The Effect of Income Taxation on Consumption and Labor Supply

James P. Ziliak, *University of Kentucky*

Thomas J. Kniesner, *Syracuse University*

We estimate the incentive effects of income taxation in a life-cycle model of consumption and labor supply without intratemporal strong separability. We find that consumption and hours worked are direct complements in utility; both increase with a compensated increase in the net wage. The compensated net wage elasticity is about 0.3, nearly double estimates for U.S. men from a linear labor supply specification. Estimated intertemporal elasticities indicate significant intertemporal smoothing of utility. The estimated marginal welfare cost of government revenue is 6%–20%, which is about half the estimated welfare cost when additivity between consumption and leisure is incorrectly imposed.

I. Introduction

Estimating the effect of income taxes on the labor supply has been a focal point of research by labor and public economists for over 3 decades (Pencavel 1986; Blundell and MaCurdy 1999). The keen economic interest

We thank Tim Besley, Nada Eissa, Bill Gale, Jerry Hausman, Tom MaCurdy, Kevin Reilly, Art Snow, Ron Warren, Dave Wildasin, and seminar participants at the 2005 American Economic Association meetings, NBER Trans-Atlantic Public Economics Workshop, the eighth annual meeting of the Society of Labor Economists, the University of Georgia, Iowa State University, the University of Kentucky, and the University of Tennessee for many helpful comments. Special thanks go to Doug Holtz-Eakin, who was a coauthor on early versions of this project but whose administrative duties at the Congressional Budget Office precluded further collaboration. All errors are our own, and nothing contained herein rep-

[*Journal of Labor Economics,* 2005, vol. 23, no. 4]

0734-306X/2005/2304-0004$10.00

stems from the well-established result that the deadweight loss from reduced incentives to work is increasing in the progressivity of the tax code (Hausman 1981; Auerbach 1985; Auerbach and Slemrod 1997; Blundell, Duncan, and Meghir 1998; Carroll et al. 1999; Ziliak and Kniesner 1999). However, there has been much disagreement over the years on the magnitude (and sometimes even the sign) of compensated wage effects—a positive compensated wage effect means that moving to a revenue-neutral flatter income tax induces more hours worked and reduces deadweight loss. Moreover, much of the research on labor supply and taxation has been conducted with static models on cross-sectional data (recent exceptions include Blundell et al. 1998; Ziliak and Kniesner 1999), and all previous empirical work on taxes and labor supply in a life-cycle setting maintains the assumption of additive separability between consumption and leisure. A more complete understanding of the economic implications of tax reform requires an evaluation of income taxation in a more flexible framework that admits interactions among consumption and leisure choices over time. We exploit the natural experiments of the tax reforms of the 1980s and 1990s in the United States to examine empirically the joint effect of income taxes on life-cycle consumption and labor supply.

The interest in identifying the impact of income taxes on labor supply was renewed in the 1990s when MaCurdy, Green, and Paarsch (1990) challenged the seminal econometric framework of Hausman (1981), who had modeled and estimated via maximum likelihood the intricacies of the piecewise linear budget set facing the worker by a simultaneous choice of segment (kink) location and hours of work. Hausman's estimates suggested that the deadweight loss of income taxation was sizable, which provided the intellectual foundation for the 1980s tax reforms. MaCurdy et al. (1990) argued that the internal consistency of Hausman's model required an upward-sloping labor supply schedule, and, upon the relaxation of some key assumptions by smoothing the budget set, the previously accepted result of a vertical or backward-bending male labor supply schedule reappeared. Ziliak and Kniesner (1999) extended the single-period linear model to the life-cycle case and found estimates closer to Hausman's, with a compensated wage elasticity ranging from 0.13 to 0.18 across wealth quartiles. The implied life-cycle deadweight loss from the 1980s U.S. income tax structure was about 20% of current income.

Unlike the case of labor supply, there is comparatively little research on how income taxes affect consumption expenditures, either independently or in conjunction with labor supply choices. Most of this work has focused on the consumption-smoothing aspects of distortionary income taxation (Varian 1980; Strawczynski 1998; Auerbach and Feenberg

resents the official positions of the University of Kentucky or Syracuse University. Contact the corresponding author, James P. Ziliak, at jziliak@uky.edu.

2000; Kniesner and Ziliak 2002; Low and Maldoom 2004). Empirical work addressing the effects of income taxes on labor supply in a framework that simultaneously models the consumption decision has been nonexistent. Other empirical research has relaxed and rejected within-period separability between consumption and leisure in the contexts of a conditional demand model (Browning and Meghir 1991; Blundell, Browning, and Meghir 1994), habit formation (Hotz, Kydland, and Sedlacek 1988), and endogenous human capital (Shaw 1989), but the research has not been concerned with income tax effects.[1] Obtaining estimates of labor-supply tax effects in the context of a flexible framework with consumption is critical to more informed tax policy, especially in light of major reforms to the U.S. tax system over the past 2 decades and recent procedural changes adopted by the Congressional Budget Office to score tax revenue effects dynamically.

Research examining the connections among taxes, consumption, and labor supply is of further interest in light of the burgeoning macroeconomics literature on precautionary saving. In aggregate data current consumption tracks current income closely, contrary to the standard life-cycle permanent income model of consumption (Carroll and Summers 1991). To explain the apparent excess sensitivity puzzle, some researchers have turned to alternative models with impatient consumers and buffer-stock saving (Deaton 1991; Carroll 1997). The recent macroeconomics literature has ignored the possibility that in the face of unanticipated wage changes households may alter their labor supply choices over time to accumulate precautionary balances instead of forgoing current consumption if consumption and leisure are direct substitutes (Low 1999). The potential importance of labor supply was first noted by Heckman (1974) in a deterministic setting, highlighting the fact that consumption tracking income may arise out of anticipated wage changes as well as uncertain wage changes.

We extend the labor supply and taxation literature by estimating a life-cycle model of consumption and labor supply under uncertainty with nonlinear wage income taxation and relaxing the standard assumption of

[1] Pistaferri (2003) is a recent exception. Using Italian data, he failed to reject the null hypothesis of additive separability between consumption and leisure within the context of a life-cycle labor supply model without income taxes. Pistaferri urges caution in interpreting his result because "we are using an unsophisticated approximation to individuals' preferences for consumption and leisure. . . . In light of the large standard errors I do not wish to put too much emphasis on this result" (745). In his test Pistaferri (2003) did not explicitly rely on consumption data as we do in this project, and thus our model should be a more robust framework for examining the interactions of consumption and leisure. In a model without income taxes, Altonji (1986) finds that food expenditures and leisure are substitutes, consistent with our findings, but his estimates are inefficiently estimated such that he cannot reject the null of separability.

strong separability in consumption and labor supply choices within periods. Unlike the conditional demand literature, we estimate within-period preferences over both consumption and labor supply via the marginal rate of substitution function and a direct translog felicity function. We then estimate intertemporal preference parameters using the Euler equation governing the first-order condition for the evolution of discounted marginal utility of wealth under uncertainty. Demographics enter the model through so-called demographic translating, which means that demographic variables directly affect the parameters governing utility (Pollak and Wales 1992). The combination of within-period preferences and intertemporal preferences permits us to identify compensated and uncompensated net wage elasticities, as well as intertemporal substitution elasticities. Although uncertainty is permitted in our framework, we do not attempt to quantify the responses of consumption and labor supply to uncertain wage and tax changes and instead focus on anticipated changes (Altonji and Ham 1990; Pistaferri 2003). Because of the endogeneity of regressors in both the first and second stages of the two-stage budgeting model, we use a generalized method-of-moments estimator (Hansen 1982).

We employ data on male heads of household from the 1980–99 waves of the Panel Study of Income Dynamics, which spans the major recent federal tax reforms in the United States from the Economic Recovery Tax Act of 1981 to the Taxpayer Relief Act of 1997. Our results indicate that consumption and hours worked are direct complements in utility, and both increase with a compensated increase in the net wage. The compensated net wage elasticity is about 0.3, which is nearly double the typical estimate for U.S. men based on a linear labor supply specification. Given our estimated intertemporal elasticity of substitution of about -1.0, the Frisch specific substitution elasticity of consumption with respect to the after-tax wage is about 0.1, and the corresponding Frisch elasticity of labor supply is about 0.5. We conclude by relating our estimated within-period preference parameters to the static marginal welfare cost of government revenue. We find that the marginal welfare cost of an additional dollar of revenue raised is upward of 20%, which is roughly half the estimated welfare cost if one incorrectly imposes strong separability between consumption and leisure ex ante.

