the choice of labor supply or taxable income this will not be a problem, but it could be otherwise. It may be that a control variable could be found here that has not been investigated. In other settings, the possible endogeneity of home ownership may not be a problem. For example, in Swedish data the choice of owning or renting only affects the lowest segment of the budget set and so is not relevant for the choice of taxable income for many individuals. In this case one would expect little effect of endogeneity on budget set regression estimators.

Productivity growth is another important specification issue when the outcome Y is taxable income. It may be possible to identify the productivity growth term $\phi(t)$ from the budget set regression for simple specifications, such as a growth rate that is constant across individuals and over time, as discussed by Blomquist et al. (2015), where plausible empirical results are obtained using this specification. More general specifications would seem to require the use of auxiliary information such as estimates of productivity growth from data other than those used by the budget set regression. An advantage of some IV specifications is that they allow for individual specific productivity growth by including individual specific constants for differenced panel data.

6. BUNCHING

There has been work on identification and estimation of taxable income effects from bunching at or near a kink. In this section we describe the implications of the RUM for bunching at a kink and how that affects identification of average compensated elasticities. To simplify the analysis and exposition, we do not treat the complications arising from optimization/measurement errors. Presence of such errors makes it even harder to deduce anything about the taxable income elasticity from bunching at a kink.

There is a RUM formula for the probability of a kink that helps to clarify what can be learned from a kink probability. Consider a budget set with kink K and slopes $\rho_1 > \rho_2$ of $A(y)$ at K from the left and right, respectively. Consider ρ with $\rho_2 \leq \rho \leq \rho_1$ and let $R(\rho) = R_1 + K(\rho_1 - \rho)$ be the virtual income for the linear budget set with slope ρ passing through the kink. Assuming that $y(\rho, R(\rho), \eta)$ is continuously distributed as η varies, let $f(y|\rho, R(\rho)) = \partial F(y|\rho, R(\rho))/\partial y$ denote its pdf and define

$$\beta(\rho, K) = E\left[\frac{\rho}{K}\left\{\frac{\partial y(\rho, R, \eta)}{\partial \rho} - K\frac{\partial y(\rho, R, \eta)}{\partial R}\right\}\bigg| y(\rho, R, \eta) = K\right]_{R=R(\rho)}, \ \tilde{\phi}(\rho) = \frac{K}{\rho} f(K|\rho, R(\rho)),$$

where the expectation is taken over the distribution of η and existence of derivatives is assumed. The $\beta(\rho, K)$ is the average compensated taxable income elasticity for those individuals facing a linear $A(y)$ with slope ρ that passes through the kink point who choose K. The $\tilde{\phi}(\rho)$ is a multiple of the pdf of individuals who would choose K for the linear budget set. The following result gives a formula for the kink probability $P_K = \Pr(Y = K)$ in terms of $\beta(\rho, K)$ and $\tilde{\phi}(\rho)$.

Proposition 4 (theorem 4 of Blomquist et al. 2015). The probability of a kink is

$$P_K = \int_{\rho_2}^{\rho_1} \tilde{\phi}(\rho)\beta(\rho, K)\mathrm{d}\rho \text{ and } \tilde{\phi}(\rho)\beta(\rho, K) = -F_\rho(K|\rho, R(\rho)) + K \cdot F_R(K|\rho, R(\rho)). \qquad 6.$$

The compensated elasticity appears in the formula because virtual income is being adjusted as ρ changes to stay at the kink. The virtual income adjustment needed to remain at the kink corresponds locally to the income adjustment needed to remain on the same indifference curve, as shown by Saez (2010). The formula for P_K in Proposition 4 bears some resemblance to the kink probability formulas of Saez (2010) but differs in important ways. Proposition 4 is global, is nonparametric, takes explicit account of general heterogeneity, and allows for nonlabor income effects, while the results of Saez (2010) are local or parametric, account for heterogeneity implicitly, and do not allow for the effects of nonlabor income.

