the review of income and wealth

Review of Income and Wealth
Series 58, Number 2, June 2012
DOI: 10.1111/j.1475-4991.2012.00493.x

ASSESSING THE WELFARE IMPACT OF TAX REFORM: A CASE STUDY OF THE 2001 U.S. TAX CUT

BY JULIE L. HOTCHKISS*

Federal Reserve Bank of Atlanta, and Andrew Young School of Policy Studies, Georgia State University

ROBERT E. MOORE AND FERNANDO RIOS-AVILA

Andrew Young School of Policy Studies, Georgia State University

This paper implements a relatively simple methodological approach to estimate the impact on family welfare of a specific tax reform. The measured impact can differ greatly from simple marginal tax rate comparisons, and conclusions about the distribution of the welfare impact can vary depending on the basis of comparison. For example, absolute welfare gains from the 2001 U.S. tax reform were concentrated among the highest and lowest income families, whereas welfare gains measured as a share of pre-tax income are found to be nearly monotonically declining in income.

JEL Codes: H23, H31, J22

Keywords: household welfare, joint labor supply, tax reform

1. Introduction

The purpose of this paper is to use the 2001 U.S. tax reform as a case study in order to illustrate a methodology of evaluating family-level welfare effects of a change in the tax code. The absolute welfare impact is compared across the income distribution and across households with different numbers of children; children were an important source of additional tax credits in the 2001 U.S. tax reform. In addition, the welfare impact is evaluated relative to pre-reform family income, and each group's share of the aggregate welfare impact is evaluated relative to the population share of the group (e.g., by income and number of children), producing a distribution of welfare gain relative to population share. The methodology presented here for evaluating the welfare impact of a tax policy change is new to the tax literature and brings the evaluation of the impact of a specific tax policy change down to its most micro level, allowing for the trade-off between leisure and consumption and the joint labor supply decisions of family members.

Note: The views expressed in this paper do not reflect those of the Federal Reserve Bank of Atlanta nor the Federal Reserve System. The paper has benefited from comments from James Alm, Michael Leeds, Roberta Moore, John Robertson, Sally Wallace, Jennifer Ward-Batts, two anonymous referees, and participants of the University of California-Merced Conference on Earnings Inequality. Excellent research assistance was provided by Daniel Melaugh, M. Laurel Graefe, Navnita Sarma, Chinying Xie, Danyang Li, and Melissa Trussell.

*Correspondence to: Julie L. Hotchkiss, Research Department, Federal Reserve Bank of Atlanta, 1000 Peachtree St, NE, Atlanta, GA 30309-4470, USA (Julie.L.Hotchkiss@atl.frb.org).

Published by Blackwell Publishing, 9600 Garsington Road, Oxford OX4 2DQ, UK and 350 Main St, Malden, MA, 02148, USA.

In the spirit of Fiorio (2008), Blundell *et al.* (2000), and Blundell (1992), among others, the analysis falls into the category of a microsimulation exercise; the tax code is not specified structurally, but, rather, a family's pre-reform utility is compared with their post-reform utility, holding everything except the tax structure (and the resulting implication for labor supply decisions and consumption) constant. The notion of "welfare" in this paper is different than that typically employed in evaluating the impact of tax reform. For example, the impact of tax reform is often measured in terms of efficiency, or the amount of excess burden a particular tax policy imposes on individuals (see, e.g., Eissa *et al.*, 2008; Fehr and Jess, 2007). In this paper, a family's welfare is measured directly as the dollar equivalent utility the family experiences under alternative tax structures. Because of this direct link between tax changes and individual family utility, this measure of welfare (as family utility) better represents the relationship between individual behavior and the direct impact of policy. The methodology also allows us to evaluate the distribution of welfare gains while accounting for differences in behavioral response across the distribution.

The evaluation of family utility, especially relative to the family's position in the income distribution and number of children, produces some intriguing results. For example, as found by Fehr and Jess (2007) in their investigation of who benefits from pension taxation reform in Germany, it might be expected that ". . . the higher the income level the higher the relative welfare gains" (p. 98). In contrast, our results indicate that families at the top *and at the bottom* of the income distribution enjoyed greater dollar equivalent welfare gains (and a share of the total welfare gain disproportionate to their population shares) from the 2001 U.S. tax reform than middle income families. In addition, relative to pre-tax income, the welfare gains from the tax reform are found to be monotonically *decreasing* in income. Further, even though families with the largest number of children typically have less income than families with no children, families with three or more children experienced greater dollar equivalent welfare gain than families with no children. This discrepancy was even more dramatic when welfare gain was measured as a percentage of pre-tax income.