II. A Model of Life-Cycle Consumption and Labor Supply

The model of life-cycle consumption and labor supply we adopt is standard in that the consumer is assumed to choose consumption and hours of work optimally to maximize the present discounted value of uncertain utility subject to an asset accumulation constraint (MaCurdy 1983). Uncertainty arises because of the unknown paths of future wages, prices, taxes, and interest rates. Intertemporal preferences are assumed to be time separable, as are budgets, which rules out preference dependence over time due to habits (Hotz et al. 1988) and rules out nonseparabilities in the budget constraint due to possible endogenous human capital and joint nonlinear taxation of wage and capital incomes (Blomquist 1985; Shaw 1989; Ziliak and Kniesner 1999).[2] We do permit nonseparabilities in within-period preferences over consumption and labor supply, which are chosen freely.[3] Added endogeneity of labor supply permits direct, unconditional assessment of the effects of wages and taxes on both margins, which is not possible in the conditional consumption demand framework.

A. Basic Theoretical Setup

The value function governing the representative household's decision problem is

$$\begin{aligned} V^t(A_t) = \max_{C_t, h_t} \{ & G[U(C_t, \bar{L} - h_t)] \\ & + \beta E_t V^{t+1}[(1 + r_t)(A_t + w_t h_t - p_t C_t - T_t(I_t))]\}, \end{aligned} \tag{1}$$

where A_t is the beginning of period t assets, $U(\cdot)$ is the within-period felicity function, and $G[\cdot]$ is a monotonic transformation of within-period preferences that governs intertemporal preferences. The variable C_t is composite nondurable consumption, $\bar{L}$ is total time available, h_t is annual hours of work, $\beta = 1/(1 + \rho)$ is the time discount rate, E_t is the expectations operator conditional on the information set at time t, r_t is a risk-free interest rate, w_t is the gross hourly wage rate, p_t is the price index on nondurable consumption, and $T_t(\cdot)$ is the household's income tax liability as a function of taxable income, $I_t = w_t h_t + N_t - D_t - Ex_t$, which is gross labor income plus property income (N_t) less deductions (D_t) and exemptions (Ex_t).[4] We assume that both the utility function and the tax function are twice continuously differentiable. Finally, we normalize by

[2] Although we ignore the budget nonseparabilities generated by joint nonlinear taxation of capital and labor, we include property income in taxable income in an attempt to measure marginal tax rates more accurately.

[3] Alternative approaches are Browning and Meghir (1991) and Blundell et al. (1994), who model consumption decisions within the context of a conditional (on labor supply) demand framework. Altonji (1986) assumes within-period separability in consumption and leisure but then tests separability by approximating nonseparability by appending the λ-constant equations with cross-substitution terms.

[4] As mentioned in n. 2 and the text, the actual measure of taxable income used to construct tax payments and marginal tax rates includes property income, but we do not address the modeling complications that arise from joint nonlinear taxation of labor and capital income.

the price of consumption so that wages and interest rates are in real terms. The value function V^{t+1} is unknown as of time t because future realizations of the function's arguments are uncertain.

The first-order conditions for consumption and hours from maximizing the value function are

$$E_t[G'U_{C,t} - \beta(1 + r_t)\lambda_A^{t+1}] = 0, \tag{2}$$

$$E_t[-G'U_{h,t} + \beta(1 + r_t)w_t(1 - \tau_t)\lambda_A^{t+1}] = 0, \tag{3}$$

and

$$\lambda_A^t = \beta E_t[(1 + r_t)\lambda_A^{t+1}], \tag{4}$$

where G' is the first derivative of the intertemporal transformation function, $U_{C,t}$ is the first derivative of within-period utility with respect to consumption, $U_{h,t}$ is the first derivative of utility with respect to hours of work, $\tau_t = \partial T_t(\cdot)/\partial h_t$ is the marginal tax rate (MTR), and $\lambda_A^{t+1} = \partial V^{t+1}/\partial A_{t+1}$ is the marginal utility of wealth.

Substituting for λ_A^{t+1} in equation (3) using equation (2) and known time t values yields the familiar first-order condition for an interior solution, which equates the marginal rate of substitution (MRS) of hours for consumption to the after-tax wage rate, $\omega_t \equiv w_t(1 - \tau_t)$,

$$-U_{h,t}/U_{C,t} = \omega_t. \tag{5}$$

It is clear from equation (5) that the monotonic transformation $G[\cdot]$ plays no role in determining within-period consumption and hours allocations, so that cross-sectional data are sufficient to identify intratemporal preferences (MaCurdy 1983; Altonji 1986). To identify intertemporal preferences, it is necessary to have panel data (or time-series or pseudo-panel data) and the Euler equation (4) governing the allocation of wealth over time.

Most of the literature on life-cycle labor supply (MaCurdy 1981; Pistaferri 2003) and life-cycle consumption, including tests of full risk sharing, of precautionary saving, and of the permanent-income hypothesis (e.g., Hall and Mishkin 1982; Cochrane 1991; Deaton 1991; Ogaki and Qiang 2001), restrict intra- and intertemporal preference parameters to be the same. An ex ante restriction that intra- and intertemporal preference parameters be the same is costly in terms of reduced flexibility of behavioral responses to wage, price, and interest rate changes (Browning 1985).

To elaborate on the importance of maximum preference function flexibility, a familiar parameter in life-cycle models of consumption is the intertemporal substitution elasticity (ISE), which is the proportional change in consumption expenditure needed to keep the marginal utility of wealth constant given an anticipated 1% change in prices. Under the standard model with time-additive preferences, the intertemporal substitution elasticity is minus the inverse of the coefficient of relative risk aversion, $U_C/(CU_{CC})$. Given the monotonic transformation in equation (1), the ISE is $U_C/\{C[U_{CC} + (G''/G')U_C^2]\}$, which will vary based on the choice of the function for G (Browning 1985). Moreover, consider the Frisch (marginal utility of wealth constant) specific-substitution elasticity between any two goods j and k,

$$e_{jk}^F = e_{jk}^U + e_j e_k s_k \Phi, \tag{6}$$

where e_{jk}^F is the Frisch elasticity, e_{jk}^U is the compensated cross-price elasticity, e_j and e_k are expenditure (income) elasticities, s_k is the share of good k in the household budget, and Φ is the ISE. If G is the identity transformation, and within-period preferences are additive, then $e_{jk}^F = e_j\Phi \approx e_{jk}^Y$, where e_{jk}^Y is the income-constant Marshallian cross-price elasticity of demand. The dual assumptions that within-period preferences are additive and transform exactly into intertemporal preferences are not innocuous, as they imply that the path of consumption is independent of the path of wages, regardless of whether wage changes are anticipated (Heckman 1974) or unanticipated (Low 1999).

B. A Tractable Empirical Representation

Our empirical strategy is to adopt the two-stage estimation method of MaCurdy (1983) where in the first stage we estimate the intratemporal equilibrium condition in equation (5) by specifying a functional form for within-period preferences that permits nonseparabilities between consumption and labor supply choices. Given the estimated within-period preference parameters, we construct the period-specific utility functions to estimate the intertemporal preference parameters from the Euler equation (4).

We specify within-period preferences with a direct translog felicity function

$$\begin{aligned} U(C, \bar{L} - h) = {} & \alpha_1 \ln(\bar{L} - h) + \alpha_2 \ln C \\ & - \alpha_3 \ln(\bar{L} - h) \ln C - \alpha_4 \ln(\bar{L} - h)^2 - \alpha_5 \ln C^2, \end{aligned} \tag{7}$$

which is a local second-order approximation to any arbitrary utility function (Christensen, Jorgensen, and Lau 1975). Important for our purposes is that the direct translog does not impose additivity between consumption and leisure—a positive coefficient on α_3 implies that consumption and leisure are direct substitutes, or that consumption and work hours are direct complements. Identification requires a normalization. We chose $\alpha_5 = 1$. Demographics are introduced into the model via the method of demographic translating whereby the utility parameters are explicit functions of demographic characteristics (x_{jt}), such that $\alpha_j = \alpha_{j0} + \sum_{k=1}^{K} \alpha_{jk} x_k, j = 1, .., 4$ (Pollak and Wales 1992). Basing our estimation on a demographically translated direct translog specification of intratemporal preferences, we then estimate the MRS condition in equation (5) as

$$\{[-\alpha_1 + \alpha_3 \ln C + 2\alpha_4 \ln(\bar{L} - h)]/(\bar{L} - h)\} - \omega\{[\alpha_2 - \alpha_3 \ln(\bar{L} - h) - 2 \ln C]/C\} + \varepsilon = 0, \tag{8}$$

where ε reflects unobserved idiosyncratic tastes.

