IDENTIFICATION OF INCOME-LEISURE PREFERENCES
AND EVALUATION OF INCOME TAX POLICY

Charles F. Manski
Department of Economics and Institute for Policy Research
Northwestern University

first public version: January 2012
this version: March 2012

Abstract

The merits of alternative income tax policies depend on the population distribution of preferences for income and leisure. Standard theory, which supposes that persons want more income and more leisure, does not predict how they resolve the tension between these desires. Empirical studies of labor supply have imposed strong preference assumptions that lack foundation. This paper examines anew the problem of inference on income-leisure preferences and considers the implications for evaluation of tax policy. I first perform a basic revealed-preference analysis assuming only that persons prefer more income and leisure. This shows that observation of a person's time allocation under a status quo tax policy may bound his allocation under a proposed policy or may have no implications, depending on the tax schedules and the person's status quo time allocation. I next explore the identifying power of two classes of assumptions that restrict the distribution of income-leisure preferences. One assumes that groups of persons who face different choice sets have the same preference distribution. The second restricts the shape of this distribution. The generic finding is partial identification of preferences. This implies partial prediction of tax revenue under proposed policies and partial knowledge of the welfare function for utilitarian policy evaluation.

1. Introduction

Economists have long recognized that the relative merits of alternative income tax policies depend on the preferences of individuals for income and leisure. The profession has also recognized the difficulty of inference on these preferences. In the conclusion to his seminal study of optimal income taxation, James Mirrlees wrote (Mirrlees, 1971, p. 207): "The examples discussed confirm, as one would expect, that the shape of the optimum earned-income tax schedule is rather sensitive to the distribution of skills within the population, and to the income-leisure preferences postulated. Neither is easy to estimate for real economies."

Income-leisure preferences play both positive and normative roles in analysis of tax policy, and it is important to distinguish them. The positive role is that preferences yield labor supply and other decisions that determine tax revenue. The normative role is that social welfare aggregates individual preferences in utilitarian policy evaluation. Thus, comparison of tax policies requires knowledge of preferences both to predict tax revenues and to compute the welfare achieved by alternative policies.

As I see it, we lack the knowledge of preferences necessary to credibly evaluate income tax policies. Analysis of the identification problem in Sections 2 and 3 will explain how I have reached this sober conclusion. As prelude, I first offer my perspective on research studying the response of labor supply to taxation, a subject that has receive enormous attention. I then outline the paper.

Taxation and Labor Supply

To begin, recall that standard economic theory does not predict the response of labor supply to income taxation. To the contrary, it shows that a worker may rationally respond in disparate ways. As tax rates increase, a person may rationally decide to work less, work more, or not change his labor supply at all.

The silence of theory on labor supply has long been appreciated; see Robbins (1930). Modern labor economics envisions labor supply as a complex sequence of schooling, occupation, and work effort decisions made under uncertainty over the life course. However, we need only consider a simple static scenario to see

that a person may respond to income taxes in disparate ways.

Suppose that a person with a predetermined wage and no unearned income allocates each day between paid work and the various non-paid activities that economists have traditionally called leisure. Let a proportional income tax reduce his wage by the prevailing tax rate, yielding his net wage. Assume that the person allocates time to maximize utility, which is an increasing function of net income and leisure.

Different utility functions imply different relationships between the tax rate and labor supply. The labor supply implied by utility functions in the Constant-Elasticity-of-Substitution (CES) family increases or decreases with the tax rate depending on the elasticity of substitution. Other utility functions imply that labor supply is *backward-bending*. That is, hours worked may initially increase as net wage rises from zero but, above some threshold, decrease as net wage rises further. Still other utility functions yield more complex non-monotone relationships between net wage and labor supply. The review article of Stern (1986) describes a broad spectrum of possibilities.

Given that theory does not predict how income taxation affects labor supply, prediction requires empirical analysis. Robbins (1930) emphasized this, concluding his article as follows (p. 129): "we are left with the conclusion . . . . that any attempt to predict the effect of a change in the terms on which income is earned must proceed by inductive investigation of elasticities."

Economists have performed numerous empirical studies of labor supply. Their methodologies, data, and findings have been summarized and critiqued in multiple lengthy review articles including Pencavel (1986), Killingsworth and Heckman (1986), Blundell and MaCurdy (1999), Meghir and Phillips (2010), Keane (2011), and Saez, Slemrod, and Giertz (2011). Attempting to distill the huge literature, Meghir and Phillips (2010) write (p. 204): "Our conclusion is that hours of work do not respond particularly strongly to the financial incentives created by tax changes for men, but they are a little more responsive for married women and lone mothers. On the other hand, the decision whether or not to take paid work at all is quite sensitive to taxation and benefits for women and mothers in particular." Saez, Slemrod, and Giertz (2011)

similarly write (p. 1): "With some exceptions, the profession has settled on a value for this elasticity close to zero for prime-age males, although for married women the responsiveness of labor force participation appears to be significant. Overall, though, the compensated elasticity of labor appears to be fairly small." Keane (2011) expresses a different perspective, writing (p. 1071): "My review suggests that labor supply of men may be more elastic than conventional wisdom suggests."

Reading the recent empirical literature, I have been struck to find that while authors may differ on the magnitude of labor-supply elasticities, they largely agree on the sign. The consensus is that increasing tax rates usually reduces work effort. Considering the effect of a rise in a proportional tax, Meghir and Phillips (2010) write (p. 207): "in most cases this will lead to less work, but when the income effect dominates the substitution effect at high hours of work it may increase effort." Keane (2011) states the directionality of the effect without reservation, writing (p. 963): "the use of labor income taxation to raise revenue causes people to work less." Here and elsewhere, empirical researchers may recognize the theoretical possibility that effort may increase with tax rates but view this as an empirical rarity rather than a regularity. This view has been accepted in official government forecasts of the response of labor supply to income taxation. See Congressional Budget Office (2007).[1]

Examining the models of labor supply used in empirical research, I have become concerned that the prevailing consensus on the sign of elasticities may be an artifact of model specification rather than an expression of reality.[2] Models differ across studies, but they generally share two key assumptions. First, they suppose that labor supply varies monotonically with net wages. Thus, model specifications do not

---

[1] Curiously, the opposite consensus prevailed early in the twentieth century. Gilbert and Pfouts (1958) cite assertions by Pigou and Knight in the 1920s that increasing tax rates increases work effort.

[2] The literature contains some precedent for my concern that empirical findings on labor supply may be artifacts of model specification. Concluding his detailed comparison of alternative labor supply functions, Stern (1986) wrote (p. 173): "Our general conclusion must be in favour of diversity of functions and great caution in drawing policy conclusions on results based on a particular form." Stern and other writers such as Blundell and MaCurdy (1999) have called attention to the potential detrimental consequences of restrictive functional-form assumptions.

generally permit backward-bending labor supply functions or other non-monotone relationships. Second, they suppose that the response of labor supply to net wage is homogeneous within broad demographic groups. With occasional exceptions, researchers specify hours-of-work equations that permit hours to vary additively across group members but that assume constant treatment response. That is, they assume that all group members would adjust hours worked in the same way in response to a conjectural change in net wage.[3]

The reality may be that persons have heterogeneous income-leisure preferences and, consequently, heterogeneous labor-supply functions. Some may increase work effort with net wage, others may decrease effort, and still others may exhibit a non-monotone wage-effort relationship. If so, estimates of models that assume monotonicity and homogeneity of labor supply can at most characterize the behavior of an artificial "representative" person. The estimates may not have even this limited interpretation.

---

[3] The review article of Keane (2011) uses this hours-of-work equation to illustrate common practices in the literature (p. 972):

$$\ln h_{it} \;=\; \beta + e \cdot \ln[w_{it}(1 - \tau_t)] + \beta_I \cdot N_{it} + \varepsilon_{it}.$$

Here i denotes a person, t is a time period, h is hours worked, w is wage, τ is a proportional tax rate, N is unearned income, and (β, e, β_I) are parameters. Keane writes: "Crucial is the addition of the stochastic term ε_{it}, which enables the model to explain heterogeneity in behavior." However, the model permits only a very restricted form of heterogeneity. Hours of work may vary additively across persons through ε_{it} but the parameters (β, e, β_I) do not vary across persons. In particular, the elasticity parameter e expresses the assumption of constant treatment response.

The occasional exceptions to such strong homogeneity assumptions begin with the important early research of Burtless and Hausman (1978), who emphasized the potential importance of heterogeneity but were more limited in their ability to operationalize it. The authors ultimately assumed homogeneous response of hours worked to changes in net wage. The recent contributions of Blundell and Shephard (2012) and Blundell, Bozio, and Laroque (2011) also emphasize heterogeneity, and are able to do more to operationalize it within a parametric random utility framework.

An exception to parametric modeling is Blomquist and Newey (2002). They initially propose general nonparametric mean regression of hours worked on wage and on tax schedule variables that characterize a person's (income, leisure) choice set. Concerned with the dimensionality of the variables needed to characterize the choice set, they subsequently study estimation of a model that takes desired hours of work to be a nonparametric but increasing function of a scalar measure of preference heterogeneity.