The focus of the analysis is on demonstrating the merits of disaggregating the measurement of tax reform effects across different demographic groups (also identified as crucial by Blundell, 1992, and others) and of the usefulness of applying a very micro-level analysis to the question of assessing welfare impacts. With that as the focus, the analysis will make a number of simplifying assumptions, about the decision making process within the household, about the form of the utility function, and about the error structure of the estimating equation. The implications of each of these simplifying assumptions are discussed in turn.

2. The 2001 U.S. Tax Reform

While the U.S. tax reform that began in 2001 is used solely as an illustration, it is necessary to understand certain details in order to put the discussion of results into context. In June 2001, the Economic Growth and Tax Relief Reconciliation Act of 2001 (2001 tax reform) was the first major tax reform since the Tax Payer Relief Act of 1997. The primary features of the 2001 tax reform were the reductions in the highest income tax brackets (from 28, 31, 36, and 39.6 percent to 25, 28, 33, and 35 percent, respectively); a new lowest tax bracket of 10 percent (the former lowest bracket was 15 percent); and expansion of the child credit and the Earned Income Tax Credit. There were other changes that affected the Alternative Minimum tax, personal exemptions, and itemized deductions.[1] All in all, the change in the tax code was extremely complicated and spawned a host of analyses that investigated the practical significance of the changes on family net income across the income distribution (see, e.g., Citizens for Tax Justice, 2002; Gale and Potter, 2002; Johnson *et al.*, 2006). The conclusions of those analyses seem to differ based on which side of the political spectrum the investigator resided. While differing in their conclusions, the one thing all of these assessments have in common is that they exclusively look at the impact on income and/or consumption, not accounting for any potential changes in labor supply behavior.

It is exactly because of the complicated and comprehensive nature of the changes to the tax code that we employ simulation methods to investigate their impact. The simulation is performed as follows. We construct a dataset of families in the year 2000, including wages earned, non-earned income, and other demographics, such as the number of children, education, etc. Using these data for families in the year 2000, we calculate the marginal tax rates on earnings (wages) and non-earned income, as well as the total tax liability under the tax structure in place in 2000 and also under the tax structure in place in 2005 (using a publicly available tax calculator developed by the National Bureau of Economic Analysis called TAXSIM). The difference in marginal tax rates and total liability under these two tax structures is used to compare the utility a family in the year 2000 enjoys under the tax structure in 2000 and the utility that same family would enjoy with the 2005 tax structure (applied to the 2000 family data). This simulation gets as close to a natural experiment as possible by keeping everything (e.g., demographics, earnings, income) constant, except the tax structure (also see Bourguignon 2011 for yet an alternative social welfare criterion for evaluating the impact of a tax reform).

3. Simulating the Impact on Family Welfare

The literature that has produced estimates of labor supply elasticities is fraught with conflict. While there are many different ways to obtain estimates of the labor supply responses to wage changes, the goal of this paper is to come up with reasonable estimates (within the range reported previously in the literature) which can then be used to simulate the change in family utility resulting from a change in the tax structure. Keane (2010) methodically lays out various issues that are addressed by the multitude of strategies that have been employed to obtain labor supply elasticities. We will address each of these issues while discussing the methodology adopted here.

Family labor supply decisions are modeled here in a neoclassical joint utility framework. This model can be thought of as a reduced-form specification of family decision making. The model yields a clean-cut expression of family welfare that allows for cross wage effects on each member's labor supply decision. The assumption of joint

[1]The 2001 tax reform was followed closely by the Jobs and Growth Tax Relief Reconciliation Act of 2003, which had the effect of accelerating some of the 2001 provisions and reducing capital gains taxes even more. The Working Families Tax Relief Act of 2004 accelerated the full phase-in of the child tax credit.

family utility (or, "collective" utility) is often empirically rejected in favor of a bargaining structure to household decision making (see, e.g., McElroy, 1990; Apps and Rees, 2009). However, there is evidence that the choice of structure for household decision making has very little implication for conclusions regarding labor supply responses or impact on welfare of tax policy changes (see Bargain and Moreau, 2003). In addition, Blundell *et al*. (2007) find that both collective and bargaining models are consistent with their household labor supply model estimated in the U.K.