For the monotonic transformation G we specify preferences as

$$G[U(C_t, \bar{L} - h_t)] = \frac{[U(C_t, \bar{L} - h_t)]^{1+\sigma} - 1}{1 + \sigma}, \tag{9}$$

where $\sigma = \sigma_0 + \sum_{j=1}^{J} \sigma_j d_{jt}$ are the intertemporal preference parameters permitting variation in risk aversion and the ISE according to time-varying demographic characteristics, d_{jt}.[5] Combining the first-order condition for consumption (2) with equation (4) that governs the evolution of the marginal utility of wealth, taking expectations and natural logs, and then first differencing, the parameterization in (9) yields the estimating equation

$$\sigma_0 \Delta \ln \hat{U}_{t+1} + \sum_j \sigma_j \Delta(d_{j,t+1} \ln \hat{U}_{t+1}) + \Delta \ln \hat{U}_{C,t+1} + \kappa_{t+1} = \nu_{t+1}, \tag{10}$$

where Δ is the first difference operator, $\hat{U}_{t+1}$ and $\hat{U}_{C,t+1}$ are the estimated values of utility and marginal utility found by replacing α_j with $\hat{\alpha}_j$ in equations (7) and (8), $\kappa_{t+1} = r_{t+1} + (\theta_t - \rho)$, $\theta_t = -E_t(\ln \zeta_{t+1})$, and $\ln \zeta_{t+1}$ is the time t forecast error uncorrelated with the model's variables. In deriving equation (10) we exploit the approximations $\ln(1 + r_{t+1}) \approx r_{t+1}$ and $\ln(1 + \rho) \approx \rho$. If ζ_{t+1} is lognormally distributed, then $\theta_t = (1/2)\psi_t^2$, where ψ_t^2 is the variance of $\ln \zeta_{t+1}$, and $(\theta_t - \rho)$ captures the trade-off between impatience and caution, which is a key parameter in determining the extent of precautionary saving in augmented life-cycle models with precautionary motives (Blundell et al. 1994). The demographics affecting the MRS equation, x_k, need not be time varying but demographics affecting intertemporal risk, d_j, must change over time, as indicated in equation (10), to have their effects identified separately from the constant term σ_0.

[5] Ogaki and Zhang (2001) show that introducing a subsistence consumption level into constant relative risk aversion preferences permits increasing, decreasing, and constant relative risk aversion and that the flexibility of risk tolerance affects tests of complete consumption insurance. We experimented with permitting a threshold level of utility in the G[$\cdot$] transformation, but the threshold parameter was not statistically significant and often created problems with convergence. MaCurdy (1983) reported similar difficulties.

III. Data and Estimation Issues

To identify the tax effects on work incentives and consumption, we use household-level data on male heads of household from the 1980–99 waves of the Panel Study of Income Dynamics (PSID; 1979–98 calendar years). The survey has followed a core set of households since 1968 plus newly formed households as members of the original core have split off into new families. Following the 1997 survey year, the PSID began interviewing households every other year, so there are no data for the 1997 calendar year. The PSID is advantageous because it contains detailed information on income and household composition and, after 1979, more detailed tax-related data. Our data are additionally desirable because they span multiple tax reforms in the United States: the Economic Recovery Tax Act of 1981, the Tax Reform Act of 1986, the Omnibus Reconciliation Tax Acts of 1990 and 1993, and the Taxpayer Relief Act of 1997. Together, the tax reforms of 1981 and 1986 reduced marginal tax rates across the board, reduced the number of tax brackets from 16 to four, and expanded the taxable income base. Although the tax reforms of the 1990s reversed the trend of the 1980s reforms by adding two new higher marginal tax rates on upper-income Americans, the tax reforms of the 1990s also significantly expanded the earned income tax credit among low-income working families.

A. Estimation Sample Details

The sample we use in estimation is an unbalanced panel treating missing observations as random events. By eliminating only a missing person-year of data, the time series for each household can be of different length within 1980–99. To be included in the sample, the household head must be (1) a male, (2) in the sample at least 5 years, (3) at least 25 years old in 1980 and no older than 60 in the last year in the sample, and (4) not a student, retired, permanently disabled, or institutionalized. Focusing on prime-age male heads of household allows us to ignore issues associated with labor force nonparticipation. To reduce further the influence of possible outliers, we follow the existing literature and delete person-years with more than a 300% increase or more than a 75% decrease in consumption and family income from the previous year. We also require annual nominal food expenditures (inclusive of food stamps) to be no less than \$520, annual real total expenditures (described below) to be no less than \$800, and annual real family income to be no less than \$1,000. Using our sample filters, we obtain 3,402 household heads in the 19-year sample. Because we require households to be present for 5 years, and because we invoke more detailed filters such as missing-data codes and extreme consumption and income changes, we retain 21,186 household-years for econometric estimation. All wage, price, income, and consumption expenditure data are deflated by the personal consumption expenditure deflator with 1998 base year.

The focal variables in the models in equations (8) and (10) are consumption expenditures, labor supply, real wage rates, taxable income, marginal tax rates, total tax payments, interest rates, and demographics. We measure consumption as total nondurable consumption expenditures. The PSID collects food expenditures only on an annual basis, and it did not collect food expenditures information in the 1988 and 1989 surveys. Blundell, Pistaferri, and Preston (2001) recently proposed a method of imputing nondurable expenditures in the PSID. Using data from the Consumer Expenditure Survey (CEX), they estimated the demand for food at home as a function of measured demographics (available in both the PSID and CEX), food prices, and nondurable expenditures. The model is

$$\ln(c_{it}^{f}) = X_{it}\varphi + \pi \ln(C_{it}) + e_{it}, \tag{11}$$

where c_{it}^{f} is food expenditures in the home and C_{it} is nondurable expenditures. Given estimates $\hat{\varphi}$ and $\hat{\pi}$ from the CEX, along with data on food and demographics in the PSID, it is possible to predict nondurable consumption as $\ln(\hat{C}_{it}) = (\ln(c_{it}^{f}) - X_{it}\hat{\varphi})/\hat{\pi}$.[6] Provided that food expenditures are monotonic in nondurable expenditures, that the point estimates from the CEX are estimated consistently, and that the trends in the variance of nondurable consumption are the same across the CEX and PSID, using equation (11) produces a consistent estimate of nondurable expenditures in the PSID.[7] Browning, Crossley, and Weber (2003) recently argued that imputation methods may be a fruitful approach to dealing with limited consumption data. As a sensitivity check on the model, we also present estimates based on food expenditures and an alternative imputed measure of nondurables consumption proposed by Skinner (1987).

Labor supply here is defined as annual hours of work from all jobs. For workers paid by the hour, the survey records the gross hourly wage rate. Given that the data after 1993 are still in the early release form, the

[6] We use a scaled-down version of the prediction equation appearing in table 4 of Blundell et al. (2001). We predict nondurable expenditures as $\ln(\hat{C}_{it}) = (\ln c_{it}^{f}) - (3.6674 - 0.5746 \ln(cpi_{t}^{f}))/0.4573$. We are grateful to Luigi Pistaferri for providing us with the necessary information.

[7] A related method of predicted nondurable consumption in the PSID appears in Skinner (1987). He, too, used data from the CEX, but many of the variables needed to construct the broadest version of Skinner's measure are no longer collected by the PSID. We use a variant of Skinner's approach in the robustness section below. Ziliak (1998) proposed a method of imputing total consumption in the PSID by netting out saving from income where it is necessary first to predict saving using wealth information in the PSID. Inferring consumption from saving measured by changes in wealth requires an additional year of data for each household to construct saving and maybe a noisier measure of consumption.

hourly wage is missing for many observations in certain years. We then follow a procedure akin to the PSID's calculation of hourly wages for salaried workers. For workers with annual hours less than 1,000, we divide annual earnings by 750; for workers with hours between 1,000 and 1,800 we divide earnings by 1,500; for workers with hours between 1,800 and 2,200, we divide earnings by 2,000; and for workers with more than 2,200 hours, we divide earnings by 2,400. Dividing earnings by standardized work-years reduces so-called division bias that plagues wages computed as the ratio of annual earnings to actual annual hours (Borjas 1981; Ziliak and Kniesner 1999).