This result helps clarify what can be nonparametrically learned from kinks. First, the compensated effects that enter P_K are only for individuals (i.e., values of η) who would choose to locate at the kink for $A(y) = R(\rho) + \rho y$, with $\rho \in [\rho_2, \rho_1]$. Thus, using kinks to provide information about compensated effects is subject to the same issues of external validity as, say, the regression discontinuity design (RDD). Just as RDD only identifies treatment effects for individuals at the jump point, so kinks only provide information about compensated effects for individuals who would locate at the kink.

Second, the kink probability depends both on the average compensated elasticity $\beta(\rho, K)$ and on a pure heterogeneity term $\tilde{\phi}(\rho)$. Intuitively, a kink probability could be large because the elasticities are large or because preferences are distributed in such a way that many like to be at the kink, that is, so that $\tilde{\phi}(\rho)$ is large. Consequently, it is not possible to separately identify compensated tax effects and heterogeneity effects from a kink. For example, consider the weighted average elasticity

$$\bar{\beta}(K) = \frac{\int_{\rho_2}^{\rho_1} \tilde{\phi}(\rho)\beta(\rho, K)\mathrm{d}\rho}{\int_{\rho_2}^{\rho_1} \tilde{\phi}(\rho)\mathrm{d}\rho} = \frac{P_K}{\int_{\rho_2}^{\rho_1} \tilde{\phi}(\rho)\mathrm{d}\rho}.$$

Evidently, $\bar{\beta}(K)$ depends on the denominator $\int_{\rho_2}^{\rho_1} \tilde{\phi}(\rho)\mathrm{d}\rho$, which is needed to normalize so that $\bar{\beta}(K)$ is a weighted average of elasticities. This denominator is not identified because no $\rho \in [\rho_2, \rho_1]$ is observed in the data. Indeed, $\tilde{\phi}(\rho)$ can be any positive function over the interval, so the denominator can vary between 0 and ∞, meaning that any $\bar{\beta}(K) \in (0, \infty)$ is consistent with the data. This result is analogous to that of Blomquist et al. (2021) for isoelastic utility, showing that a single budget set is not informative about the taxable income elasticity. Similarly, P_K is not informative about the size of $\bar{\beta}(K)$ for the nonparametric model with general heterogeneity.

Third, nonparametric restrictions on $\tilde{\phi}(\rho)$ can make P_K informative about $\bar{\beta}(K)$. For example, if the slope of the budget set is ρ_1 on an open interval to the left of K, and $f(y|\rho_1, R(\rho_1))$ is continuous in y at K, then we have $\phi(\rho_1) = \lim_{y \uparrow K} f(y|\rho_1, R(\rho_1))$, so that $\phi(\rho_1)$ is identified. Similarly, $\phi(\rho_2)$ is identified if the slope of the budget set is ρ_2 on an interval to the right of K and $f(y|\rho_2, R(\rho_2))$ is continuous at K. If, in addition, $\phi(\rho)$ is monotonic on $[\rho_2, \rho_1]$, then we have $\max\{\phi(\rho_1), \phi(\rho_2)\} \le \phi(\rho) \le \max\{\phi(\rho_1), \phi(\rho_2)\}$, and hence there are identified upper and lower bounds to $\phi(\rho)$ on the interval $[\rho_2, \rho_1]$. Substituting these bounds in the denominator and integrating gives

$$P_K/[K \ln(\rho_1/\rho_2) \max\{\phi(\rho_1), \phi(\rho_2)\}] \le \bar{\beta}(K) \le P_K/[K \ln(\rho_1/\rho_2) \min\{\phi(\rho_1), \phi(\rho_2)\}].$$

Of course, such a bound depends on the monotonicity of $\phi(\rho)$, which generally may not be credible a priori. Also, in practice kinks often do not have positive probability due to measurement and/or optimization errors, making it difficult to estimate P_K.