Within the framework of the neoclassical family labor supply model, a family maximizes a utility function that represents the household welfare. Assuming, for simplicity, that there are only two working members of the household (husband and wife), the family chooses levels of leisure for each member and a joint consumption level in order to solve the following problem:

$$(1) \qquad \max_{(L_1, L_2, C)} U = U(L_1, L_2, C) \quad \textit{subject to } C = w_1 h_1 + w_2 h_2 + Y.$$

Define T as total time available for an individual; $L_1 = T - h_1$ will be referred to as the husband's leisure, and $L_2 = T - h_2$ will be referred to as the wife's leisure; h_1 is the labor supply of the husband; h_2 is the labor supply of the wife; C is total money income (or consumption with price equal to one); w_1 is the husband's after-tax market wage; w_2 is the wife's after-tax market wage; and Y is after-tax unearned income. Although we refer to L_1 and L_2 as the "leisure" of the husband and wife, respectively, they actually correspond to all uses of non-market time, including home production activities.[2]

The solution to the maximization problem in equation (1) can be expressed in terms of the indirect utility function, which is solely a function of the wages of the husband and wife and unearned income of the family:

$$(2) \qquad V(w_1, w_2, Y) = U\{[T - h_1^*(w_1, w_2, Y)], [T - h_2^*(w_1, w_2, Y)], [w_1 h_1^*(w_1, w_2, Y) + w_2 h_2^*(w_1, w_2, Y) + Y]\},$$

where $h_1^*(w_1, w_2, Y)$ and $h_2^*(w_1, w_2, Y)$ correspond to the optimal labor supply equations (desired hours) for the husband and wife, respectively. In order to capture the total effect of the change in after-tax income resulting from the 2001 tax reform, the indirect utility function is totally differentiated (see also Apps and Rees, 2009, p. 263):

$$(3) \qquad dV = -U_1 dh_1 - U_2 dh_2 + U_3 dC,$$

where U_1 is the family's marginal utility of the husband's leisure, U_2 is the family's marginal utility of the wife's leisure, and U_3 is the family's marginal utility of consumption. Equation (3) makes it clear that the change in welfare depends not only on the individual labor supply responses, but also on the family's marginal evaluation of a change in leisure and income.

[2]Apps and Rees (2009) are highly critical of family utility models that do not include measures of household production, but even they acknowledge that not much can be done without the availability of richer data (p. 108). Since the focus of the analysis in this paper is utility at the household level, the absence of home production activities is not crucial.

Theoretically, labor is supplied to the extent that the marginal utility of leisure is equal to the market wage, suggesting that, within a family, if husbands are paid more than wives, $U_1 > U_2$. In addition, lower estimates of the marginal utilities of leisure would be consistent with greater values of labor supply, *ceteris paribus*. As the additional utility gained from an additional dollar of income increases at a decreasing rate, we would expect U_3 to be smaller for higher income families. It is because of these theoretically expected differences across family characteristics (namely, wages and preferences) that the utility function parameters will be estimated separately by family type.

Expressed in terms of changes in wages and unearned income, and rearranging terms to illuminate the contribution of those changes to family welfare through their impact on husband's labor supply, wife's labor supply, and total family income, the total derivative in equation (3) becomes:

$$\text{(4)} \quad \begin{aligned} dV = &-U_1\left(\frac{\partial h_1}{\partial w_1}dw_1 + \frac{\partial h_1}{\partial w_2}dw_2 + \frac{\partial h_1}{\partial Y}dY\right) \\ &- U_2\left(\frac{\partial h_2}{\partial w_1}dw_1 + \frac{\partial h_2}{\partial w_2}dw_2 + \frac{\partial h_2}{\partial Y}dY\right) \\ &+ U_3\left[\left(w_1\frac{\partial h_1}{\partial w_1} + h_1 + w_2\frac{\partial h_2}{\partial w_1}\right)dw_1 + \left(w_1\frac{\partial h_1}{\partial w_2} + h_2 + w_2\frac{\partial h_2}{\partial w_2}\right)dw_2 \right. \\ &+ \left.\left(w_1\frac{\partial h_1}{\partial Y} + 1 + w_2\frac{\partial h_2}{\partial Y}\right)dY\right]. \end{aligned}$$