When constructing annual taxable income, we assume that married men filed joint tax returns and unmarried men filed as head of household.[8] Adjusted gross income (AGI) is the sum of labor earnings, cash transfers, and property income. To approximate the actual marginal tax rate facing the household, we include property income in AGI, inclusive of wife's earnings in cases where married men have working wives. However, as mentioned in the model section, for tractability we abstract from the fact that an inclusive property income measure may generate nonseparabilities both within periods in spousal labor supply choices and across periods in intertemporal labor supply, as in Ziliak and Kniesner (1999). Taxable income is adjusted gross income less deductions and exemptions. The PSID provides the number of tax exemptions for de-

[8] The assumption that all unmarried men file as heads of household regardless of the presence or absence of dependents is for tractability. We selected two tax years, 1985 and 1995, to gauge the potential scope of tax misclassification and the possible implications for the estimates described below. In 1985, 106 out of 1,200 men are single with no dependents. We calculate that we misclassify about 70 single men, or only about 6% of the total number of observations in 1985. For 1995, we likely misclassify about 21% of single, childless men, or about 1%–2% of the total sample for that year. Misclassifying occurs more often in 1985 because there were 16 tax brackets in 1985 as opposed to six brackets in 1995, and the head of household rate brackets rose with a lag behind the single individual tax brackets. Although misclassifying occurs more often in the mid-1980s, the economic consequences are likely to be less than in the mid-1990s because the rates increase more sharply after TRA86. To assess how important any filing status misclassification might be for our estimated compensated wage elasticity, we performed two experiments. First, we assumed that we underestimated the federal MTR for single men without dependents by 10% and then increased the individuals' federal MTRs by 10%. Second, we assumed that we understated the MTR by 50% and then increased the individuals' federal MTRs by 50%. The baseline compensated wage elasticity of labor supply is 0.328. When we make the 10% MTR correction the elasticity remains unchanged. When we make the 50% MTR correction the elasticity falls to 0.32, which is only 3% below the baseline. In other words, the economic consequences of our tractability assumption are negligible because the MTRs will not be off by 50%.

pendents taken in each year, but how we calculate deductions requires additional explanation.

Computing the value of deductions depends on the year under consideration. To evaluate annual deductions prior to and including 1983, we follow the convention established in the PSID. With information from the Internal Revenue Service's *Statistics of Income*, we generate the typical value of itemized deductions based on adjusted gross income. We then calculate the difference between typical itemized deductions and the standard deduction, known as excess itemized deductions. For the years prior to and including 1983, when excess itemized deductions are positive we subtract them from adjusted gross income; when excess itemized deductions are nonpositive, we apply the standard deduction.

Beginning in 1984, the PSID records whether the family itemized. For known itemizers we subtract excess itemized deductions from adjusted gross income and use the standard deduction for the men who did not itemize deductions. Prior to the Tax Reform Act of 1986 (TRA86) the standard deduction was built into the tax tables; we need only subtract the value of deductions exceeding the standard deduction from taxable income. After TRA86 the standard deduction is no longer built into the tax tables, so we subtract either the standard deduction or total itemized deductions from adjusted gross income depending on whether the family itemized.

We approximate the income tax liability via several steps. Using a method derived by MaCurdy et al. (1990) and implemented by Ziliak and Kniesner (1999), we approximate federal income tax payments with a smooth cubic polynomial in taxable income. The idea is to act as if the household faces a smooth tax function, rather than a piecewise-linear function, and use the smooth tax function to approximate the marginal tax rate. Because the marginal tax rate is also a smooth and continuously differentiable function of taxable income, we can integrate the function back to obtain total tax payments. From total federal tax payments we net out the imputed earned income tax credit for each year (assuming a 100% take-up rate) and add in Federal Insurance Contribution Act (FICA payroll) taxes and the relevant state income tax payments, which for tractability we take as a proportional tax on income with the tax rate determined by the average income tax rate in the state (U.S. Department of Commerce, Bureau of the Census 1980–99).[9] An alternative to our approach is to use the marginal tax rate and tax payments constructed by the PSID. The shortcoming with the PSID tax rates for our purposes is that they stopped calculating household income tax liability after the 1991 interview year. Our tax imputation method produces tax rates that coincide well with the PSID's tax rates in the years that permit a comparison.

[9] Details of the tax calculations are available from the authors upon request.

Last, for the demographics moderating the parameters α_j in the MRS equation (8) we use a parsimonious specification with the number of children in the household, the race of the male head, and the age of the youngest child. To maintain tractability we admit only the demographics in α_1 and α_2, assuming the remaining two parameters are homogeneous across the sample. The parallel demographics that affect risk aversion and the ISE are the age of the household head and the health status of the head. Appendix table A1 contains selected summary statistics for the variables used in our econometric model.

B. Econometric Issues

Estimations of the MRS equation (5) and the Euler equation (4) are complicated both because the models are nonlinear in the parameters and because the regressors are endogenous (hours, consumption, and wages in the MRS equation and utility in the Euler equation).[10] Although the empirical counterparts in equations (8) and (10) are linear functions of parameters, we still must address endogeneity. It is possible to rearrange equations (8) and (10) into a linear instrumental variables framework by using the normalization $\alpha_5 = 1$ in equation (8) to make $-2\omega \ln(C)/C$ the left-hand-side variable and using the change in marginal utilities ($\Delta \ln \hat{U}_{C,t+1}$) as the left-hand-side variable in Euler equation (10). The particular instrumental variable estimator we adopt is the generalized method of moments (GMM) estimator (Hansen 1982). Given a ($1 \times Q$) vector of instrumental variables for the MRS equation, z_{it}, the population orthogonality conditions we estimate for the first stage are $E(z_{it}'\varepsilon_{it}) = 0$. The analogous conditions for equation (10) are $E(m_{it}'\nu_{it+1}) = 0$, where m_{it} is a ($1 \times M$) vector of instrumental variables. The two-stage GMM estimator we employ admits conditional heteroskedasticity where in the first stage we estimate equations (8) and (10) via the two-stage least-squares method (2SLS) and use the estimated residuals to form the second-stage optimal weight matrix for the GMM estimator.

In selecting instrumental variables for the MRS equation, we assume that ε is not autocorrelated but may be conditionally heteroskedastic. We use as instruments a constant and the $(t-1)$ values of the head's age, the family size, the number of kids, the age of the youngest child, and dummy indicators for marital status, education, race, self-employment status, health status, home ownership, union status, industry, occupation, and region of country. For the Euler equation (10) we use the $(t-2)$ values of the variables in the MRS instrument set along with time dummies and

[10] Endogeneity is not unique to the MRS-Euler equation estimation approach; a model that estimates consumption or labor supply directly still needs to address the issue of wage endogeneity (Altonji 1986).

twice lagged real after-tax wages, nondurable expenditures (or food expenditures), and hours of work.

Because the Euler equation (10) is a function of estimated parameters from the first stage, it is necessary to correct the second-stage standard errors for the additional sampling variation. Although asymptotic approximations to the variance-covariance matrix of sequential method-of-moments estimators are available (Newey and McFadden 1994), we instead use bootstrapping to construct the second-stage standard errors. The typical regression-based bootstrap is a multistep procedure whereby the researcher resamples with replacement the estimated residuals, constructs a new dependent variable as the sum of the fitted value from the regression plus the bootstrapped residual, reestimates the model, and repeats the exercise B times ($b = 1, .., B$). There are then B observations from which to compute measures of bias, variability, or confidence intervals. The basic bootstrap approach is consistent under the assumptions of conditional homoskedasticity, no serial dependence, and nonstochastic regressors.

When the regressors are stochastic or there is conditional heteroskedasticity, as is typical in instrumental variables (IV) estimation, Freedman (1984) suggests an alternative procedure. Instead of resampling the residuals, one resamples simultaneously the estimated residuals along with the regressors and instruments. More specifically, one resamples with replacement from $(\hat{\nu}, \hat{P}, m)$, where $\hat{\nu}$ is the Euler equation residual, $\hat{P}$ is the matrix of regressors in the Euler equation, and m is the matrix of instruments. Call the constructed information bootstrap data $(\hat{\nu}^*, \hat{P}^*, m^*)$. One then constructs the new dependent variable, $\Delta \ln \hat{U}^*_{C,t+1}$, from the bootstrap data (the bootstrapped residuals and accompanying regressors), which is in turn reestimated with the accompanying instruments, m^*. Defining the vector of bootstrapped parameters estimates as $\hat{\delta}_b$, the bootstrap standard error is $\{[\sum_{b=1}^{B} (\hat{\delta}_b - \frac{1}{B} \sum_{b=1}^{B} \hat{\delta}_b)^2] / (B-1)\}^{1/2}$. We set B equal to 1,000 replications. The multistage approach, in which each observation has equal probability weight 1/N of being drawn from the discrete empirical distribution function, is an asymptotically valid method of bootstrapping an IV estimator and offers efficiency gains over first-order asymptotics (Hall and Horowitz 1996; Ziliak 1997).