This Paper

In light of the above, this paper examines anew the problem of identification of income-leisure preferences and draws implications for evaluation of tax policy. I study inference when data on time allocation under a status quo tax policy are interpreted through the lens of standard theory. To illuminate elemental issues, I find it productive to study the classical static model in which persons must allocate one unit of time to work and leisure in an environment with predetermined production of public goods.

The classical static model is obviously simplistic, but it serves well the purpose of showing the difficulty of inference on preferences. Inference becomes even more difficult when one considers models of dynamic labor supply (e.g., MaCurdy 1985), models that consider how income taxation may affect decisions other than time allocation (e.g., Feldstein, 1995), models that conjecture socially interdependent preferences (e.g., Lindbeck, Nyberg, and Weibull, 1999), or models that suppose persons may be boundedly rational (e.g., Chetty, 2009). Inference also becomes more difficult when one wants to study policies that vary production of public goods.[4] These developments all generalize the classical static model. Hence, they logically pose more difficult inferential problems.

Sections 2 and 3 consider the use of revealed preference analysis to predict labor supply and tax revenue under a proposed policy that would alter the status quo tax schedule. A central objective is to shed light on how maintained assumptions affect the conclusions that one may draw. As in my past research, I find it illuminating to begin with weak assumptions and then to characterize the identifying power of stronger assumptions. I have exposited this approach to empirical research in Manski (2007a) and elsewhere.

Section 2 assumes only that persons prefer to have more income and leisure. Basic revealed-

---

[4] The literature on income taxation largely abstracts from this important question. Over thirty years ago, Wildasin (1979) observed (p.63-64): "The proper way of taking the effects of distortionary taxes into account in evaluating public expenditure depends sensitively on complement-substitute relations between public and private goods." He went on (p. 64): "Most bothersome of all is the fact that we have very little empirical information on the interaction between public good provision and private demand." This observation remains accurate today.

preference analysis then shows that observation of a person's time allocation under a status quo tax policy may bound his allocation under a proposed policy or may have no implications, depending on the tax schedules and the person's status quo time allocation. Basic analysis assuming only that more-is-better generically does not predict the sign of labor-supply response to change in the tax schedule.

Section 3 explores the identifying power of assumptions restricting the population distribution of preferences. For this analysis, I find it analytically helpful to suppose that persons choose among a finite set of feasible (income, leisure) values rather than the continuum often assumed in the literature. I then use the discrete-choice framework of Manski (2007b) to characterize preferences and to predict aggregate labor supply and tax revenue when various assumptions restrict the distribution of preferences. I show how to determine the identifying power of two classes of assumptions. The first assumes that groups of persons who face different choice sets and/or have different covariates have the same distribution of preferences. The second restricts the shape of this distribution. The generic finding is partial identification of the preference distribution. This implies partial ability to predict tax revenue under proposed policies. I use a computational experiment to illustrate.

The analysis of Sections 2 and 3 reaches highly cautionary conclusions about present knowledge of income-leisure. The concluding Section 4 discusses some implications for evaluation of tax policy. A familiar exercise in normative public economics poses a utilitarian social welfare function and ranks tax policies by the welfare they achieve. This requires knowledge of income-leisure preferences to predict tax revenues and compute welfare. I observe that partial knowledge of preferences implies two difficulties for policy evaluation. One can only partially predict tax revenue and one can only partially evaluate the utilitarian welfare of policies. We may potentially improve our capacity to evaluate tax policy by collecting richer data that sheds more light on preferences. I pose some possibilities for data collection.

2. Basic Analysis of Revealed Preference

To begin, Section 2.1 formalizes the conventional description of income taxation and labor supply that I use. Section 2.2 applies basic revealed-preference analysis to derive partial predictions of individual labor supply under counterfactual alternatives to a status quo tax policy. Section 2.3 studies simple cases where the net-income functions under the status quo and proposed tax schedules cross once. I call this section a "basic" analysis of revealed preference because I maintain no assumptions about preferences except that individual utility is an increasing function of net income and leisure. In short, more is better.

In the absence of assumptions that restrict the population distribution of preferences, predicting population labor supply under a proposed tax policy simply requires aggregation of individual predictions. Hence, the analysis below focuses on one person.

2.1. Tax Policy and Labor Supply

Suppose that a person labeled j is endowed with wage w_j, unearned income z_j, and one unit of time. The person must allocate his time endowment between leisure and work. If he allocates a fraction $L \in [0, 1]$ to leisure and $1 - L$ to work, he receives gross income $w_j(1 - L) + z_j$.

The status quo tax policy, denoted S, subtracts the work-dependent tax revenue $R_{jS}(L)$ from gross income, leaving j with net income

(1) $Y_{jS}(L) \equiv w_j(1 - L) + z_j - R_{jS}(L).$

The $R_{jS}(\cdot)$ notation allows the tax schedule to be specific to person j. Thus, the tax may depend on unearned income and on personal attributes that determine eligibility for exemptions and deductions. Taxes may take

positive or negative values.

Person j chooses a value of L from a set $\Lambda_j \subset [0, 1]$ of feasible leisure alternatives. Most analysis of labor supply supposes that $\Lambda_j = [0, 1]$. However, it may be more realistic to suppose that only a few allocations are feasible. For example, $\Lambda_j = \{0, \frac{1}{2}, 1\}$ means that the feasible options are full-time work (L = 0), half-time work (L = ½), and no work (L = 1).

Preferences are expressed in person j's utility function $U_j(\cdot, \cdot)$, whose arguments are (net income, leisure). Utility is strictly increasing in both arguments. Let $L_{js} \in \Lambda_j$ denote the amount of leisure that j chooses under tax schedule $R_{js}(\cdot)$. Utility maximization implies these inequalities:

(2) $U_j[Y_{js}(L_{js}), L_{js}] \geq U_j[Y_{js}(L), L], \quad \text{all } L \in \Lambda_j.$

Using $U_j(\cdot, \cdot)$ to express preferences suppresses the possible dependence of preferences on the public goods produced with tax revenue under policy S. This is innocuous if the objective is to predict labor supply under an alternative policy that alters the tax schedule but leaves public-good production unchanged.

2.2. Predicting Labor Supply under a Proposed Tax Schedule

Suppose that one observes the wage, unearned income, and other tax-relevant attributes of person j. One also observes the leisure L_{js} chosen by j under tax schedule $R_{js}(\cdot)$. Let $R_{jT}(\cdot)$ denote the tax schedule if j were to face a proposed tax policy T. What can one predict about time allocation under $R_{jT}(\cdot)$? The answer depends on the value of L_{js} and on the budget sets $\{[Y_{js}(L), L], L \in \Lambda_j\}$ and $\{[Y_{jT}(L), L], L \in \Lambda_j\}$ that j faces under the status-quo and proposed tax schedules.

Illustration

I begin with an illustration, which modestly extends the original revealed-preference argument of Samuelson (1938). Consider the tax schedules in Figure 1. Policy S has a two-rate progressive schedule and T has a proportional one, the latter crossing the former from above when leisure equals L_j^*. Person j has no unearned income.

Net Income

$Y_{jT}(\cdot)$

$Y_{jS}(\cdot)$

0   L_{jS} L_j^* L   1
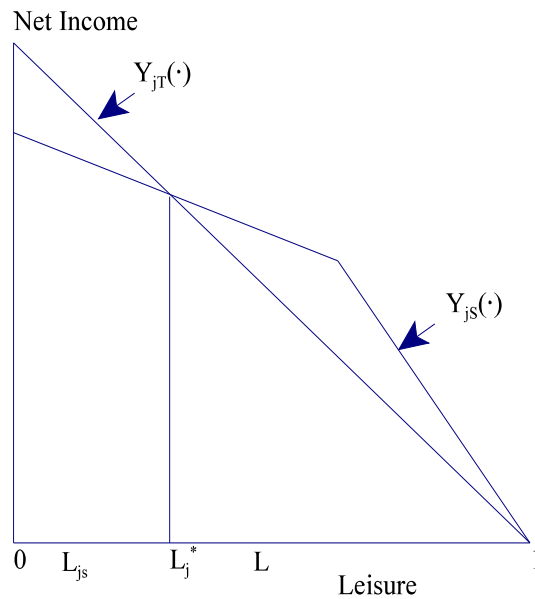
Leisure

Figure 1: Net Income with Progressive and Proportional Tax Schedules

Suppose that $L_{jS} \in [0, L_j^*]$, and consider any feasible $L > L_j^*$. The (income, leisure) pair $[Y_{jT}(L_{jS}), L_{jS}]$ is feasible under policy T. Since more income is better than less, this pair is preferred to the pair $[Y_{jS}(L_{jS}), L_{jS}]$ chosen under S. Pair $[Y_{jS}(L), L]$ is feasible under policy S. Since more income is preferred to less, this pair is preferred to $[Y_{jT}(L), L]$. The observation that j chose $[Y_{jS}(L_{jS}), L_{jS}]$ reveals that j prefers this pair to $[Y_{jS}(L), L]$. Combining these preference inequalities implies that he prefers $[Y_{jT}(L_{jS}), L_{jS}]$ to $[Y_{jT}(L), L]$. Thus, if person j were to face the proposed tax schedule, he would not choose any $L > L_j^*$. It will be shown

below that this exhausts the predictive power of basic revealed preference analysis.[5]

General Analysis

I now formally perform basic revealed-preference analysis without restricting the tax schedules generated by the status quo and proposed policies. The analysis requires a few definitions.