The 2001 U.S. tax reform had the effect of changing workers' after-tax wages (dw_1 and dw_2) and families' after-tax unearned incomes (dY). The impact of these changes on family welfare can be calculated using simulated wage and income changes and estimated utility function parameters. In order to determine the impact of the tax reform on family welfare, utility function parameters are estimated for the sample of families in 2000, given their net, after-tax wages and unearned income in 2000. We then calculate what each member's 2000 after-tax wage and unearned income would be under the 2005 tax code (using the TAXSIM software described below). The differences in these wages and income are then used to calculate the difference in welfare that the family experiences under the 2005 tax regime compared with the tax regime prior to the 2001 tax reform. The advantage of this simulation exercise (as opposed to merely calculating observed changes in hours pre- and post-tax regime change) is that everything except the tax regime is held constant.[3]

Crucial to the simulation exercise, of course, is estimating the own and joint labor supply responses to changes in net wages $\left(\frac{\partial h_1}{\partial w_1}, \frac{\partial h_1}{\partial w_2}, \frac{\partial h_2}{\partial w_2}, \frac{\partial h_2}{\partial w_1}\right)$ and net non-labor income $\left(\frac{\partial h_1}{\partial Y}, \frac{\partial h_2}{\partial Y}\right)$. This is discussed in the next section along with multiple empirical issues identified by Keane (2010) that have plagued the labor supply literature (also see Apps and Rees, 2009).

[3]Wages and income in 2000 are not inflated to 2005 values in order to keep the impact of tax regime changes from being confounded by the effects of inflation.

4. Estimating the Joint Labor Supply Responses

The direction (sign) of the change in utility at the optimal leisure choices that results from changes in the husband's and wife's after-tax wages, and family after-tax unearned income cannot always be determined analytically; it depends on the direction of the wage changes and the size of labor supply responses of the husband and wife to own and to spouse wage changes, as well as on the relative size of the additional utility the family attains from the leisure enjoyed by the husband and wife and from additional unearned income.

As mentioned previously, there are many divergent empirical issues raised in the literature in attempts to obtain estimates of labor supply responses to wage changes, i.e., estimates of labor supply elasticities. While the focus of this paper is on the simulation exercise itself, the simulation does require labor supply elasticities and it is therefore worthwhile to address some of the empirical issues. Again, estimation methodology is not the focus of this paper; the goal here is to produce labor supply elasticities that are consistent with the literature. Toward that end, the methodology adopted takes as simple an approach as possible while maintaining basic theoretical and empirical integrity.

The requirement of simplicity here primarily derives from the goal of quantifying the family-level utility gain from the tax reform. In order to obtain estimates of the pieces of the total derivative in equation (4), a specific functional form of utility must be specified. Following others (e.g., Ransom, 1987; Hotchkiss *et al.*, 1997; Heim, 2009), we estimate a quadratic form of the utility function:

$$U(Z) = \alpha(Z) - (1/2)Z'\mathrm{B}Z, \tag{5}$$

where Z is a vector with elements $Z_1 = T - h_1$, $Z_2 = T - h_2$, and $Z_3 = w_1h_1 + w_2h_2 + Y$; α is a vector of parameters; and B is a symmetric matrix of parameters. This functional form has the advantage of belonging to the class of flexible functional forms in the sense that it can be thought of as a second order approximation to an arbitrary utility function (when B is positive definite). In addition, it is possible to produce analytical closed-form solutions for both the husband's and wife's labor supply functions. The problem becomes an unconstrained optimization problem resulting in first order conditions which are a system of equations linear in h:[4]

$$\frac{\partial U}{\partial h_1} = \Omega_1 h_1 + \Omega_2 h_2 + \Omega_3 \tag{6}$$

$$\frac{\partial U}{\partial h_2} = \Omega_2 h_1 + \Omega_4 h_2 + \Omega_5. \tag{7}$$

This system can be solved simultaneously, and the desired hours become $h_1^* = f(w_1, w_2, Y)$ and $h_2^* = g(w_1, w_2, Y)$, which represent the desired number of

[4]The components of and solution for desired hours are found in Appendix A.

hours the members of a household would like to work, given the parameters that define their household utility function.