IV. Results

In table 1 we record the estimates of both the intratemporal preferences from the MRS equation (8) and the intertemporal preferences from the Euler equation (10). We set the value of total time, $\bar{L}$, equal to the number of hours in a year (8,760).

The estimates in table 1 show that the marginal rate of substitution between hours of work and consumption is increasing in the number of

Table 1
GMM Estimates of Intratemporal and Intertemporal Preference Parameters

Direct Translog MRS Preference Parameters (Eq. [8])		Euler Equation Preference Parameters (Eq. [10])	
Variable:			
α_1 (constant)	77.496 (28.676)	Constant	.844 (.230) {.214}
α_1 (number of kids)	3.035 (.784)	$\theta_t - \rho$	.070 (.022) {.024}
α_1 (race = 1 if white)	8.877 (1.922)	Age	−.039 (.035) {.0004}
α_1 (age of youngest child)	1.470 (.254)	Health (work limited = 1)	.006 (.030) {.035}
α_2 (constant)	51.407 (3.954)		
α_2 (number of kids)	.822 (.299)		
α_2 (race = 1 if white)	2.285 (.425)		
α_2 (age of youngest child)	.499 (.088)		
α_3	4.263 (.434)		
α_4	3.085 (1.573)		
Sargan [df]	110 [28]		138 [55]

Note.—Asymptotic SEs corrected for conditional heteroskedasticity are in parentheses. Bootstrap SEs from 1,000 replications are reported in braces. The number of observations is 21,186 person years.

children and in the age of the youngest child and is larger for white men. Ceteris paribus, labor supply is then higher for men with more children, higher for men with older children relative to men with younger (or no) children, and higher for white men relative to nonwhite men. The parameter governing the within-period relationship between consumption and work hours, α_3, is positive and statistically different from zero, which implies that consumption and leisure hours are direct substitutes in utility. We explore the implications of the inverse dependence between consumption and leisure choices below.

Although the p-value from the Sargan test of the validity of the overidentifying restrictions in the first-stage 2SLS does not reject our model specification, the test statistic from the second-stage GMM model reported in table 1 indicates possible model misspecification owing to invalid instruments. As one check on our instrument set, we replaced the initial set of instruments with their corresponding values at $(t-2)$, but we obtained equally weak test results. It is important to note that the GMM Sargan test based on a relatively large number of moment conditions is poorly sized and tends to overreject (Hall and Horowitz 1996; Ziliak 1997). Given that the 2SLS version of the Sargan test does not reject the overidentifying conditions and that the GMM variant is poorly sized, we have reasonable confidence in our instrument choice.

In the second column of table 1 we record the estimates of the Euler equation for nondurable expenditures.[11] The estimate of $(\theta_t - \rho)$ equals 0.07, suggesting that prudence outweighs impatience and that precautionary saving motives are present. The nondurable consumption Euler equation model suggests (weakly) that risk aversion is declining with age but that risk preferences are not affected economically or statistically by health-induced work limitations.

A. Intra- and Intertemporal Elasticities

It is informative to characterize the estimates in table 1 into terms useful for labor-market and tax policy; namely, compensated and uncompensated wage elasticities for within-period preferences and the ISE and Frisch specific substitution elasticities for intertemporal preferences. When closed-form solutions for within-period demand and supply functions are not available, MaCurdy (1983) observed that it is still possible to derive the implied compensated and uncompensated wage effects by exploiting a result in Phlips (1974) known as the fundamental matrix equation. We follow the fundamental matrix equation method closely and summarize it here for completeness.

Ignoring for the time being the monotonic transformation, $G[.]$, define the Hessian matrix for the utility function as H and the marginal utility of income as $\mu = U_{C,t}/p_t$. Furthermore, define the price vector of interest as $q' \equiv (p_t, \omega_{it})$, where p_t is the price of consumption normalized to 1 and ω_{it} is the real after-tax wage rate. The implied income effects, compensated effects, and uncompensated effects are

$$\begin{pmatrix} \partial C/\partial Y \\ -\partial h/\partial Y \end{pmatrix} = \frac{1}{n} H^{-1} q,$$

$$\begin{pmatrix} \partial C/\partial q'|_U \\ -\partial h/\partial q'|_U \end{pmatrix} = \mu H^{-1} - \frac{\mu}{n} H^{-1} q q' H^{-1}, \tag{12}$$

$$\begin{pmatrix} \partial C/\partial \omega|_Y \\ -\partial h/\partial \omega|_Y \end{pmatrix} = \begin{pmatrix} \partial C/\partial \omega|_U \\ -\partial h/\partial \omega|_U \end{pmatrix} + \begin{pmatrix} \partial C/\partial Y \\ -\partial h/\partial Y \end{pmatrix} h,$$

[11] There is an unintended by-product of the flexibility of the direct translog utility function. The marginal utilities of consumption and leisure are not restricted to be positive for all observations, which creates obvious problems when we take the log of the marginal utility of consumption for the second-stage Euler equation. In cases with nonpositive marginal utilities we assumed that the person-years contribute nothing to intertemporal substitution and set the difference in log marginal utilities of these observations to zero.

Table 2
Selected Intratemporal and Intertemporal Elasticities

Changes in Real After-Tax Wages (ω_t)	Consumption	Labor Supply
Income elasticity	.035	−.517
	(.015)	(.078)
Compensated elasticity	.086	.328
	(.014)	(.064)
Uncompensated elasticity	.232	−.468
	(.080)	(.098)
Intertemporal substitution elasticity	−.964	. . .
	(.009)	
Frisch specific substitution elasticity	.072	.535
	(.010)	(.124)

Note.—The elasticities, which are based on the parameter estimates in table 1, are evaluated at the mean values of the functions. The SEs, reported in parentheses, are based on 1,000 bootstrap replications of the MRS and Euler equations.

where $n \equiv q'H^{-1}q$. The values in equation (12) are evaluated at the estimated parameters from the MRS equation (8), $\hat{\alpha}_j$. For ease of interpretation we convert the marginal effects in equation (12) into point elasticities.

The intratemporal elasticities derived from equation (12) tell only part of the story because lifetime considerations are a critical component in evaluating tax reforms. The estimates of the monotonic transformation from the Euler equation for consumption in table 1 provide the information necessary to construct the ISE, which uses $\hat{U}_C/\{C(\hat{U}_{CC} + (\hat{G}''/\hat{G}')\hat{U}_C^2)\}$. Combining the compensated elasticities from equation (12) with the ISE, along with the associated consumption and hours of work nonlabor income elasticities, it is possible to construct the Frisch specific substitution elasticities of equation (6). The elasticities are complicated nonlinear functions of parameters. Procedures such as the delta method, although straightforward with numerical gradient methods, may not yield very efficient standard errors. We adopt instead the bootstrap procedure described in Section III to calculate standard errors for both the first- and second-stage model elasticities.

In table 2 we report the within-period and intertemporal elasticities implied by our point estimates from table 1, evaluated at the sample means of hours, net wages, and nondurable consumption. The nonlabor income elasticities for consumption and for labor supply are 0.035 and −0.517, indicating that both consumption and leisure are normal goods. Note that the property income elasticity of consumption is not the same as the total income elasticity reported in consumption studies such as Browning and

Meghir (1991).[12] The corresponding utility-constant compensated wage elasticities of consumption and labor supply are 0.086 and 0.328.

Our estimated compensated wage elasticity of labor supply exceeds that typically reported in the literature and implies a sizable deadweight loss of taxation. For example, in a model based on linear preferences and additive separability between consumption and hours, Ziliak and Kniesner (1999) find a compensated wage elasticity about one-half that reported here. Below we explore whether the difference is driven more by functional form differences than by the possibility of nonseparability between consumption and labor supply. Because of the sizable nonlabor income effect relative to the compensated wage effect, we find that the uncompensated wage elasticity of labor supply is negative. Male labor supply bends backward. Although the income elasticity of labor supply is large, it is in the range of previous estimates reported in the literature, as is the finding of backward-bending male labor supply (Pencavel 1986; Blundell and MaCurdy 1999). Important for estimates of the economic efficiency of the tax system is that we do find an upward-sloping compensated labor supply function.

The estimate of the ISE at the means is about -1.0 for nondurable expenditures, which is consistent with strictly concave intertemporal preferences. The estimated ISE here is similar to the ISE estimated by Blundell et al. (1994) in their application to U.K. data. Given the ISE and compensated wage elasticities, the Frisch-specific substitution elasticity of labor supply is 0.54. The parallel Frisch net wage elasticity of consumption is 0.072. Our table 1 estimates imply that consumption and leisure are substitutes within periods. Intertemporally, the elasticities in table 2 confirm that with an anticipated increase in the real after-tax wage hours of market work increase, leisure falls, and consumption rises. Collectively the elasticity estimates in table 2 imply that welfare gains from increased labor supply and consumption are possible from revenue-neutral tax reforms that raise the after-tax wages.