Proposition 4 can be used to relate positivity of the kink probability P_K to the Slutsky condition. We see from the integral formula for P_K and the second equality that the Slutsky condition is sufficient for $P_K \ge 0$. However, the Slutsky condition is not necessary for $P_K \ge 0$. The integral of $\tilde{\phi}(\rho)\beta(\rho, K)$ may be positive even when $\beta(\rho, K)$ is negative over some interval. Thus in the RUM setting the Slutsky condition is sufficient but not necessary for a kink probability to be nonnegative. Blomquist (1995) showed this result in a parametric likelihood setting. Blomquist et al. (2015) extend the nonparametric analysis to show that a coherent nonparametric model, one with a positive pdf and kink probabilities, can be constructed without imposing all the conditions of utility maximization. In particular, the distribution of taxable income implied by a particular $F(y|\rho, R)$ consistent with the RUM and a smooth, concave budget frontier can be coherent without the Slutsky condition being satisfied.

An important feature of identification is that variation in the after-tax schedule across time and place will generally not suffice to identify elasticities from kink probabilities. In particular, if different kinks always have different individuals, then kinks provide no information about how a change in the budget set affects any individual. Furthermore, Proposition 4 clearly shows that the kink is a weighted average of elasticities over different individuals, so kinks can provide at most partial identification of elasticities from budget set changes. The order condition for identification posed by Blomquist et al. (2021) also suggests that it is impossible to identify a taxable income elasticity from the effect of budget set changes on bunching. A finite number of kinks will not suffice to identify a taxable income elasticity, because the distribution of preferences is an infinite dimensional nuisance parameter.

7. INSTRUMENTAL VARIABLES ESTIMATION

IV estimation can provide useful estimates of preference parameters. For example, there is a large literature on IV estimation of the taxable income elasticity, beginning with Feldstein (1995, 1999). This literature includes the works of Gruber & Saez (2002), Blomquist & Selin (2010), and Burns & Ziliak (2017), among others. IV seems less useful for estimating policy effects, like the effect of changes in the tax schedule on taxable income, but information about preferences is important. A recent example is provided by Kumar & Liang (2020) in regard to the IV estimation of weighted averages of heterogeneous taxable income elasticities. In this section we summarize those interesting results and discuss the use of IV in relation to the RUM.

Kumar & Liang (2020) allow the elasticity of taxable income (ETI) to vary across individuals and propose potentially valid instruments. The model they consider is the isoelastic utility model of Equation 1, where β varies with individuals. In that model the choice of taxable income for a linear budget set is given by

$$\ln Y(\rho, R, \eta) = \beta(\eta) \ln \rho + \eta_1,$$

where $\beta(\eta)$ is the ETI for individual η. This model allows for heterogeneity of the ETI across individuals, though it is not the full RUM.

Kumar & Liang (2020) use IV constructed from the net of tax rate at fixed values of y, that is, they use $\mathrm{d}A(y_k)/\mathrm{d}y$ at fixed y_k for a set $\{y_1, \ldots, y_K\}$ of fixed values of taxable income. They use panel data IV where the left-hand side variable is the difference of Y across time periods, the right-hand side variable is the difference of $\ln(\mathrm{d}A(Y)/\mathrm{d}y)$ across the same time periods, and the IV is the difference of fixed income net of tax rates across the same time periods. The differencing eliminates any fixed effect in η_2 that is allowed to be correlated with the instrumental variables while also controlling for a common time trend that would generally be present due to productivity growth. They show that, under certain plausible monotonicity conditions, IV estimates weighted averages of η_1, with weights that depend on the choice of instrumental variable.