Observed hours ($\tilde{h}$), however, might differ from the optimum hours due to stochastic errors, such that:

$$\tilde{h}_1 = \begin{cases} h_1^* + e_1 & \text{if } h_1^* + e_1 > 0 \\ 0 & \text{otherwise} \end{cases}$$

$$\tilde{h}_2 = \begin{cases} h_2^* + e_2 & \text{if } h_2^* + e_2 > 0 \\ 0 & \text{otherwise} \end{cases},$$

where we assume that (e_1, e_2) follows a bivariate Normal distribution with mean 0 and covariance matrix Σ. This model can be thought of as a simultaneous Tobit model, where we have four kinds of families: those where both husband and wife work, those where only one of the spouses works (two cases), and those where neither of them work. Allowing for hours adjustment along the extensive margin for the wife when assessing labor supply responses to wage changes or tax reform has been found to make a significant difference when assessing total labor supply response (see, e.g., Eissa *et al.*, 2008; Heim, 2009); however, extensive margin hours adjustments appear to be unimportant for men (see, e.g., Blundell *et al.*, 1988; Heim, 2009). We have estimated the model both ways, and since it appears to make a difference here, we opt to include husbands that are not observed working.

The presence of non-working wives and husbands raises one empirical issue identified by Keane (2010) that must be addressed: market wages are not observed for family members who do not work. To obtain estimates of those wages, we take the standard approach in the literature of estimating a selectivity-corrected wage equation (Heckman, 1974) on the sample of working men and women, using regressors observable for both working and non-working individuals. The resulting parameter estimates are then used to predict wages for non-working men and women based on their observable characteristics.

The maximum likelihood function corresponding to the joint labor supply optimization problem can be written as follows:

$$(8) \qquad L = \prod_{i=1}^{N}\left[\left(\frac{1}{\sigma_1\sigma_2}\right)\psi\left(\frac{\tilde{h}_1 - h_1^*}{\sigma_1}, \frac{\tilde{h}_2 - h_2^*}{\sigma_2}, \rho\right)\right]^{(H=1,W=1)}$$
$$\cdot\left[\frac{1}{\sigma_1}\varphi\left(\frac{\tilde{h}_1 - h_1^*}{\sigma_1}\right)\left\{1-\Phi\left(\frac{\sigma_1 h_2^* - \rho\sigma_2(\tilde{h}_1 - h_1^*)}{\sigma_2\sigma_1\sqrt{1-\rho^2}}\right)\right\}\right]^{(H=1,W=0)}$$
$$\cdot\left[\frac{1}{\sigma_2}\varphi\left(\frac{\tilde{h}_2 - h_2^*}{\sigma_2}\right)\left\{1-\Phi\left(\frac{\sigma_2 h_1^* - \rho\sigma_1(\tilde{h}_2 - h_2^*)}{\sigma_2\sigma_1\sqrt{1-\rho^2}}\right)\right\}\right]^{(H=0,W=1)}$$
$$\cdot\Psi\left(\frac{-h_1^*}{\sigma_1}, \frac{-h_2^*}{\sigma_2}, \rho\right)^{(H=0,W=0)}$$

where φ and Φ correspond to the probability density and cumulative distribution functions of a univariate normal, and ψ and Ψ represent the probability density and cumulative distribution functions of the bivariate normal. Also, $H = 1$ if the husband is working and $W = 1$ if the wife is working (0 otherwise), σ_i ($i = 1,2$) represents the standard deviations of (e_1, e_2), and ρ is the correlation between the stochastic errors.

The stochastic errors accounted for in equation (8) represent errors in optimization—observed hours do not exactly reflect desired hours. Keane (2010) points out that there may exist measurement error in observed wages and non-labor income. This classical measurement error may bias elasticity estimates toward zero. Heim (2009), using a methodology most similar to the one used here, presents results showing that accounting for measurement error produces elasticities practically identical to when it is not accounted for. In addition, since we restrict our sample to workers paid weekly or by the hour, we mitigate what Keane refers to as "denominator bias."