B. Robustness

We consider a number of specification checks on our base-case results. First, we reduce the time endowment for work and leisure from 24 hours per day to 16 hours per day. The assumption is that 8 hours per day are overhead or human capital maintenance in the form of nonwork, nonleisure sleep time. We reestimated the model in equations (8) and (10) and report the relevant elasticities in the first two columns of table 3. The estimated elasticities evaluated at the mean values of the functions are

[12] The formula of the point elasticity is revealing here. The elasticity is $(\partial C/\partial Y)(\bar{Y}/\bar{C})$, and because the mean of nonlabor income is small in relation to the mean of nondurable consumption, the elasticity is small.

Table 3
Robustness of Elasticities to Alternative Model Assumptions

Changes in Real After-Tax Wages (ω_t)	Time Endowment = 16 Hours per Day: Consumption	Time Endowment = 16 Hours per Day: Labor Supply	Food Expenditures as Proxy for Nondurable Consumption: Consumption	Food Expenditures as Proxy for Nondurable Consumption: Labor Supply
Income elasticity	.046	−.481	.492	−.251
	(.042)	(.087)	(.061)	(.017)
Compensated elasticity	.081	.309	.213	.094
	(.017)	(.067)	(.021)	(.011)
Uncompensated elasticity	.274	−.424	2.582	−.442
	(.189)	(.154)	(.254)	(.033)
Intertemporal substitution elasticity	−.899	...	−1.038	...
	(.010)		(.013)	
Frisch specific substitution elasticity	.065	.478	.107	.148
	(.021)	(.127)	(.031)	(.016)
	Skinner's (1987) Measure as Proxy for Nondurable Consumption: Consumption	**Skinner's (1987) Measure as Proxy for Nondurable Consumption: Labor Supply**	**Direct Quadratic Utility Function for First-Stage MRS Equation ($\alpha_3 = 0$): Consumption**	**Direct Quadratic Utility Function for First-Stage MRS Equation ($\alpha_3 = 0$): Labor Supply**
Income elasticity	.102	−.191	.036	−.781
	(.008)	(.014)	(.044)	(.200)
Compensated elasticity	.134	.220	.128	.652
	(.006)	(.017)	(.080)	(.443)
Uncompensated elasticity	.671	−.313	.270	−.157
	(.023)	(.025)	(.260)	(.492)
Intertemporal substitution elasticity	−.859	...	−.725	...
	(.191)		(.004)	
Frisch specific substitution elasticity	.120	.246	.112	1.004
	(.006)	(.019)	(.058)	(.655)

Note.—The elasticities are evaluated at the mean values of the functions. The SEs, reported in parentheses, are based on 1,000 bootstrap replications of the MRS and Euler equations.

both qualitatively and quantitatively smaller, differing from the base case by no more than 5%–7%.[13]

Second, we replace imputed nondurable expenditures with food expenditures as the measure of consumption. Food is the prevalent measure of expenditures used in consumption-based analyses in the PSID, though more by default than choice, as food may be a poor proxy for nondurable consumption (Altonji 1986; Skinner 1987; Attanasio and Weber 1995;

[13] With the time endowment set to 16 hours per day, some observations had negative leisure hours. For the observations with negative implied leisure hours, we top-coded annual hours of work at 5,740, which leaves 100 hours of annual leisure time. We also set the time endowment at 19 hours per day, which did not require any top-coding of labor supply. The results were virtually the same as the base case.

Ziliak 1998). The property income effect for food consumption based on equation (12) is about 0.5; because the point elasticity involves multiplying the marginal effect by the ratio of property income to food consumption, the elasticity is also about 0.5 because average food spending is of comparable magnitude to average property income. Using food consumption leads to a significantly larger uncompensated wage elasticity of consumption. As in the case of nondurables, the Frisch specific substitution elasticity is positive, reflecting that food consumption and leisure are substitutes. Indeed, the coefficient on the food consumption–leisure interaction term is 15.14 with a standard error of 0.90, as compared to the base-case estimate of 4.26 (0.43). Although our results coincide with Altonji's (1986) estimates qualitatively, he is not able to reject the null of separability due to large standard errors. The implications for labor supply elasticities in the case of food consumption are to cut the estimated property income elasticity in half and to cut the compensated wage elasticity by about 70%.[14] Although the qualitative results remain unchanged when we switched from nondurable consumption to food consumption, the magnitudes clearly depend on the consumption measure.

We explore the sensitivity of the estimated elasticities to the consumption measure further in table 3 by replacing nondurable consumption with another variant of nondurable expenditures. Skinner (1987) predicts nondurable consumption in the PSID using data on food consumed at home and away from home, house value, expenditures on rent and utilities, and number of automobiles. The PSID stopped collecting data on utilities and automobiles in the early part of our sample. We therefore use a more parsimonious variant (Skinner 1987, table 1, col. 1): $\hat{C}_{it} = 1.930 * Food$ (home) + $2.928 * Food$ (away) + $0.1374 * House\ Value + 1.828 * Rent$, which imposes linear homogeneity by suppressing the constant term and frees the researcher from updating the coefficients for inflation. Although there is strong evidence of within-period nonseparability ($\hat{\alpha}_3 = 0.615\,(\mathrm{SE} = 0.107)$) using Skinner's measure, the estimated elasticities based on Skinner's consumption variant in table 3 are dampened somewhat relative to the benchmark measure. The dampening is not surprising because the scaled-down version of Skinner's measure that can be constructed in the PSID in the 1990s does not encompass as broad a metric of nondurable consumption as does the base case.[15] How-

[14] The magnitudes for labor supply elasticities that we find when food is the consumption measure are similar to others who use food consumption, such as Altonji (1986).

[15] On the surface it seems as though the Skinner measure we use is broader than the Blundell et al. measure because the latter just involves translating food expenditures but the former involves food plus housing. However, the Blundell measure involves deflating food spending by the elasticity of food consumption with respect to nondurable expenditures, $\hat{\pi}$ in eq. (11). Although the mean of

ever, the Skinner consumption-based estimates are much closer to the base case compared to food consumption.

The final robustness check we perform is to impose the common assumption of additivity between consumption and leisure to examine how important allowing for nonseparabilities in within-period preferences is for key parameters used in policy analysis. Specifically, we return to our base-case model of translog preferences and nondurable consumption but modify the functional form of utility by setting $\alpha_3 = 0$. We record the resulting elasticities in the second two columns of table 3. Focusing on the labor supply results, we estimate significantly larger nonlabor income, compensated wage, and Frisch wage elasticities of labor supply and a correspondingly smaller (in absolute value) uncompensated wage elasticity of labor supply.

The pattern of results in tables 1 and 3 reveals something akin to the classic omitted variable bias problem. We demonstrated in table 1 that consumption and hours of work are not separable and are direct complements. Given that consumption and property income are positively correlated, as are consumption and labor supply, omitting consumption imparts a downward (negative) bias on the nonlabor income elasticity of labor supply and an upward bias on the compensated wage elasticity of labor supply. Allowing for nonseparability between consumption and labor supply is important economically. Models that ignore consumption-hours interactions likely provide upper bounds on labor supply elasticities.

To explore the nonseparability issue further, we examined whether a similar pattern emerges in the standard linear labor supply model with and without consumption. Specifically, we regress annual hours of work on the log of the real net wage, virtual nonlabor income, and the same demographics as in equation (8), with and without consumption.[16] The linear labor supply model with consumption is similar to the conditional demand framework described in Browning and Meghir (1991) where consumption is not formally modeled as above but simply serves as a conditioning variable for labor supply outcomes. Although the magnitudes of the elasticities are significantly lower in the linear case, which highlights a further potential cost of choosing an inflexible specification of prefer-

nondurable expenditures using the Blundell et al. approach is over \$48,000 as reported in app. table A1, it is only about \$30,000 using the Skinner method.