Given two bivariate real vectors (a, b) and (a′, b′), define (a, b) < (a′, b′) and say that (a, b) is strictly smaller than (a′, b′) if (a ≤ a′ and b ≤ b′) and (a < a′ or b < b′). Define (a, b) ≤ (a′, b′) and say that (a, b) is weakly smaller than (a′, b′) if (a ≤ a′ and b ≤ b′). Define strictly and weakly larger analogously.

Let $\Lambda_{j<}$ denote the feasible leisure values such that the (income, leisure) pairs in the $R_{jT}(\cdot)$ budget set are strictly smaller than some pair in the $R_{jS}(\cdot)$ budget set. That is,

(3) $\Lambda_{j<} \equiv \{L_< \in \Lambda_j: [Y_{jT}(L_<), L_<] < [Y_{jS}(L), L]$ for some $L \in \Lambda_j\}$.

Let $\Lambda_{j>}$ denote the leisure values such that the (income, leisure) pairs in the $R_{jT}(\cdot)$ budget set are strictly larger than $[Y_{jS}(L_{jS}), L_{jS}]$, the pair that j chooses under the status quo policy. That is,

(4) $\Lambda_{j>} \equiv \{L_> \in \Lambda_j: [Y_{jT}(L_>), L_>] > [Y_{jS}(L_{jS}), L_{jS}]\}$.

Define $\Lambda_{j\leq}$ and $\Lambda_{j\geq}$ analogously, with weak inequalities replacing the strict ones. The sets $(\Lambda_{j<}, \Lambda_{j>}, \Lambda_{j\leq}, \Lambda_{j\geq})$ are functions of the tax schedules $R_{jS}(\cdot)$ and $R_{jT}(\cdot)$, but I keep this implicit to simplify the notation.

Revealed preference analysis combines three facts. First, choice of L_{jS} yields the inequalities (2).

---

[5] I have written that this derivation modestly extends Samuelson (1938). Samuelson considered use of data on the commodity bundle chosen under a status quo linear budget set to predict the bundle that would be chosen under a counterfactual linear budget set. Figure 1 differs only in that the budget set produced by policy T is not linear. Linearity of either budget set is immaterial. Samuelson's argument holds whenever the status quo and counterfactual budget sets cross one another.

Second, the assumption that more is better implies that, for each $L_< \in \Lambda_{j<}$, there exists an $L \in \Lambda_j$ such that

(5) $U_j[Y_{jS}(L), L] > U_j[Y_{jT}(L_<), L_<]$.

Similarly, for each $L_\leq \in \Lambda_{j\leq}$, there exists an $L \in \Lambda_j$ such that

(5′) $U_j[Y_{jS}(L), L] \geq U_j[Y_{jT}(L_\leq), L_\leq]$.

Third, more-is-better implies that for each $L_> \in \Lambda_{j>}$,

(6) $U_j[Y_{jT}(L_>), L_>] > U_j[Y_{jS}(L_{jS}), L_{jS}]$.

Similarly, for each $L_\geq \in \Lambda_{j\geq}$,

(6′) $U_j[Y_{jT}(L_\geq), L_\geq] \geq U_j[Y_{jS}(L_{jS}), L_{jS}]$.

Combining (2), (5), and (6′) yields

(7) $U_j[Y_{jT}(L_\geq), L_\geq] \geq U_j[Y_{jS}(L_{jS}), L_{jS}] > U_j[Y_{jT}(L_<), L_<]$,   all $L_< \in \Lambda_{j<}$ and $L_\geq \in \Lambda_{j\geq}$.

Similarly, combining (2), (5′), and (6) yields

(7′) $U_j[Y_{jT}(L_>), L_>] > U_j[Y_{jS}(L_{jS}), L_{jS}] \geq U_j[Y_{jT}(L_\leq), L_\leq]$,   all $L_\leq \in \Lambda_{j\leq}$ and $L_> \in \Lambda_{j>}$.

Proposition 1 summarizes the implications for prediction of time allocation.

Proposition 1: Let $\Lambda_{j<}$ and $\Lambda_{j\geq}$ be non-empty. Then if person j were to face tax schedule $R_{jT}(\cdot)$, he would not choose any leisure value in $\Lambda_{j<}$. Let $\Lambda_{j\leq}$ and $\Lambda_{j>}$ be non-empty. Then if j were to face $R_{jT}(\cdot)$, he would not choose any leisure value in $\Lambda_{j\leq}$.   □

These conclusions exhaust the predictive power of basic revealed preference analysis for time allocation under the proposed tax schedule.

Welfare Implications

We may also draw partial conclusions about the welfare implications of replacing the status quo tax schedule with the proposed one. If $\Lambda_{j<} = \Lambda_j$, replacement of the status quo tax schedule with the proposal strictly decreases the welfare of person j. This follows purely from the assumption that more is better and does not require observation of j's time allocation under the status quo. If $\Lambda_{j>}$ is non-empty, replacement of the status quo with the proposal strictly increases welfare. This holds because non-emptiness of $\Lambda_{j>}$ means that the budget set under the proposed schedule contains an (income, leisure) pair larger than the one j obtained under the status quo. Basic analysis yields no welfare conclusions if $\Lambda_{j<}$ is a proper subset of Λ_j and $\Lambda_{j>}$ is empty.

2.3. Prediction When Downward-Sloping Net-Income Functions Cross Once

The above analysis was entirely general. I henceforth suppose that the net income functions $Y_{jS}(\cdot)$ and $Y_{jT}(\cdot)$ implied by the status quo and proposed tax schedules are downward sloping. This condition is substantively innocuous, it being violated only if the marginal tax rate is above 100 percent at some level of

work effort.

Given that the schedules are downward sloping, basic revealed preference analysis can have predictive power for labor supply only if the two net-income functions cross at least once. Formally, $Y_{jT}(\cdot)$ crosses $Y_{js}(\cdot)$ from above at some $L_j^* \in (0, 1)$ if $Y_{jT}(L_j^*) = Y_{js}(L_j^*)$, $Y_{jT}(L) > Y_{js}(L)$ for $L < L_j^*$, and $Y_{jT}(L) < Y_{js}(L)$ for $L > L_j^*$. Similarly, $Y_{jT}(\cdot)$ crosses $Y_{js}(\cdot)$ from below at L_j^* if the inequalities are reversed.

To see that a crossing is necessary for predictive power, consider a policy T such that $Y_{jT}(L) \geq Y_{js}(L)$ for all $L \in [0, 1]$. Thus, at all L a person pays weakly less tax under T than S. Then set $\Lambda_{j<}$ is empty because $Y_{jT}(L_<) < Y_{js}(L) \Rightarrow L_< > L$. Similarly, consider T such that $Y_{jT}(L) \leq Y_{js}(L)$ for all $L \in [0, 1]$. Thus, at all L a person pays weakly more tax under T than S. Then $\Lambda_{j\geq}$ is empty because $Y_{jT}(L_<) > Y_{js}(L_{js}) \Rightarrow L_< < L_{js}$.

Policy comparison is simple when the two tax schedules cross exactly once, as in Figure 1. When, as in the figure, $Y_{jT}(\cdot)$ crosses $Y_{js}(\cdot)$ from above, $\Lambda_{j<} = [L \in \Lambda_j: L > L_j^*]$ and $\Lambda_{j\leq} = [L \in \Lambda_j: L \geq L_j^*]$.[6] The set $\Lambda_{j>}$ is non-empty if and only if $L_{js} < L_j^*$, and $\Lambda_{j\geq}$ is non-empty if and only if $L_{js} \leq L_j^*$.[7] Hence, observation that $L_{js} < L_j^*$ implies that if person j were to face tax schedule $R_{jT}(\cdot)$, he would not choose any $L \geq L_j^*$. Observation that $L_{js} \leq L_j^*$ implies that he would not choose any $L > L_j^*$. Observation that $L_{js} > L_j^*$ has no predictive power.[8]

---

[6] To see this, first observe that $L > L_j^* \Rightarrow [Y_{jT}(L), L] < [Y_{js}(L), L]$ and $L = L_j^* \Rightarrow [Y_{jT}(L), L] = [Y_{js}(L), L]$. Hence, $[L \in \Lambda_j: L > L_j^*] \subset \Lambda_{j<}$ and $[L \in \Lambda_j: L \geq L_j^*] \subset \Lambda_{j\leq}$. Now consider $L < L_j^*$. Then $Y_{jT}(L) > Y_{js}(L)$. Given that $Y_{js}(\cdot)$ is downward sloping, $Y_{jT}(L) < Y_{js}(L') \Rightarrow L' < L$. Hence, there exists no L' such that $[Y_{jT}(L), L] < [Y_{js}(L'), L']$. Hence, $L < L_j^* \Rightarrow L \notin \Lambda_{j<}$.