Using the NBER tax panel for 1979–1990, they estimate a weighted average ETI of 0.57 when using all available variation in budget set changes provided by the tax reforms in the data. This estimate represents a weighted average of local average ETI estimates from single fixed-dollar instruments ranging between 0.28 and 1.48. These estimates correspond to different weighted averages of the ETI, so the results indicate important economic variation in individual taxable income elasticities. The standard errors of their ETI estimators are quite small, suggesting that variation in elasticities across individuals is highly significant statistically as well as economically. Their finding of large preference heterogeneity motivates the use of RUM in estimating policy effects.

The budget set exogeneity assumptions of Kumar & Liang (2020) and most of the taxable income IV literature are weaker than those for the budget set regression. Most of the taxable income literature relies on panel data to estimate taxable income effects and on linear models. The linear model and panel data allow for individual additive effects that are endogenous and are "differenced out" in the panel data estimation. The budget set regression is nonlinear, and so additive individual effects are not a natural feature of that regression.

The labor supply and taxable income literatures have used quite different econometrics. The labor supply literature focused on nonlinear models is based on utility maximization that takes careful account of the nonlinearity of the budget sets. Starting with Feldstein (1995, 1999), the taxable income literature focused on estimation of linear models. IV econometrics is simpler but is not fully consistent with utility maximization. Typically, the net of tax rate used as the right-hand side variable for IV is the net of tax rate for the chosen income value. At or near a kink this is incorrect. The correct variable for those who are at a kink is the slope of the indifference curve at the kink point, which is not observed. It is true that in many data sets positive kink probabilities are not observed. Burtless & Hausman (1978) and others have explained this phenomenon by the presence of measurement/optimization errors. If taxable income of each individual is a utility-maximizing value plus an independent measurement error, then using the net of tax rate at the chosen income is incorrect. This claim is straightforward to verify in the isoelastic specification of Equation 1 with ETI that does not vary over individuals. The bunching near kinks that is observed in some data sets suggests that this misspecification could be important in some applications. It could be useful to check for sensitivity to this misspecification. To the best of our knowledge, such checks are not available. Blundell et al. (1998) did IV estimation combined with ignoring data near a kink and correcting for the resulting sample selection.

An interesting extension of previous results would be to analyze IV estimation in the RUM model. Instruments could be formed from budget set variables that are slopes at fixed income values. Smooth approximation of budget sets as done by MaCurdy et al. (1990) and use of the marginal net of tax rate at the chosen taxable income, rather than the marginal rate for the actual budget set, could partially correct the kink misspecification problem noted in the previous paragraph, though that does not allow for measurement and/or optimization errors.

Measurement error is a potentially important problem in any IV specification based on differencing panel data. Nonlabor income may have this problem, and the gross wage may also have it. Differencing the data, especially first-differencing, may cause significant errors in variables problems, as shown by Griliches & Hausman (1986). The use of IV could help with this problem, as long as the instruments were independent of the measurement error. Explicit allowance for measurement error would be useful to consider in IV estimation with nonlinear budget sets.

8. CONCLUSION

Among topics that might be considered for future research are the properties of IV estimators, the effect of costs of work (i.e., the extensive margin) on the budget set regression, and panel versions of the budget set regression. The effect of measurement error on IV could be studied. Also, it would be interesting to see if adjusting for misspecification near kinks had an effect on IV estimation. It would also be interesting to seek an interpretation of IV estimates in terms of the RUM. Furthermore, cost of work effects could change the budget set regression, and it would be good to know how this happens in the RUM context, especially for the study of groups of individuals in which many do not work. Also, panel data versions of the budget set regression could potentially control for endogeneity of the budget sets.

DISCLOSURE STATEMENT

The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

ACKNOWLEDGMENTS

The NSF provided financial support via grant 1757140. A. Kumar provided helpful comments and D. Ahimbisibwe and H. Kono excellent research assistance.