[16] Virtual nonlabor income is the adjustment to nonlabor income (y_t) necessary to compensate the worker to act as if he or she faced the same marginal tax rate for all taxable income; virtual income is $y_t + \tau_t w_t h_t - T(\cdot)$. The instruments for the linear model are the same as the instruments in the MRS eq. (8) but with the addition of $(t-1)$ lagged wages, virtual income, and consumption. The additional instruments were necessary for the model to satisfy Slutsky integrability; without the additional instruments the linear model yielded negative compensated wage elasticities.

ences such as the linear labor supply model, the estimated compensated wage elasticity of labor supply without consumption is 0.024 and with consumption is 0.02. While the 20% difference in the linear estimates with and without consumption is smaller than the difference between the translog and quadratic log specifications reported in tables 2 and 3, the result is the same. Imposing additivity between consumption and leisure has important consequences for estimates of labor-market behavior.[17]

C. Implications for the Marginal Welfare Cost of Taxation

We close the results section by examining one avenue through which our results can be informative to discussions of tax reform. We now map our within-period estimates into the marginal welfare cost of government revenue (MWC), which is how much welfare changes in response to a change in tax revenue produced when a tax rate changes. The calculations are static and provide only a portion of the potential behavioral response to a tax change. The other obvious behavioral margins of interest are intertemporal changes, which may include both anticipated components and the unanticipated components occurring in the case of uncertain tax policy. A more detailed simulation is beyond the scope of the current project but should be a high priority for future research. In the two-stage budgeting formulation that we use the within-period preferences, and thus the corresponding MWC calculations, are life-cycle consistent.

The bulk of the econometric estimates of the welfare cost of taxation stemming from models of labor supply and taxes have emphasized tax reforms that are revenue neutral (Hausman 1981; Triest 1994; Ziliak and Kniesner 1999). Econometric research has largely presented so-called differential tax calculations where there is no balanced-budget spending or revenue effects so that the MWC reflects pure distortions of labor supply (Browning 1987; Ballard 1990). In contrast, much of the theoretical research on the marginal cost of public funds has focused on balanced-budget tax policy in which a marginal dollar of public spending is financed by raising an additional dollar of tax revenue (Snow and Warren 1996). We follow the econometric literature and focus on a transparent calculation of the marginal welfare cost of government revenue in the event of revenue-neutral reforms (Browning 1987, eq. [10]). Browning defines

[17] Another potential source of model sensitivity in eq. (8) is the omission of unobserved person-specific heterogeneity that affects the marginal rate of substitution between consumption and leisure. To investigate the potential for so-called fixed effects in the MRS model, we estimated a first-differenced variant of eq. (8). The results, not tabulated, indicate that the qualitative results in table 2 remain with additional latent heterogeneity included, although there are some differences. In the case of the nondurable consumption model, the compensated wage elasticity of labor supply increases by a factor of five so that the resulting uncompensated labor supply schedule is upward-sloping.

Table 4
Alternative Estimates of the Marginal Welfare Cost of Taxation (%)

	(1)	(2)	(3)
Progressive Tax:	20.9	5.9	41.7
$d\tau/d\bar{t} = 1.32$	(4.1)	(.73)	(28.3)
Proportional Tax:	15.9	4.5	31.6
$d\tau/d\bar{t} = 1$	(3.1)	(.56)	(21.4)

Note.—All estimates are based on eq. (10) in Browning (1987), where the marginal welfare cost of taxation is $MWC = [(\tau + 0.5d\tau)/(1 - \tau)]\eta_w^c d\tau/d\bar{t}$, with τ as the marginal tax rate, $d\tau$ the change in the marginal tax rate, η_w^c the compensated wage elasticity of labor supply, $\bar{t}$ the average tax rate, and $d\tau/d\bar{t}$ the change in the progressivity of the tax code in response to the tax reform. For each calculation we set $\tau = .323$, $d\tau = .01$, and $d\tau/d\bar{t}$ equal to 1.32 for progressive tax reforms (the ratio of the sample average marginal tax rate to the sample average tax rate) or equal to 1.0 for proportional tax reforms. In specification (1) we set $\eta_w^c = .328$ based on the direct translog MRS elasticities with nondurable consumption in table 2, in specification (2) we set $\eta_w^c = .092$ for the direct translog MRS elasticities with food consumption in table 3, and in specification (3) we set $\eta_w^c = .652$ for the quadratic direct MRS elasticities with nondurable consumption in table 3. The SEs, reported in parentheses, are based on 1,000 bootstrap replications of the MRS and Euler equations.

the marginal welfare cost as $MWC = [(\tau + 0.5d\tau)/(1 - \tau)]\eta_w^c d\tau/d\bar{t}$, with τ the marginal tax rate, $d\tau$ the change in the marginal tax rate, η_w^c the compensated wage elasticity of labor supply, $\bar{t}$ the average tax rate, and $d\tau/d\bar{t}$ the change in the progressivity of the tax code in response to the tax reform. The *MWC* formula highlights that only substitution effects and no income effects matter for revenue-neutral welfare calculations.

For each calculation we set $\tau = 0.323$, which is the sample average marginal tax rate, $d\tau = 0.01$, which is a one percentage point change in the marginal tax rate, and $d\tau/d\bar{t}$ equal to 1.32, for progressive tax reforms (the ratio of the sample average marginal tax rate to the sample average tax rate) or equal to 1.0 for proportional tax reforms. We consider three specifications for the marginal welfare cost of taxation. In specification (1) we set $\eta_w^c = 0.328$ based on the direct translog MRS elasticities with nondurable consumption in table 2; in specification (2) we set $\eta_w^c = 0.092$ for the direct translog MRS elasticities with food consumption in table 3; in specification (3) we set $\eta_w^c = 0.652$ for the quadratic direct MRS elasticities with nondurable consumption in table 3. There are six calculations in table 4, then, three for each of the progressive and proportional changes in the tax code.

In the base-case model with nonseparable preferences in the direct translog model in table 4 the marginal welfare cost of an additional dollar of taxation ranges from 16% to 21% depending on whether the reform is a proportional or a progressive change in the tax structure. The deadweight welfare losses are sizable and suggest possibilities for welfare-improving revenue neutral tax reforms in the United States. When we turn to specification (2), it becomes clear that how we measure consumption has a large impact on our estimates of welfare loss. With food as our measure, the MWC of taxation is a modest 4.5%–6%. Specification (3), however, pushes the estimated MWC in the opposite direction. Imposing additivity between consumption and leisure yielded a larger estimate of the compensated wage elasticity of labor supply in table 3, which translates into a doubling of the marginal welfare cost of taxation relative to the base-case model that relaxes separability. Models with additive preferences between consumption and labor supply likely yield upper-bound estimates of the deadweight loss of taxation.

V. Conclusion

We estimated a model of life-cycle consumption and labor supply where the empirical equilibrium conditions governing the optimal interior consumption and work choices identify intratemporal preferences and the empirical Euler equation for consumption identifies intertemporal preferences.

Our estimates based on direct translog preferences for within-period utility reject the separability of consumption choices from labor supply choices. This rejection held up under a variety of specification checks, including changing the proxy measure of consumption and changing the time endowment. The implied elasticities indicate that labor supply responds positively to (compensated) after-tax wage increases both within periods and across periods. The estimated complementarity of consumption and labor supply, coupled with the positive Frisch elasticity of consumption with respect to the net wage rate, is informative for the macroeconomic literature on consumption and saving because it suggests an avenue for why consumption tracks income over time.

We also further clarified the scope for improved labor-market efficiency with beneficial revenue-neutral tax reforms. Our base-case estimates with nondurable consumption suggest that the marginal welfare cost of taxation is 16%–21%, depending on whether the reform results in a proportional or progressive change in the tax structure. Our research has highlighted that the functional form of preferences—whether that be by imposing linearity in the labor supply response to a wage change or by imposing additivity between consumption and leisure—has a significant impact on estimated wage elasticities of labor supply. We find that imposing separability between consumption and leisure choices, within the context of a linear or nonlinear labor supply model, leads to an upward bias (as much as double in the translog model) in compensated wage elasticities used in evaluating labor market and tax policies. Further empirical research on models that identify the insurance aspects of progressive income taxation from the efficiency cost aspects would be the logical next step in pinning down more completely the welfare implications of tax policy.

Appendix

Table A1
Selected Summary Statistics

	Mean	SD
Nondurable expenditures	48.775	168.231
Annual hours of work	2.241	.575
After-tax wage	12.478	7.940
Total marginal tax rate	.323	.088
After-tax interest rate	.005	.015
Age	38.024	6.548
Family size	3.573	1.388
Number of children	1.464	1.216
Age of youngest child	4.930	5.136
Marital status (=1 if married)	.878	.327
Health (= 1 if work limited)	.071	.257
Race (= 1 if white)	.749	.433
Less than high school	.186	.389
High school graduate	.312	.463
More than high school	.503	.500
Self-employed	.133	.339
Home owner	.746	.435
Union member	.244	.430
Live in Northeast	.173	.378
Live in North Central	.242	.428
Live in South	.405	.491
Live in West	.179	.384

NOTE.—All income and price data are in real (1998) dollars using the personal consumption expenditure deflator. Number of person-years = 21,186.