[7] To see this, first observe that $L_{js} < L_j^* \Rightarrow [Y_{jT}(L_{js}), L_{js}] > [Y_{js}(L_{js}), L_{js}]$ and $L_{js} = L_j^* \Rightarrow [Y_{jT}(L_{js}), L_{js}] = [Y_{js}(L_{js}), L_{js}]$. Hence, $L_{js} < L_j^*$ implies that $\Lambda_{j>}$ is non-empty and $L_{js} = L_j^*$ implies that $\Lambda_{j\geq}$ is non-empty. Now consider $L_{js} > L_j^*$. Then $Y_{jT}(L_{js}) < Y_{js}(L_{js})$. Given that $Y_{jT}(\cdot)$ is downward sloping, $Y_{jT}(L') \geq Y_{js}(L_{js}) \Rightarrow L' < L_{js}$. Hence, there exists no L' such that $[Y_{jT}(L'), L'] \geq [Y_{js}(L_{js}), L_{js}]$. Hence, $L_{js} > L_j^*$ implies that $\Lambda_{j\geq}$ is empty.

[8] Analogous findings hold when $Y_{jT}(\cdot)$ crosses $Y_{js}(\cdot)$ from below. Then $\Lambda_{j<} = [L_< \in \Lambda_j: L < L_j^*]$ and $\Lambda_{j\leq} = [L_< \in \Lambda_j: L \leq L_j^*]$. The set $\Lambda_{j>}$ is non-empty if $L_{js} > L_j^*$ and $\Lambda_{j\geq}$ is non-empty if $L_{js} \geq L_j^*$. Hence, observation that $L_{js} > L_j^*$ implies that if j were to face schedule $R_{jT}(\cdot)$, he would not choose any $L \leq L_j^*$. Observation that $L_{js} \geq L_j^*$ implies that he would not choose any $L < L_j^*$. Observation that $L_{js} < L_j^*$ has no predictive power.

The above shows that, when the proposed tax schedule crosses the status quo from above, the sign of the response of labor supply to the change in tax schedule is predictable only in two cases. If a person works full time under the status quo schedule, then obviously labor supply cannot increase under the proposed schedule. If status quo leisure is $L_{jS} = L_j^*$, then basic analysis of revealed preference predicts that labor supply will not decrease under the proposed schedule. Thus, basic analysis predicts the sign of the labor-supply response only in the latter very special case.

Numerical Illustration

To illustrate, suppose that the progressive schedule taxes income at rate 15 percent up to $50,000 per year and at rate 25 percent above $50,000. Let the proportional schedule tax all income at rate 20 percent. Thus, the tax schedules cross when gross income equals $100,000, when both take $20,000 tax and yield net income $80,000. As in Figure 1, suppose that person j has no unearned income. The conclusions that may be drawn about time allocation and welfare under the proposed tax schedule depend on person j's wage, his feasible set of time allocations, and the allocation he chooses under the status quo schedule. In what follows, I measure wage as the annual gross labor income that j would receive for full-time work and I suppose that all time allocations are feasible.

Suppose first that w < $100,000. Then net income under the proposed schedule is lower than net income under the status quo schedule for all $L \in [0, 1]$. Hence, the proposed schedule definitely lowers welfare. Revealed preference analysis yields no prediction of time allocation.

Next let w = $150,000. Then $L^* = 1/3$, which is the roughly the case in Figure 1. Basic revealed preference analysis has predictive power if $L_{jS} \le 1/3$; that is, if j works at least 2/3 of the time in the status quo. Then we can conclude that leisure under the proposed schedule would lie in the interval [0, 1/3] and that welfare would increase. However, no prediction of time allocation or welfare is possible if $L_{jS} > 1/3$.

Finally let w = $500,000. Then $L^* = 4/5$. If $L_{jS} \le 4/5$, we can conclude that leisure under the

proposed schedule would lie in the interval [0, 4/5] and that welfare would increase. No prediction is possible if $L_{jS} > 4/5$.

Inequalities on Aggregate Labor Supply

Although basic analysis rarely predicts the sign of individual labor-supply response, it generically yields some simple inequalities relating population labor supply under the status quo and proposed tax policies. Let the proposed tax schedule cross the status quo from above. Then, among persons whose tax schedules cross at the same point ℓ, the fraction who choose leisure less than ℓ is higher under the proposed schedule than under the status quo schedule.

To see this, let J denote a population of persons formalized as a probability space (J, Ω, P). Let L_{jT} denote the leisure that person j would choose under the proposed schedule $R_{jT}(\cdot)$. We have found that, for each $j \in J$, $L_{jS} < L_j^* \Rightarrow L_{jT} < L_j^*$ and $L_{jS} \leq L_j^* \Rightarrow L_{jT} \leq L_j^*$. It follows that, for all $\ell \in [0, 1],$[9]

(8a) $P(L_S < L^* | L^* = \ell) \leq P(L_T < L^* | L^* = \ell)$

(8b) $P(L_S \leq L^* | L^* = \ell) \leq P(L_T \leq L^* | L^* = \ell).$

Inequalities (8a)-(8b) formalize a rather limited sense in which decreasing the progressivity of a tax schedule must weakly increase aggregate labor supply. If schedule T crosses S from above at point ℓ, then T is less progressive than S in the specific sense that, among people with crossing point ℓ, T yields higher net income than S when leisure is less than ℓ and lower net income than S when leisure is larger than ℓ.

I say that the inequalities formalize a "rather limited" relationship between progressivity and labor

---

[9] Hoderlein and Stoye (2012), Lemma 1 reports a version of this inequality in the context of the classical problem of consumer demand for two goods and cites earlier, more abstract, research that implies the result. In the classical demand setting, one observes demand under status quo product prices and wants to predict demand under counterfactual prices. Among persons with the same income, the status quo and counterfactual budget lines necessarily cross once, at the same point.

supply because they hold only at the crossing point of the two tax schedules. Consider (8a). It is consistent with rational choice and the more-is-better assumption to have the reverse inequality hold at any L other than ℓ. That is, it can happen that $P(L_S < L | L^* = \ell) > P(L_T < L | L^* = \ell)$ for any $L \neq \ell$.

3. Restrictions on the Preference Distribution

A huge distance separates basic revealed preference analysis from the practice of empirical analysis of labor supply. As noted in the Introduction, the models used in empirical studies usually suppose that labor supply responds monotonically to variation in net wage. Moreover, it is common to assume that time allocation differs across persons only via a person-specific additive constant.

To explore the vast middle ground between basic analysis and current practice, I find it productive to use the discrete-choice framework developed in Manski (2007b). Section 3.1 lays out basic concepts and notation. Section 3.2 shows how to determine the identifying power of alternative assumptions placing restrictions on the population distribution of preferences. Section 3.3 discusses the mathematics and the substance of the identification problem. Section 3.4 presents a computational experiment using the framework to predict tax revenue under a proposed tax policy.

3.1. Preferences on a Finite Universe of Alternatives

Consider again a population J of persons formalized as a probability space (J, Ω, P). Suppose that the status quo and proposed tax policies together generate a finite universe of potential values for (net income, leisure). That is, the universe

(9) $A \equiv \{[Y_{jS}(L), L], [Y_{jT}(L), L], L \in \Lambda_j, j \in J\}$

of distinct (Y, L) pairs that members of the population can realize under either policy has finite cardinality. Current practice usually treats net income and leisure as continuous variables, but some authors have viewed time-allocation as a choice among a discrete set of leisure options. See, for example, Blundell and Shephard (2012), Hoynes (1996), and Keane and Moffitt (1998).

Viewing the universe of feasible (Y, L) pairs as finite rather than continuous enables one to explore flexibly the identifying power of a spectrum of assumptions on the preferences of population members. Analysis is most straightforward if almost all persons have strict preference orderings. There are $|A|!$ potential strict orderings on A. The number of feasible orderings is smaller than $|A|!$ because the assumption that more-is-better excludes some orderings, the number excluded depending on the composition of A.[10]

Numerical Illustration

Consider again the illustration of Section 2.3. A status quo progressive schedule taxes income at rate 15 percent up to $50,000 per year and at rate 25 percent above $50,000. A proposed proportional schedule taxes all income at rate 20 percent.

Consider a population of high-skilled, prime-age persons. Let each person have two time-allocation alternatives, full-time or half-time work. Thus, $\Lambda_j = \{0, \frac{1}{2}\}$ for all $j \in J$. Let the population divide into two groups with different annual wages for full-time work, w = $150,000 or w = $500,000. Then the (Y, L)

---

[10] The assumption that almost all persons have strict preference orderings can be weakened. The number of possible ordering is still finite, but larger, if we allow weak orderings that may exhibit indifference among some (Y, L) pairs. However, when considering prediction, it is essential that at least some known positive fraction of orderings be strict rather than weak. The concept of rationality makes no prediction about choice among equally ranked alternatives. Revealed preference analysis is vacuous if one permits an unknown fraction of the population to be indifferent among all elements of A.

choice sets are as follows:

		Status Quo Schedule		Proposed Schedule
Group 1 (w = $150,000)				
Full-time work	a	[117,500, 0]	e	[120,000, 0]
Half-time work	b	[61,250, ½]	f	[60,000, ½]
Group 2 (w = $500,000)				
Full-time work	c	[380,000, 0]	g	[400,000, 0]
Half-time work	d	[192,500, ½]	h	[200,000, ½].