Keane (2010) also identifies two potential sources of endogeneity. First, it is reasonable to expect that observed wages and non-labor income are correlated with a person's taste for work (reflected through hours of work). Both fixed effects and instrumental variables have been used to resolve this issue, but are simply not possible in this case since we do not have panel data and because of the highly non-linear nature of the labor supply functions. In addition to the inclusion of variables expected to affect the taste for work (e.g., children), we expect that the inclusion of spousal variables (hours of work and education, in particular) will help to remove additional sources of correlation from the error term (i.e., because of positive assortative mating, people with similar taste for work will be married to each other; see Lam, 1988; Hernstein and Murray, 1994). Second, in light of the kinked budget constraint created by the progressivity of the tax system, the after-tax wage rate and after tax non-labor income depend on in which tax bracket the worker falls, which depends on the number of hours worked. The simplest approach to this issue is that first proposed by Hall (1973), which basically "linearizes" the budget constraint segment on which each person is observed by simply recalculating unearned income to find its "virtual" intercept as if it were extended beyond the specific segment. This strategy amounts to assuming that preferences are strictly convex;[5] see equation (5), which means that family members would make the same hours choice facing this linearized budget constraint that they would have made facing the non-linear budget constraint. If this assumption is binding, Keane points out that labor supply elasticities will be biased in a negative direction.

An additional concern Keane (2010) identifies in the literature is making sure the hours/wage combinations observed in the data are coming off workers' labor supply curve, rather than off employers' labor demand curve. Identification of the labor supply relationship boils down to including regressors (determinants of hours) that reflect the demand for a person's skills (which thus determine the

[5]This assumption of strictly convex preferences is supported by a positive definite B matrix. As will be seen in Table 2, and Table B2 in Appendix B, all the eigenvalues of the estimated B matrices are positive, indicating that the matrix itself is positive definite.

observed wage) that are not reflective of that person's taste for work. Toward that end, we include an indicator for race that could affect observed wage through employer discrimination, but, *ceteris paribus* (e.g., at the same education levels), should not affect taste for work.

Keane (2010) also presents concerns related to estimating lifetime labor supply elasticities that are not relevant for this cross-sectional analysis. Further, the issue of the presence of fixed costs of working is raised by Apps and Rees (2009). We only marginally control for fixed costs by including the presence of children in the determination of hours. However, Heim (2009), again, using the methodology in the literature most similar to the one employed here, presents results showing that once demographics are controlled for, additional consideration of fixed costs only very slightly impacts estimates of the parameters of the utility function (Heim, table 3).

5. Data

The outgoing rotation groups from the Current Population Survey (CPS) in March, April, May, and June 2000 were used to construct the sample for which the family labor supply model is estimated. Detailed unearned income was obtained by matching each family to the March survey, which is the month in which this information is collected.[6] Multiple months of outgoing rotation groups were used in order to expand the sample size. Only respondents who reported hourly or weekly wages were included in the sample to ensure a more accurate reporting/construction of the hourly wage; consequently, the results are generalizable only to non-salaried workers. Table 1 contains the means and standard deviations of the variables for the sample.

The families are split into pre-tax total family income quartiles. The differences in means across quartiles are not surprising. The representation of blacks is lowest among high income families; the number of children is higher among families in the upper section of the income distribution compared to families in the lower section; hours of work, wages, and unearned income all increase with family income.

Information on the detailed sources of family income, number of children, and earnings available from the CPS is used to calculate the marginal tax rate on earnings (wages) and the total tax liability (in any year of interest) using the National Bureau of Economic Research (NBER) TAXSIM tax calculator (http://www.nber.org/~taxsim/; see also Feenberg and Coutts, 1993). The calculator is more complete than we have information for from the CPS, so we made assumptions for the missing values as recommended by the managers of the tax calculator. For example, there is no information in the CPS that would allow one to calculate itemized deductions (mortgage payments, charitable contributions, etc.), so values of zero are entered for the missing information. Although this means we do not have as accurate a calculation of the family's actual tax liability as we would like, it is important to remember that the assumptions for each family do not change across years of comparison.

[6]Since the CPS only allows for identification of unearned income in the previous year, it must be considered a proxy for current family unearned income.