LITERATURE CITED

Aaberge R, Dagsvik J, Ström S. 1995. Labor supply responses and welfare effects of tax reforms. *Scand. J. Econ.* 97:635–59

Beffy M, Blundell R, Bozio A, Laroque G, Tô M. 2019. Labour supply and taxation with restricted choices. *J. Econom.* 211:16–46

Bertanha M, McCallum AH, Seegert N. 2017. *Better bunching, nicer notching*. Unpublished manuscript, Univ. Notre Dame, Notre Dame, IN

Blomquist S. 1983. The effect of income taxation on the labor supply of married men in Sweden. *J. Public Econ.* 22:167–97

Blomquist S. 1995. Restrictions in labor supply estimation: Is the MaCurdy critique correct? *Econ. Lett.* 47:229–35

Blomquist S, Kumar A, Liang CY, Newey WK. 2014. *Individual heterogeneity, nonlinear budget sets, and taxable income*. CEMMAP Work. Pap. CWP 21/14, Cent. Microdata Methods Pract., Inst. Fiscal Stud., London

Blomquist S, Kumar A, Liang CY, Newey WK. 2015. *Individual heterogeneity, nonlinear budget sets, and taxable income*. CEMMAP Work. Pap. CWP 21/15, Cent. Microdata Meth. Pract., Inst. Fiscal Stud., London

Blomquist S, Newey WK. 2017. *The bunching estimator cannot identify the taxable income elasticity*. NBER Work. Pap. 24136

Blomquist S, Newey WK. 2002. Nonparametric estimation with nonlinear budget sets. *Econometrica* 70:2455–80

Blomquist S, Newey WK, Kumar A, Liang CY. 2021. On bunching and identification of the taxable income elasticity. *J. Political Econ.* 129:2320–43

Blomquist S, Selin H. 2010. Hourly wage rate and taxable labor income responsiveness to changes in marginal tax rates. *J. Public Econ.* 94:878–89

Blundell RW, Duncan A, Meghir C. 1998. Estimating labor supply responses using tax reforms. *Econometrica* 66:827–61

Blundell RW, Kristensen D, Matzkin R. 2014. Bounding quantile demand functions using revealed preference inequalities. *J. Econom.* 179:112–27

Blundell RW, Matzkin R. 2014. Control functions in nonseparable simultaneous equations models. *Quant. Econ.* 5:271–95

Blundell RW, Powell JL. 2006. Endogeneity in nonparametric and semiparametric regression models. In *Advances in Economics and Econometrics: Theory and Applications*, Vol. II, ed. M Dewatripont, LP Hansen, SJ Turnovsky, pp. 312–57. Cambridge, UK: Cambridge Univ. Press

Blundell RW, Shephard A. 2012. Employment, hours of work and the optimal taxation of low-income families. *Rev. Econ. Stud.* 79:481–510

Burns SA, Ziliak JP. 2017. Identifying the elasticity of taxable income. *Econ. J.* 127:297–329

Burtless G, Hausman J. 1978. The effect of taxation on labor supply: evaluating the Gary negative income tax experiment. *J. Political Econ.* 86:1103–30

Chernozhukov V, Newey WK, Singh R. 2022. Automatic debiased machine learning of causal and structural effects. *Econometrica* 90:967–1027

Chetty R. 2012. Bounds on elasticities with optimization frictions: a synthesis of micro and macro evidence on labor supply. *Econometrica* 80:969–1018

Chetty R, Friedman J, Olsen T, Pistaferri L. 2011. Adjustment costs, firm responses, and micro versus macro labor supply elasticities: evidence from Danish tax records. *Q. J. Econ.* 126:749–804

Dagsvik J. 1994. Discrete and continuous choice, max-stable processes and independence from irrelevant attributes. *Econometrica* 62:1179–205

Dette H, Hoderlein S, Nuemeyer N. 2011. *Testing multivariate economic restrictions using quantiles: the example of Slutsky negative semidefiniteness*. Work. Pap., Boston Coll., Chestnut Hill, MA