The universe A of potential values of (Y, L) contains eight elements, denoted a to h. Thus, the total number of strict preference orderings is 8! = 40,320. However, the assumption that more-is-better eliminates many orderings. In particular, it implies that feasible orderings satisfy the conditions $g \succ c \succ e \succ a$, $h \succ d \succ b \succ f$, and $d \succ e$. Eliminating the orderings that do not satisfy these conditions leaves 53 feasible orderings.

3.2. Identification Analysis

Let A_k, $k \in K$ denote the orderings of set A that are consistent with the assumption that more-is-better. Then the population distribution of preferences is multinomial with at most $|K|$ mass points, where $|K| \leq |A|!$. All assumptions about population preferences take the form of restrictions on this multinomial distribution.[11]

---

[11] In Manski (2007b) I observed that prediction of counterfactual choice behavior does not require that one distinguish all feasible preference orderings from one another. Multiple preference orderings may yield identical choices when facing all status quo and proposed choice sets. Thus, it is only necessary to distinguish groups of preference orderings that have different *choice functions;* that is, groups yielding different choices in some relevant choice setting. This permits one to reduce the set K of distinct preference orderings to a smaller set of distinct choice functions, which can be computationally advantageous in the analysis below.

I will, however, not reduce K to the set of distinct choice functions. The reason is that prediction of counterfactual choice behavior is not the sole concern of this paper. A further concern, discussed in Section 4, is to compare policies via a utilitarian social welfare function. Such policy comparison requires one to distinguish persons by their utility functions, not just by their choice functions.

I show here how to determine the identifying power of two layers of assumptions. The first assumes that groups of persons who face different status quo or proposed choice sets have the same distribution of preferences. The second adds restrictions on the shape of this distribution. In all that follows, I suppose that one observes the wage, unearned income, and other tax-relevant attributes of the members of the population.

To begin, divide the population into a collection $(J_m, m \in M)$ of mutually exclusive and exhaustive groups according to the (Y, L) choice sets they face under the status quo and proposed tax schedules. Thus, all members of group m face a common choice set $C_m \subset A$ of (Y, L) pairs under the status quo schedule and a common set $D_m \subset A$ under the proposed schedule. Groups m and m′ are distinct if $C_m \neq C_{m'}$ or $D_m \neq D_{m'}$. Substantively, groups are composed of persons with the same tax-relevant attributes. Persons who have the same wage, unearned income, exemptions/deductions, and live in the same tax jurisdiction face the same income-leisure options.

Next, divide the population into a finite collection X of groups of persons with common observed covariates. These covariates need not be tax relevant. For example, covariate groups may be defined by gender, race, and/or education. Let J_{mx} denote the sub-population of persons who are members of tax group m and covariate group x.

Let $\pi_{mx} \equiv (\pi_{mxk}, k \in K)$ denote the multinomial distribution of preferences in (tax, covariate) group (m, x). Thus, for each $k \in K$, π_{mxk} is the fraction of group (m, x) who have preference ordering A_k. With no assumptions on the population distribution of preferences, $(\pi_{mx}, m \in M, x \in X)$ is an unrestricted set of multinomial distributions, one for each (tax, covariate) group. The analysis of this section first assumes that a specified set of groups have the same preference distribution and then places restrictions on the shape of this distribution. Formally, consider

Assumption 1: For a specified class of (tax, covariate) groups $N \subset M \times X$, there exists a single multinomial distribution $\pi_N \equiv (\pi_{Nk}, k \in K)$ such that $\pi_{mx} = \pi_N$, all $(m, x) \in N$. $\quad \square$

Assumption 2: π_N lies in a specified set Π_N of multinomial distributions with at most $|K|$ mass points. □

Assumption 1 poses a substantive requirement whose credibility depends on the specification of N. One may think it credible to assume that some (m, x) groups have the same preference distribution but not others. Assumption 1 does not restrict the shape of the distribution π_N of (income, leisure) preferences. Assumption 2 generically adds such restrictions. Particularly simple are *linear* restrictions that take Π_N to be a set of multinomial distributions satisfying specified linear equalities or inequalities. These include assumptions restricting the preference orderings that appear in the population. Suppose one believes that some orderings, say $K_{N0} \subset K$, do not occur in class N. Then $\pi_{Nk} = 0, k \in K_{N0}$.

The identifying power of Assumptions 1 and 2 is easy to characterize, but we first need some new notation. For each non-empty set $B \subset A$, let $c_k(B)$ denote the (Y, L) pair that a person with preference ordering k would choose if he were to face choice set B. Let $c_j(B)$ denote the pair that person j would choose and let the random variable c(B) denote the choice of a person randomly drawn from J. For each (m, x) $\in$ N and (Y, L) $\in C_m$, let $P[c(C_m) = (Y, L)|J_{mx}]$ give the observed fraction of sub-population J_{mx} who make choice (Y, L) when facing their status quo choice set C_m. Similarly, for each (m, x) $\in$ N and (Y, L) $\in D_m$, let $P[c(D_m) = (Y, L)|J_{mx}]$ give the unobserved fraction of persons in J_{mx} who would choose (Y, L) if they were to face choice set D_m.

Assumption 1 implies that the observed choice probabilities $P[c(C_m) = (Y, L)|J_{mx}]$, (Y, L) $\in C_m$, (m, x) $\in$ N are related to the preference distribution π_N through the linear equations

$$(10) \quad P[c(C_m) = (Y, L)|J_{mx}] = \sum_{k \in K} 1[c_k(C_m) = (Y, L)] \cdot \pi_{Nk}, \qquad (Y, L) \in C_m, \quad (m, x) \in N.$$

For each value of (m, x) and (Y, L), the right-hand side adds up the fractions of J_{mx} whose preference orderings make (Y, L) their most preferred alternative in C_m. Distribution π_N also satisfies the adding-up and

non-negativity conditions

(11a) $\quad \sum_{k \in K} \pi_{Nk} = 1,$

(11b) $\quad \pi_{Nk} \geq 0, \quad k \in K.$

Finally, Assumption 2 states that

(12) $\quad \pi_N \in \Pi_N.$

Proposition 2 states the implications for identification of the distribution of preferences.

Proposition 2: Given Assumptions 1 and 2, the identification region for π_N is the set $H(\pi_N)$ of multinomial distributions that solve (10), (11), and (12). $\quad \square$

The set $H(\pi_N)$ has a simple structure when Π_N places linear restrictions on π_N. Then $H(\pi_N)$ is the convex set of vectors in $R^{|K|}$ that solve a set of linear equalities and inequalities. $H(\pi_N)$ is non-empty if the preference assumptions are correct; that is, if almost all persons have strict preferences satisfying more-is-better and if Assumptions 1 and 2 hold. If $H(\pi_N)$ is empty, then some assumption is incorrect.

When $H(\pi_N)$ is non-empty, it is straightforward to predict the choice behavior of class N under the proposed tax schedule. The unobserved choice probabilities $P[c(D_m) = (Y, L)|J_{mx}]$, $d \in D_m$, $(m, x) \in N$ are related to π_N through the linear equations

(13) $P[c(D_m) = (Y, L)|J_{mx}] = \sum_{k \in K} 1[c_k(D_m) = (Y, L)] \cdot \pi_{Nk}, \qquad (Y, L) \in D_m, \quad (m, x) \in N.$

The feasible preference distributions are the elements of $H(\pi_N)$. Hence, the identification region for the unobserved choice probabilities is

(14) $H\{P[c(D_m) = (Y, L)|J_{mx}], (Y, L) \in D_m, (m, x) \in N\}$

$$= \{\{ \sum_{k \in K} 1[c_k(D_m) = (Y, L)] \cdot p_k, (Y, L) \in D_m, (m, x) \in N\}, \text{ all } (p_k, k \in K) \in H(\pi_N)\}\}.$$

A question of policy interest is to predict tax revenue under the proposed tax schedule. Recall that L_{jT} denotes the leisure that person j would choose under the proposed schedule $R_{jT}(\cdot)$. Let the random variable $R_T(L_T)$ denote the tax paid by a randomly drawn member of N. Then the per capita tax paid by the persons in N under the proposed schedule is $E[R_T(L_T)|J_N]$. This is related to the preference distribution as follows:

(15) $E[R_T(L_T)|J_N] = \sum_{(m, x) \in N} P(m, x) \sum_{(Y, L) \in D_m} R_{mT}(L) \sum_{k \in K} 1[c_k(D_m) = (Y, L)] \cdot \pi_{Nk},$

where $R_{mT}(\cdot)$ is the common tax schedule faced by members of tax group m and where $P(m, x)$ is the fraction of the population who belong to group (m, x). Hence, the identification region for per capita tax revenue in class N is the set

(16) $H\{E[R_T(L_T)|J_N]\} = \{ \sum_{(m, x) \in N} P(m, x) \sum_{(Y, L) \in D_m} R_{mT}(L) \sum_{k \in K} 1[c_k(D_m) = (Y, L)] \cdot p_k,$

$$\text{all } (p_k, k \in K) \in H(\pi_N)\}.$$

When Π_N places linear restrictions on π_N, convexity of $H(\pi_N)$ and the linearity of equation (14) in π_N imply that $H\{E[R_T(L_T)]\}$ is an interval. The lower and upper bounds of this interval solve two linear programming problems. The lower bound solves

$$(17) \quad \min_{(p_k, \, k \in K) \, \in \, H(\pi_N)} \quad \sum_{(m, \, x) \, \in \, N} P(m, x) \sum_{(Y, \, L) \, \in \, D_m} R_{mT}(L) \sum_{k \, \in \, K} 1[c_k(D_m) = (Y, L)] \cdot p_k.$$

The upper bound solves the analogous problem, with max replacing min.