References

Altonji, Joseph. 1986. Intertemporal substitution in labor supply: Evidence from micro data. *Journal of Political Economy* 94:S176–S215.

Altonji, Joseph, and John Ham. 1990. Intertemporal substitution, exogeneity, and surprises: Estimating life cycle models for Canada. *Canadian Journal of Economics* 23, no. 1:1–43.

Attanasio, Orazio, and Guglielmo Weber. 1995. Is consumption growth consistent with intertemporal optimization? Evidence from the consumer expenditure survey. *Journal of Political Economy* 103, no. 6: 1121–57.

Auerbach, Alan J. 1985. The theory of excess burden and optimal taxation. In *Handbook of public economics*, vol. 1, ed. Alan J. Auerbach and Martin Feldstein. Amsterdam: North-Holland.

Auerbach, Alan, and Daniel Feenberg. 2000. The significance of federal taxes as automatic stabilizers. *Journal of Economic Perspectives* 14: 37–56.

Auerbach, Alan, and Joel Slemrod. 1997. The economic effects of the tax reform act of 1986. *Journal of Economic Literature* 35, no. 2:589–632.

Ballard, Charles. 1990. Marginal efficiency cost calculations: Differential analysis vs. balanced-budget analysis. *Journal of Public Economics* 41: 263–76.

Blomquist, N. Soren. 1985. Labor supply in a two-period model: The effect of a nonlinear progressive income tax. *Review of Economic Studies* 52:515–24.

Blundell, Richard, Martin Browning, and Costas Meghir. 1994. Consumer demand and the life-cycle allocation of household expenditures. *Review of Economic Studies* 61:57–80.

Blundell, Richard, Alan Duncan, and Costas Meghir. 1998. Estimating labor supply responses using tax reforms. *Econometrica* 66, no. 4: 827–62.

Blundell, Richard, and Thomas MaCurdy. 1999. Labor supply: A review of alternative approaches. In *Handbook of labor economics,* vol. 3A, ed. Orley Ashenfelter and David Card. Amsterdam: North-Holland.

Blundell, Richard, Luigi Pistaferri, and Ian Preston. 2001. Partial insurance and consumption dynamics. Unpublished manuscript, Department of Economics, University College London.

Borjas, George. 1981. The relationship between wages and weekly hours of work: The role of division bias. *Journal of Human Resources* 15, no. 3:409–23.

Browning, Edgar. 1987. On the marginal welfare cost of taxation. *American Economic Review* 77, no. 1:11–23.

Browning, Martin. 1985. Which demand elasticities do we know and which do we need to know for policy analysis? Working Paper no. 85-13, Department of Economics, McMaster University.

Browning, Martin, Thomas Crossley, and Guglielmo Weber. 2003. Asking consumption questions in general purpose surveys. *Economic Journal* 113:F540–F567.

Browning, Martin, and Costas Meghir. 1991. The effects of male and female labor supply on commodity demands. *Econometrica* 59, no. 4: 925–51.

Carroll, Christopher. 1997. Buffer-stock saving and the life-cycle/permanent income hypothesis. *Quarterly Journal of Economics* 111, no. 1: 1–57.

Carroll, Christopher, and Lawrence Summers. 1991. Consumption growth parallels income growth: Some new evidence. In *National saving and economic performance*, ed. B. Douglas Bernheim and John Shoven, 305–43. Chicago: University of Chicago Press.

Carroll, Robert, Douglas Holtz-Eakin, Mark Rider, and Harvey Rosen.

1999. Income taxes and entrepreneurs' use of labor. *Journal of Labor Economics* 18, no. 2:324–51.

Christensen, Lauritis, Dale Jorgensen, and Lawrence Lau. 1975. Transcendental logarithmic utility functions. *American Economic Review* 65, no. 3:367–83.

Cochrane, John. 1991. A simple test of consumption insurance. *Journal of Political Economy* 99, no. 5:957–76.

Deaton, Angus. 1991. Saving and liquidity constraints. *Econometrica* 59, no. 5:1121–42.

Freedman, David. 1984. On bootstrapping two stage least squares estimates in stationary linear models. *Annals of Statistics* 12, no. 3:827–42.

Hall, Peter, and Joel Horowitz. 1996. Bootstrap critical values for tests based on generalized method-of-moments estimators. *Econometrica* 64, no. 4:891–916.

Hall, Robert, and Fredric Mishkin. 1982. The sensitivity of consumption to transitory income: Estimates from panel data on households. *Econometrica* 50, no. 2:461–81.

Hansen, Lars P. 1982. Large sample properties of generalized method-of-moments estimators. *Econometrica* 50, no. 3:1029–54.

Hausman, Jerry. 1981. Labor supply. In *How taxes affect economic behavior*, ed. Henry Aaron and Joseph Pechman. Washington, DC: Brookings Institution.

Heckman, James J. 1974. Life cycle consumption and labor supply: An explanation of the relationship between income and consumption over the life cycle. *American Economic Review* 64, no. 2:188–94.

Hotz, V. Joseph, Finn E. Kydland, and Guilerme L. Sedlacek. 1988. Intertemporal preferences and labor supply. *Econometrica* 56:35–60.

Kniesner, Thomas J., and James P. Ziliak. 2002. Tax reform and automatic stabilization. *American Economic Review* 92, no. 3:590–612.

Low, Hamish. 1999. Self-insurance and unemployment benefit in a life-cycle model of labour supply and savings. Working Paper no. 98/24, Institute for Fiscal Studies, London.

Low, Hamish, and Daniel Maldoom. 2004. Optimal taxation, prudence, and risk sharing. *Journal of Public Economics* 88:443–64.

MaCurdy, Thomas. 1981. An empirical model of labor supply in a life-cycle setting. *Journal of Political Economy* 89:1059–85.

———. 1983. A simple scheme for estimating an inter-temporal model of labor supply and consumption in the presence of taxes and uncertainty. *International Economic Review* 24:265–89.

MaCurdy, Thomas E., David Green, and Harry Paarsch. 1990. Assessing empirical approaches for analyzing taxes and labor supply. *Journal of Human Resources* 25:415–90.

Newey, Whitney, and Daniel McFadden. 1994. Large sample inference and testing. In *Handbook of econometrics*, vol. 4, no. 2, ed. Robert Engle and Daniel McFadden, 2111–45. Amsterdam: North-Holland.

Ogaki, Masao, and Zhang Qiang. 2001. Decreasing relative risk aversion and test of risk sharing. *Econometrica* 69:515–26.

Pencavel, John. 1986. Labor supply of men. In *Handbook of labor economics*, vol. 1, no. 5, ed. Orley Ashenfelter and Richard Layard, 3–102. Amsterdam: North-Holland.

Phlips, Louis. 1974. *Applied consumption analysis*. Amsterdam: North-Holland.

Pistaferri, Luigi. 2003. Anticipated and unanticipated wage changes, wage risk, and intertemporal labor supply. *Journal of Labor Economics* 21, no. 3:729–54.

Pollak, Robert, and Terence Wales. 1992. *Demand system specification and estimation*. Oxford: Oxford University Press.

Shaw, Kathryn. 1989. Life-cycle labor supply with human capital accumulation. *International Economic Review* 30, no. 2:431–56.

Skinner, Jonathan. 1987. A superior measure of consumption from the panel study of income dynamics. *Economics Letters* 23, no. 2:213–16.

Snow, Arthur, and Ronald Warren. 1996. The marginal welfare cost of public funds: Theory and estimates. *Journal of Public Economics* 61: 289–305.

Strawczynski, Michel. 1998. Social insurance and the optimum piecewise linear income tax. *Journal of Public Economics* 69, no. 3:371–88.

Triest, Robert. 1994. The efficiency cost of increased progressivity. In *Tax progressivity and income inequality*, ed. J. Slemrod, 137–70. Cambridge: Cambridge University Press.

U.S. Department of Commerce, Bureau of the Census. 1980–99. *State Government Tax Collections*. Washington, DC: Bureau of the Census.

Varian, Hal. 1980. Redistributive taxation as social insurance. *Journal of Public Economics* 14, no. 1:49–68.

Ziliak, James P. 1997. Efficient estimation with panel data when instruments are predetermined: An empirical comparison of moment-condition estimators. *Journal of Business and Economic Statistics* 15, no. 4: 419–31.

———. 1998. Does the choice of consumption matter? An application to the permanent-income hypothesis. *Journal of Monetary Economics* 41, no. 1:201–16.

Ziliak, James P., and Thomas J. Kniesner. 1999. Estimating life cycle labor supply tax effects. *Journal of Political Economy* 107, no. 2:326–59.