Eklöf M, Sacklen H. 2000. The Hausman-MaCurdy controversy: Why do the results differ across studies? *J. Hum. Resourc.* 35:204–20

Feldstein M. 1995. The effects of marginal tax rates on taxable income: a panel study of the 1986 Tax Reform Act. *J. Political Econ.* 103(3):551–72

Feldstein M. 1999. Tax avoidance and the deadweight loss of the income tax. *Rev. Econ. Stat.* 81(4):674–80

Griliches Z, Hausman JA. 1986. Errors in variables in panel regression. *J. Econom.* 31:93–118
Gruber J, Saez E. 2002. The elasticity of taxable income: evidence and implications. *J. Public Econ.* 84:1–32
Hausman JA. 1978. Specification tests in econometrics. *Econometrica* 46:1251–71
Hausman JA. 1979. The econometrics of labor supply on convex budget sets. *Econ. Lett.* 3:171–74
Hausman JA. 1981. The effect of taxes on labor supply. In *How Taxes Affect Economic Behavior*, ed. HJ Aaron, JA Pechman, pp. 27–72. Washington, DC: Brook. Inst.
Hausman JA. 1985. The econometrics of nonlinear budget sets. *Econometrica* 53:1255–82
Hausman JA, Newey WK. 2016. Individual heterogeneity and average welfare. *Econometrica* 84:1225–48
Keane MP. 2011. Labor supply and taxes: a survey. *J. Econ. Lit.* 49:961–1075
Keane MP, Moffitt R. 1998. A structural model of multiple welfare program participation and labor supply. *Int. Econ. Rev.* 39:553–89
Kitamura Y, Stoye J. 2018. Nonparametric analysis of random utility models. *Econometrica* 86:1883–909
Kleven HJ, Waseem M. 2013. Using notches to uncover optimization frictions and structural elasticities: theory and evidence from Pakistan. *Q. J. Econ.* 128:669–723
Kline P, Tartari M. 2016. Bounding the labor supply responses to a randomized welfare experiment: a revealed preference approach. *Am. Econ. Rev.* 106:972–1014
Kumar A. 2008. Labor supply, deadweight loss and Tax Reform Act of 1986: a nonparametric evaluation using panel data. *J. Public Econ.* 92:236–53
Kumar A, Liang CY. 2020. Estimating taxable income responses with elasticity heterogeneity. *J. Public Econ.* 188:104209
MaCurdy T, Green D, Paarsch H. 1990. Assessing empirical approaches for analyzing taxes and labor supply. *J. Hum. Resourc.* 25:414–90
Manski C. 2014. Identification of income-leisure preferences and evaluation of income tax policy. *Quant. Econ.* 5:145–74
McCallum AH, Seegert N. 2017. *Better bunching, nicer notching*. Unpublished manuscript, Board Gov. Fed. Reserve Syst., Washington, DC
McFadden D. 2005. Revealed stochastic preference: a synthesis. *Econ. Theory* 26:245–64
McFadden D, Richter K. 1991. Stochastic rationality and revealed stochastic preference. In *Preferences, Uncertainty, and Rationality*, ed. J Chipman, D McFadden, K Richter, pp. 161–86. Boulder, CO: Westview Press
Saez E. 2010. Do taxpayers bunch at kink points? *Am. Econ. J. Econ. 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. *J. Econ. Lit.* 50:3–50
Van Soest A. 1995. Structural models of family labor supply: a discrete choice approach. *J. Hum. Resourc.* 30:63–88
Van Soest A, Das M, Gong X. 2002. A structural labour supply model with flexible preferences. *J. Econom.* 107:345–74
Wales TJ, Woodland AD. 1979. Labour supply and progressive taxes. *Rev. Econ. Stud.* 46:83–95

Annual Review of Economics

Volume 15, 2023

Contents

Indexes

Errata

An online log of corrections to *Annual Review of Economics* articles may be found at http://www.annualreviews.org/errata/economics