When the restrictions imposed by Π_N are not linear, minimization problem (16) still gives the sharp lower bound and the corresponding maximization problem gives the sharp upper bound. However, the identification region may or may not be the entire interval connecting these bounds, depending on the structure of Π_N.

3.3. Discussion

The identification analysis of Section 3.2 is mathematically simple. When Π_N places linear restrictions on π_N, computation is tractable as well. However, two matters are less straightforward, one analytical and the other substantive.

Analytically, one would like to be able to characterize succinctly the identifying power of alternative assumptions. When Π_N places linear restrictions on π_N, one can crudely count the number of linear equations and inequalities that (10), (11), and (12) place on the $|K|$-dimensional vector π_N of multinomial probabilities. Unfortunately, doing so does not reveal much about the size and shape of $H(\pi_N)$, nor about the width of bounds on derived quantities such as tax revenue under proposed tax schedules. The structure of $H(\pi_N)$ depends delicately on the choice sets that the various tax groups in class N face under the status quo tax policy. The implied bounds on outcomes under a proposed tax schedule depend as well on the choice sets

that arise under the proposed schedule.

To learn the identifying power of alternative assumptions, it appears necessary to impose them and compute the resulting identification regions. There does not seem to be an effective shortcut to characterize $H(\pi_N)$ or derived quantities. This should not be surprising. A body of research that began with Marschak (1960) present has found that even simple configurations of observed binary choice probabilities imply subtle restrictions on preference distributions. See Fishburn (1992) and McFadden (2005).

The substantive issue is that we lack a credible basis for placing informative restrictions on preference distributions. Standard economic theory only suggests that persons have preference orderings that satisfy more-is-better. This is a compelling starting point for analysis, but it is not clear how to go beyond it. The strong assumptions regarding the homogeneity and structure of preferences used in empirical research on labor supply have been motivated by a desire to obtain tractable point estimates, not by empirical evidence.

Researchers sometimes try to motivate preference assumptions by viewing time allocation as a neoclassical problem of budget allocation, with income and leisure being goods akin to apples and oranges. They may thus introspect on whether leisure is a normal good, whether the income effect of a change in net wage outweighs the substitution effect, and so on. Such introspection may perhaps be a helpful way to conjecture plausible forms for income-leisure preferences. However, it is important to understand that the analogy of time allocation and neoclassical budget allocation is inexact.

The neoclassical problem presumes that a person allocates a specified money budget between two goods with linear pricing determining the set of feasible allocations. In contrast, the time-allocation problem presumes that a person begins with two asymmetric endowments, one of time and the other of unearned income. The asymmetry is that the person can transform time into income through paid work, but he cannot transform unearned income into time. Moreover, non-linear pricing determines the set of feasible

transformations of time into income except in the special case of a proportional tax.[12] Further distinctions

arise if the set of feasible time allocations is finite rather than a continuum as assumed in neoclassical theory.

All of these factors make me wary of using neoclassical consumer theory to introspect about time allocation.

3.4. A Computational Experiment

The classical static model is too simplistic to warrant application in substantive study of labor

supply. Nevertheless, it is revealing to perform computational experiments that show the identifying power

of alternative assumptions on the preference distribution. I describe a representative such experiment here.

Given data on labor supply under a status-quo progressive tax schedule, the problem is to predict the

tax revenue per capita that would materialize under a proposed proportional schedule. I show the predictive

power of a sequence of increasingly strong assumptions: (1) more is better; (2) additionally, persons in

specified wage groups have the same distribution of preferences; (3) additionally, preferences have the CES

form; and (4) additionally, all CES utility functions in the specified wage groups have the same elasticity of

substitution.

Tax Policies, Wages, Choice Sets, and Preferences

The progressive status quo policy taxes income at rate 20 percent up to $100,000 per year and at rate

30 percent above $100,000. The proposed proportional policy taxes all income at rate 25 percent. Thus, the

tax schedules cross when gross income equals $200,000, when both take $50,000 tax and yield net income

---

[12] Beginning with Hall (1973), it has been common to model non-proportional tax schedules as if they are proportional. The idea is to view an individual facing a non-proportional status quo schedule as if he were facing a proportional tax equal to the marginal tax rate at his chosen time allocation. His inframarginal net income is treated as a "virtual" form of unearned income. This modeling strategy has become so prevalent that researchers often neglect to cite its maintained assumptions. These include that income-leisure preferences have convex indifference curves and that all time allocations in the [0, 1] interval be feasible.

$150,000.

The population is composed of persons who have no unearned income and who have one of eight annual wages for full-time work: {50, 100, 150, 200, 250, 300, 350, 400} thousand dollars per year. The assumed population wage distribution is

$$P(w = 50) = 0.70, \quad P(w = 100) = 0.20, \quad P(w = 150) = 0.05, \quad P(w = 200) = 0.02$$

$$P(w = 250) = 0.01, \quad P(w = 300) = 0.0075, \quad P(w = 350) = 0.0075, \quad P(w = 400) = 0.005.$$

This distribution roughly approximates that of the 2011 population of adult males in the United States. The tax schedule under the status quo policy is a simplified idealization of the prevailing American tax schedule.

I suppose that each person can work full-time, three-quarter time, half-time, or not at all. Thus, Λ_j = {0, ¼, ½, 1} for all $j \in J$. Appendix Table A1 gives the resulting universe of potential values of (Y, L).

When using the assumption that persons have CES preferences, I suppose that person j has a utility function of the form

(18) $U_j(Y, L) = [\alpha_j(Y/400,000)^{\rho_j} + (1 - \alpha_j)L^{\rho_j}]^{1/\rho_j}.$

Dividing Y by 400,000, which is the maximum annual wage for full-time work, re-scales income to the interval [0, 1], to make income and leisure have comparable range. The identification analysis presumes that the parameter α may take one of the 101 values {0, 0.01, 0.02, . . . , 0.99, 1} and that ρ may take one of the 211 values {-100, -90, . . . , -20, -10, -1, -0.99, -0.98, . . . , 0.98, 0.99, 1}. Hence, there are 101 × 211 = 21,311 possible CES preferences. Income and leisure are essentially perfect complements when $\rho = -100$. They are perfect substitutes when $\rho = 1$. The value $\rho = 0$ gives Cobb-Douglas preferences.

The final step in specification of the experiment is to choose the actual distribution of preferences

in the population. I suppose that the population actually contains persons having twenty distinct CES preferences, with $(\alpha, \rho) \in \{0.25, 0.5, 0. 65, 0.75\} \times \{-100, -0.5, 0, 0.5, 1\}$. Moreover, I suppose that the population distribution of (α, ρ) is uniform conditional on wage. Thus, the population subgroups with different wages share the same preference distribution, placing probability 1/20 on each feasible (α, ρ) pair. This distribution of (α, ρ) values was selected so the population would contain a modest fraction of persons who choose different time allocations under policies S and T.[13]

Findings

Given this specification, actual per capita tax revenues under the status quo and proposed tax policies are $7,652 and $9,211 respectively. Appendix Table A2 details the labor supply and taxes paid under the two policies by persons with different wages.

If one were to know the wage distribution but not know status quo labor supply, one would only be able to predict that tax revenue per capita under the proposed policy must lie in the interval [0, $18969]. The upper and lower bounds are the taxes paid if everyone were to work full time or not at all, respectively.

Given data on status quo labor supply, one's ability to predict tax revenue under the proposed policy depends on the assumptions imposed on preferences. The basic assumption that more-is-better yields the revenue bound [$593, $18969]. This result follows from the analysis of Section 2.3. The upper bound is the same as the upper bound with no labor supply data because working full time under the proposed policy is always consistent with the assumption that more-is-better. The lower bound is larger than zero because any person j for whom $L_{jS} \leq L_j^*$ would not choose $L_{jT} > L_j^*$. In the context of the computational experiment, this condition only raises the lower bound on per capita tax revenue from 0 to $593.