TABLE 1
Sample Means by Family Education Type, Families in 2000

	Total	Quartile 1	Quartile 2	Quartile 3	Quartile 4
Age1	49.51	52.52	51.19	46.85	47.49
Age2	47.01	49.73	48.54	44.54	45.24
Black1	6.6%	6.8%	7.5%	7.0%	5.1%
Black2	6.2%	6.4%	7.1%	6.5%	4.9%
White1	89.1%	88.5%	88.6%	89.0%	90.3%
White2	89.0%	88.3%	88.5%	88.9%	90.2%
Other race 1	4.3%	4.7%	3.9%	4.0%	4.5%
Other race 2	4.8%	5.2%	4.3%	4.6%	4.9%
Lesshs1	13.9%	26.2%	19.8%	7.4%	2.0%
Lesshs2	11.5%	23.3%	15.6%	6.0%	1.3%
Hs1	31.3%	33.7%	37.3%	35.8%	18.4%
Hs2	35.5%	39.3%	43.2%	38.1%	21.5%
Scoll1	25.0%	20.8%	26.0%	29.9%	23.5%
Scoll2	27.3%	22.9%	27.2%	30.5%	28.6%
Coll1	29.8%	19.3%	16.9%	26.9%	56.1%
Coll2	25.6%	14.5%	14.1%	25.4%	48.6%
Nkids	0.87	0.82	0.82	0.92	0.90
Nkids (if Nkids > 0)	1.94	2.06	2.00	1.89	1.83
preschl	0.21	0.21	0.22	0.22	0.20
preschl (if preschl > 0)	0.45	0.50	0.52	0.43	0.39
Husband's weekly hours	44.01	45.41	41.81	43.22	45.62
Wife's weekly hours (if working)	36.57	33.35	35.31	36.77	38.54
Husband's pre-tax wage	$11.1	$1.3	$6.2	$12.7	$24.3
Wife's pre-tax wage	$12.5	$5.0	$8.9	$11.7	$18.8
Family pretax weekly non labor income	$237.6	$129.0	$221.1	$207.3	$393.1
Total annual pre-tax income	$51,723.0	$11,746.0	$32,808.0	$54,891.0	$107,455.0
Total after tax income	$38,240.0	$12,610.0	$29,859.0	$43,269.0	$67,225.0
Change in marginal tax rate on wages (%age change)[a]	−2.74	−2.98	−1.35	−2.74	−3.88
dw1[b]	0.79	0.54	0.81	0.80	1.01
dw2[b]	0.27	0.18	0.07	0.29	0.55
dy[b]	29.46	11.87	31.86	25.99	48.12
Number of observations	24,338	6,085	6,084	6,085	6,084

Notes: The "1" subscript corresponds to the husband's values of the variable and the "2" subscript corresponds to the wife.

[a]Includes state and federal taxes; average change between 2000 and 2005.

[b]These correspond to how after-tax wages, virtual non-labor income, and total annual family income and would be different in 2005 compared with 2000, changing only the tax regime under which the wages and income are received.

The bottom rows of Table 1 present comparisons of the families' marginal tax rates, after-tax total annual family incomes, after-tax wages, and after-tax weekly unearned incomes between 2000 and 2005. These are calculated by evaluating the 2000 gross values under the tax regimes in 2000 and in 2005. The average marginal tax rate decreased for families in all income quartiles, with the greatest decline going to families with the highest and lowest incomes. Husbands enjoyed a bigger gain in their after-tax wage than wives (which actually has implications for gender wage inequality which will not be explored here), and families with the highest family income levels also experienced the largest absolute (and percentage)

TABLE 2

Maximum Likelihood Parameter Estimates and Utility Function Parameters, Families in 2000