I next combine the assumption that more-is-better with the assumption that persons within each of

[13] Persons with the six (α, ρ) pairs $\{(0.5, -0.5), (0.65, -0.5), (0.65, 0), (0.65, 0.5), (0.65, 1), (0.75, 0.5)\}$ change their choices if they have certain values of w. Those with the other fourteen (α, ρ) pairs never change regardless of w.

two broad wage groups have the same distribution of preferences. I assume that persons with w ≤ $200,000 have the same preference distribution and that those with w > $200,000 have the same preference distribution. I do not assume that persons with wage below and above $200,000 have the same preference distribution. The resulting bound on per capita tax revenue is [$3744, $14149]. This result follows from application of equation (16) to each of the two wage groups and aggregation of the findings across the two groups.[14]

The above result applies a version of Assumption 1 but does not use any version of Assumption 2. I next suppose that all persons have CES utility functions. Then application of equation (16) to each wage group and aggregation across the two groups yields the bound [$6883, $10444] on per capita tax revenue.[15]

The assumption that everyone has CES preferences is strong. However, it is still much weaker than the assumptions of homogeneous preferences that have been traditionally maintained in parametric empirical analyses of labor supply. My final step is to assume that the persons in each broad wage group have the same elasticity of substitution. Thus, I assume that all persons with w ≤ 200000 have the same value of ρ in their CES utility functions and, similarly, all persons with w > 200000 have the same ρ. I place no homogeneity assumption on personal values of α. It turns out that assuming homogeneity of ρ yields an empty identification region for the preference distribution within the second broad wage group. Thus, the assumption of homogeneous ρ within each wage group is rejected.

---

[14] Within the groups with w ≤ $200,000 and w > $200,000, the bounds on per capita tax revenue are [$2941, $12455] and [$29701, $68906] respectively. The fractions of the population with wages in these two groups are 0.97 and 0.03 respectively. Averaging the bounds with these fractional weights yields the aggregate result [$3744, $14149].

[15] Within the groups with w ≤ $200,000 and w > $200,000, the bounds on per capita tax revenue are [$5947, $9143] and [$37148, $52500] respectively. The fractions of the population with wages in these two groups are 0.97 and 0.03 respectively. Averaging the bounds with these fractional weights yields the aggregate result [$6883, $10444].

Computation of the Bounds

Computation of the above bounds ranges from easy to challenging. The bound making only the basic assumption that more-is-better is simple to compute. It uses only the analytical result of Section 2.3 and does not require the identification analysis of Section 3.2.

The bound assuming that preferences have the CES form uses the analysis of Section 3.2, but it is not hard to compute because the number of distinct choice functions is fairly small. Although persons may have 21,311 distinct values of (α, ρ), this set partitions into a much smaller number of equivalency classes, within which distinct (α, ρ) values yield the same choices under both tax schedules. It turns out that 108 and 99 equivalency classes suffice to describe the behavior of persons with $w \leq \$200,000$ and $w > \$200,000$ respectively. Thus, the relevant multinomial distributions π place probability on 108 and 99 mass points respectively. Solution of the resulting linear programming problem (17) is tractable.

The most challenging task is to compute the intermediate bound assuming that persons within each wage group have the same distribution of preferences, but not assuming that preferences have the CES form. Inspection of Table A1 shows that the members of each wage group face 24 distinct (Y, L) pairs under the status quo and proposed tax schedules. Thus, the total number of strict preference orderings in each group is $24! \cong 6.2 \times 10^{23}$. Laborious application of an algorithm that determines and excludes orderings inconsistent with the more-is-better assumption drastically reduces the number of feasible orderings to 7,654 and 4,000 in the groups with $w \leq \$200,000$ and $w > \$200,000$ respectively. Hence, the relevant multinomial distributions π place probability on 7,654 and 4,000 mass points respectively. Solution of the resulting linear programming problem (17) is still tractable.[16]

---

[16] I am grateful to Jörg Stoye for making available the MATLAB code used to compute the intermediate bounds. This code was developed for his analysis of related testing and prediction problems that arise in applications of neoclassical consumer theory. See Kitamura and Stoye (2011).

4. Implications for Policy Evaluation

A familiar exercise in normative public economics poses a utilitarian social welfare function and ranks tax policies by the welfare they achieve. Performing this ranking requires knowledge of income-leisure preferences both to predict tax revenues and to compute the welfare achieved by alternative policies. Thus, Mirrlees (1971) wrote in his introductory section (p. 176): "The State is supposed to have perfect information about the individuals in the economy, their utilities and, consequently, their actions." This assumption has been standard in the subsequent literature on optimal income taxation.

Consider a status-quo policy S with tax schedules $[R_{jS}(\cdot), j \in J]$ and a proposed policy T with corresponding $[R_{jT}(\cdot), j \in J]$. Given cardinal representations for persons' ordinal utility functions, these policies yield utilitarian welfare $E\{[U[(Y_S(L_S), L_S]\}$ and $E\{U[Y_T(L_T), L_T]\}$ respectively. Policy T is strictly preferred to S if

(19) $\quad E\{[U[(Y_T(L_T), L_T]\} \ - \ E\{U[Y_S(L_S), L_S]\} \ > \ 0.$

Based on the analysis in Sections 2 and 3, I conclude that we lack the knowledge of preferences necessary to credibly rank policies. Considering the classical static model, Section 2 showed that basic revealed preference analysis has little power to predict labor supply under proposed policies. Importantly, it does not predict whether increasing tax rates reduces or increases work effort. Section 3 showed that Assumptions 1 and 2 add predictive power, but credible versions of these assumptions generally yield bounds rather than point predictions. The computational experiment of Section 3.4 is only an illustration, but I think it typifies the situations that empirical researchers using the classical static model face in practice. Studies using more general models of time allocation face more difficult inferential problems.

Partial identification of income-leisure preferences creates two distinct difficulties for policy

evaluation. First, partial knowledge of the tax revenue yielded by a proposed policy makes it impossible to ensure that the policy yields a balanced budget. Whatever the intent ex ante, a budget surplus or deficit may materialize ex post. This makes it infeasible to evaluate policies within a deterministic static framework of the type used by Mirrlees and others. One requires a suitable dynamic perspective on private time-allocation and social planning under uncertainty.

Second, partial identification of preferences implies partial knowledge of the utilitarian welfare achieved by a policy. Thus, choice of tax policy becomes a problem of planning under ambiguity. Basic decision theory prescribes that a planner should first eliminate dominated policies and then use some reasonable criterion to choose among the undominated policies. Studying other planning problems, I have found it instructive to compare the policy choices that result from maximization of subjective expected welfare and from application of the maximin and minimax-regret criteria (Manski, 2011). A utilitarian planner may use these or other criteria to choose a tax policy.

A challenge for future research is to develop tractable approaches to the study of choice of tax policy under ambiguity. One issue, already mentioned above, is the need for a dynamic perspective on planning. Another is that the knowledge of preferences required to compute welfare is more extensive than the knowledge needed to predict labor supply. The transformation from preferences to time-allocation under status quo and proposed tax policies is many-to-one. That is, distinct preferences may imply the same choices and, hence, be equivalent from the perspective of prediction of labor supply.[17] However, such distinct preferences may have different utilitarian welfare implications.

Yet another complication is that specification of a welfare function requires that the planner select cardinal representations of ordinal individual utility functions. These representations are already consequential for policy comparison in the Mirrlees setting where all members of the population have the

---

[17] The computational experiment in Section 3.4 gave an example. Recall that the 21,311 distinct values of the CES preference parameters (α, ρ) partitioned into a much smaller number of 108 or 99 equivalency classes, within which distinct (α, ρ) values yielded the same choices under both tax schedules.

same known preferences. They are even more consequential when preferences may be heterogeneous and the planner has only partial knowledge of the preference distribution.

Enriching the Data for Identification of Preferences

I have in this paper assumed that the available data are observations of time allocation under a status quo tax policy. Collection of richer data can potentially sharpen knowledge of income-leisure preferences.

One avenue, although a difficult one to operationalize, is to resume the performance of randomized experiments with tax policy that began with the negative income tax experiments of the 1970s. See Hausman and Wise (1985) and Munnell (1986) for assessment of the lessons learned. Randomized experiments can contribute to inference by making it credible to assume that groups of persons who face different tax schedules have the same distribution of income-leisure preferences. Thus, experiments can legitimize applications of Assumption 1.[18]

A second avenue is to obtain richer data from individuals than observation of their time allocation under a single status quo tax policy. If one takes seriously the classical static model, it is useful to repeatedly observe the labor supply of an individual in circumstances with varying wages, unearned income, or tax schedules. Assuming that the person's preferences do not vary over time, transitivity implies that basic revealed preference analysis of the type performed in Section 2 has increasing predictive power as more choices are observed (Samuelson, 1938, 1948; Afriat, 1967). However, this interpretation of repeated observations rests critically on the static model. Suppose, to the contrary, that one views labor supply in

---

[18] Randomized experiments may also be used to directly learn labor supply under proposed policies, without reference to economic theory. This is the traditional use of experiments in analysis of treatment response. However, such analysis of experimental data yields no ability to extrapolate findings to policies not administered in the experiment. Moreover, it provides none of the knowledge of preferences necessary to compute utilitarian welfare.

dynamic terms, as in MaCurdy (1985). Then repeated observations of labor supply provide data on a single sequential choice path rather than data on choice from multiple independent choice sets.

An alternative way to obtain richer data from individuals is to interview them, pose scenarios with hypothetical wages or tax schedules, and ask them to predict their time allocation in these scenarios. Analysis of stated-preference data of this sort has a long history in empirical choice analysis, illustrative studies using various methods ranging from Beggs, Cardell, and Hausman (1981) and Ben Akiva and Morikawa (1990) to Blass, Lach, and Manski (2010). The presumption in these and other studies using stated preference data is that individuals introspect sufficiently well as to provide informative reports of their preferences. As far as I am aware, the approach has not been used to study income-leisure preferences.

References

Afriat, S. (1967), "The Construction of Utility Functions from Expenditure Data," *International Economic Review*, 8, 67-77.

Beggs, S., S. Cardell, and J. Hausman (1981), "Assessing the Potential Demand for Electric Cars," *Journal of Econometrics*, 16, 1-19.

Ben-Akiva, M. and T. Morikawa (1990), "Estimation of Switching Models from Revealed Preferences and Stated Intentions," *Transportation Research A* 24A, 485-495.

Blass, A., S. Lach, and C. Manski (2010), "Using Elicited Choice Probabilities to Estimate Random Utility Models: Preferences for Electricity Reliability," *International Economic Review*, 51, 421-440.

Blomquist, S. and W. Newey (2002), "Nonparametric Estimation with Nonlinear Budget Sets," *Econometrica*, 70, 2455-2480.

Blundell, R., A. Bozio, and G. Laroque (2011), "Extensive and Intensive Margins of Labour Supply: Working Hours in the US, UK and France," Institute for Fiscal Studies Working Paper 01/11.

Blundell, R. and T. MaCurdy (1999), "Labor Supply: A Review of Alternative Approaches," in O. Ashenfelter and D. Card (editors), *Handbook of Labor Economics* Vol.3, Amsterdam: North-Holland, 1559-1695.

Blundell, R. and A. Shephard (2012), "Employment, Hours of Work and the Optimal Taxation of Low-Income Families," *Review of Economic Studies*, forthcoming.

Burtless, G. and J. Hausman (1978), "The Effect of Taxes on Labor Supply," *Journal of Political Economy,* 86, 1103-1130.

Chetty, R. (2009). "Bounds on Elasticities with Optimization Frictions: A Synthesis of Micro and Macro Evidence on Labor Supply," National Bureau of Economic Research Working Paper 15616.

Congressional Budget Office (2007), "The Effect of Tax Changes on Labor Supply in CBO's Microsimulation Tax Model," Congress of the United States, Congressional Budget Office.

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.

Fishburn, P. (1992), "Induced Binary Probabilities and the Linear Ordering Polytope: A Status Report," *Mathematical Social Sciences* 23, 67-80.

Gilbert, F. and R. Pfouts (1958), "A Theory of the Responsiveness of Hours of Work to Changes in Wage Rates," *Review of Economics and Statistics*, 40, 116-121.

Hall, R. (1973), "Wages, Income and Hours of Work in the US Labor Force," in G. Cain and H. Watts (editors), *Income Maintenance and Labor Supply,* Chicago: University of Chicago Press.

Hausman, J. and D. Wise (eds) (1985), *Social Experimentation*, Chicago: University of Chicago Press.

Hoderlein and Stoye, J. (2012), "Testing Stochastic Rationality and Predicting Stochastic Demand: The Case of Two Goods," Department of Economics, Cornell University.

Hoynes, H. (1996), "Welfare Transfers in Two-Parent Families: Labor Supply and Welfare Participation Under AFDC-UP", *Econometrica*, 64, 295–332.

Keane, M. (2011), "Labor Supply and Taxes: A Survey," *Journal of Economic Literature*, 49, 961–1075.

Keane, M. and R. Moffitt (1998), "A Structural Model of Multiple Welfare Program Participation and Labor Supply," *International Economic Review*, 39, 553–589.

Killingsworth, M. and J. Heckman (1986), "Female Labor Supply: a Survey," in O. Ashenfelter and R. Layard (editors), *Handbook of Labor Economics* Vol.1, Amsterdam: North-Holland, 103-204.

Kitamura, Y. and J. Stoye (2011), "Nonparametric Analysis of Random Utility Models," Department of Economics, Cornell University.

Lindbeck, A., Nyberg, and J. Weibull (1999), "Social Norms and Economic Incentives in the Welfare State," *Quarterly Journal of Economics*, 114, 1-35.

MaCurdy, T. (1985), "Interpreting Empirical Models of Labor Supply in an Intertemporal Framework with Uncertainty," in J. Heckman and B. Singer (editors), *Longitudinal Analysis of Labor Market Data*, Cambridge: Cambridge University Press.

Manski, C. (2007a), *Identification for Prediction and Decision*, Cambridge, MA: Harvard University Press.

Manski, C. (2007b), "Partial Identification of Counterfactual Choice Probabilities," *International Economic Review*, 48, 1393-1410.

Manski, C. (2011), "Choosing Treatment Policies under Ambiguity," *Annual Review of Economics*, 3, 25-49.

Marschak, J. (1960), "Binary Choice Constraints on Random Utility Indicators," in K. Arrow (editor) *Stanford Symposium on Mathematical Methods in the Social Sciences,* Stanford, Ca.: Stanford University Press.

McFadden, D. (2005), "Revealed Stochastic Preference: A Synthesis," *Economic Theory*, 26, 245-264.

Meghir, C. and D. Phillips (2010), "Labour Supply and Taxes," in T. Besley, R. Blundell, M. Gammie, and J. Poterba (editors), Dimensions of Tax Design: the Mirrlees Review. Oxford: Oxford University Press, 202-274.

Mirrlees J. (1971), "An Exploration in the Theory of Optimal Income Taxation," *Review of Economic Studies*, 38, 175-208.

Munnell, A. (editor) (1986), *Lessons from the Income Maintenance Experiments*, Boston: Federal Reserve Bank of Boston.

Pencavel, J. (1986), "Labor supply of men: a survey," in O. Ashenfelter and R. Layard (editors), *Handbook of Labor Economics* Vol.1, Amsterdam: North-Holland, 3-102.

Robbins, L. (1930), "On the Elasticity of Demand for Income in Terms of Effort," *Economica*, 29, 123-129.

Saez, E., J. Slemrod, and S. Giertz (2011), "The Elasticity of Taxable Income with Respect to Marginal Tax Rates: A Critical Review," *Journal of Economic Literature*, forthcoming.

Samuelson, P. (1938), "A Note on the Pure Theory of Consumer Behavior," *Economica*, 5, 61-71.

Samuelson, P. (1948), "Consumption Theory in Terms of Revealed Preferences," *Economica*, 15, 243-253.

Stern, N. (1986), "On the Specification of Labour Supply Functions," in R. Blundell and I. Walker (editors) *Unemployment, Search and Labour Supply*, Cambridge: Cambridge University Press, 143-189.

Wildasin, D. (1979), "Public Good Provision with Optimal and Non-Optimal Commodity Taxation," *Economic Letters*, 4, 59-64.

Table A1: Universe of (Y, L) Values in the Computational Experiment

Wage	(Net Income, Leisure) Status Quo	Proposed	Wage	(Net Income, Leisure) Status Quo	Proposed
50000	[40000, 0]	[37500, 0]	250000	[185000, 0]	[187500, 0]
	[30000 ¼]	[28125, ¼]		[141250, ¼]	[140625, ¼]
	[20000 ½]	[18750, ½]		[97500, ½]	[93750, ½]
	[0, 1]	[0, 1]		[0, 1]	[0, 1]
100000	[80000, 0]	[75000, 0]	300000	[220000 0]	[225000, 0]
	[60000, ¼]	[56250, ¼]		[167500, ¼]	[168750, ¼]
	[40000, ½]	[37500, ½]		[115000, ½]	[112500, ½]
	[0, 1]	[0, 1]		[0, 1]	[0, 1]
150000	[115000, 0]	[112500, 0]	350000	[255000, 0]	[262500, 0]
	[88750, ¼]	[84375, ¼]		[193750, ¼]	[196875, ¼]
	[60000, ½]	[56250, ½]		[132500, ½]	[131250, ½]
	[0, 1]	[0, 1]		[0, 1]	[0, 1]
200000	[150000, 0]	[150000, 0]	400000	[290000, 0]	[300000, 0]
	[115000, ¼]	[112500, ¼]		[220000, ¼]	[225000, ¼]
	[80000, ½]	[75000, ½]		[150000, ½]	[150000, ½]
	[0, 1]	[0, 1]		[0, 1]	[0, 1]

Table A2: Actual Labor Supply and Tax Revenues in the Computational Experiment

Wage	Status Quo Policy						Proposed Policy				
	Leisure Choice				Tax per Capita		Leisure Choice				Tax per Capita
	L = 0	L = ¼	L = ½	L = 1			L = 0	L = ¼	L = ½	L = 1	
50,000	0	0.45	0.25	0.30	4,625		0	0.45	0.25	0.30	5,781
100,000	0	0.45	0.30	0.25	9,750		0	0.45	0.30	0.25	12,188
150,000	0	0.45	0.30	0.25	15,188		0	0.45	0.30	0.25	18,281
200,000	0.05	0.40	0.35	0.20	23,500		0.05	0.45	0.30	0.20	26,875
250,000	0.05	0.40	0.35	0.20	31,375		0.05	0.45	0.30	0.20	33,594
300,000	0.05	0.25	0.55	0.15	37,625		0.10	0.30	0.45	0.15	41,250
350,000	0.10	0.25	0.50	0.15	47,938		0.10	0.30	0.45	0.15	48,125
400,000	0.10	0.25	0.50	0.15	56,000		0.10	0.25	0.50	0.15	53,750