Variable	Full Sample	Quartile 1	Quartile 2	Quartile 3	Quartile 4
a1					
Black1	−5.1096***	−20.000***	−8.5330***	−5.38226***	−3.02767**
Lesshs1	−3.5706***	−6.3972***	−7.7625***	−1.72995	0.92264
Scoll1	−1.7901***	−3.2043	−1.7210	−0.86707	−1.85677**
Coll1	0.18539	−0.7050	2.3033	3.55938***	0.0913
Age1	0.6655***	0.63752	0.66366**	0.40265*	1.45753***
Age1sq	−0.0104***	−0.00553	−0.01101***	−0.00919***	−0.02037***
Nkids	1.2219***	6.55175***	4.00427***	1.72716***	0.62534*
Preschl	0.7623	10.56585**	−8.02855***	−2.66071*	−2.51909**
_cons	−7.9293***	−29.03526*	−5.58964	16.97291***	11.23905**
a2					
Black2	1.49787***	2.78627	−1.41707	−0.11539	1.02296*
Lesshs2	−2.48695***	−5.36245***	−3.74845***	−0.2949	2.11451*
Scoll2	−0.4418	−1.94675*	0.46891	0.08964	−0.36035
Coll2	−2.1155***	−8.01064***	0.42659	−1.81277***	−0.96039***
Age2	0.0372	−0.33534*	0.18299	−0.10295	0.18036*
Age2sq	−0.0002	0.00595**	−0.0022	−0.0004	−0.00289**
Nkids	−0.8024***	−1.51243***	1.34976***	0.27122	−1.03335***
Preschl	−2.5390***	−2.07767	−2.64185**	−4.29960***	−3.79319***
_cons	−19.1674***	−44.9948***	−39.4483***	−9.28867**	−2.92417
a3	4.53668***	9.03263***	17.48377***	12.15483***	2.27820***
b11	1	1	1	1	1
b12	−0.28622***	−0.69848***	−0.16475***	−0.22790***	−0.18368***
b13	−0.00304***	−0.06661***	−0.02306***	−0.01459***	−0.00916***
b22	0.53073***	0.98211***	0.60004***	0.49297***	0.36267***
b23	0.00011	0.07399***	0.02037***	−0.00051	−0.00369***
b33	0.00182***	0.01348***	0.02246***	0.01072***	0.00053***
ρ	0.10033***	0.1199***	−0.1015***	−0.2059***	0.0235
σ_1	21.56220***	40.17071***	19.74481***	13.44439***	13.38413***
σ_2	22.04147***	34.86156***	21.16596***	15.71302***	15.95515***
Statistics					
log-likelihood	−149,138.0	−25,791.9	−32,916.5	−39,893.5	−43,274.6
No. of observations	24,338	6,085	6,084	6,085	6,084
Eigenvalues B matrix					
$\lambda 1$	1.1355	1.6955	1.0599	1.0875	1.0492
$\lambda 2$	0.3953	0.2926	0.5412	0.4057	0.3136
$\lambda 3$	0.0018	0.0075	0.0214	0.0105	0.0004

Notes: The "1" subscript corresponds to the husband's hours equation and the "2" subscript corresponds to the wife's hours equation.

*, **, *** denote significance at the 90, 95, and 99% confidence levels, respectively.

increase in hourly pay and unearned income. This greatest (absolute and percentage) benefit of the 2001 tax reform going to the highest income groups has been the source of much controversy and criticism of the policy change.[7]

6. Labor Supply and Welfare Function Estimation Results

Table 2 contains the maximum likelihood parameter estimates of the utility function parameters corresponding to equation (5). The regressors that affect

[7]For example, see Hashemzadeh and Saubert (2004).

TABLE 3
PREDICTIVE ACCURACY OF THE MODEL

		Wife		
		Doesn't Work	Works	Total
Husband	Doesn't Work	21.4%	9.4%	30.8%
		(18.6%)	(1.1%)	(19.7%)
	Works	20.8%	48.4%	69.2%
		(11.7%)	(68.6%)	(80.3%)
	Total	42.2%	57.8%	100.0%
		(30.4%)	(69.6%)	
		Average hours of work		
	Overall	Actual	Predicted	
	Husband	30.5	30.0	
	Wife	21.1	20.9	
	If working			
	Husband	44.0	38.8	
	Wife	36.6	33.0	

Notes: The values in parentheses correspond to the model predictions.

labor supply typically do so in expected ways. Generally, the presence of children, especially young children, significantly decreases the labor supply of wives and increases (or decreases by a smaller amount) the labor supply of husbands. Whereas black women typically work the same or more hours than white women, black men work less than white men for all levels of income. In addition, college educated men typically work more hours than those with less education, but college educated women appear to work fewer hours. Both husbands and wives exhibit the typical inverted-U hours/age profile.

Table 3 presents some statistics summarizing the goodness of fit of the model for the overall sample. The model does a fairly good job predicting the proportion of households where neither husband nor wife work, although it tends to under predict the probability of not working for both household members. Additionally, it does a good job predicting overall average hours of work, although it predicted slightly lower than actual hours among workers only.

The estimates in Table 2 are used to calculate the marginal utilities of leisure and income, as well as the labor supply elasticities along both the intensive (hours) margin and the extensive (labor force participation) margin.[8] These are reported in Table 4 and graphically interpolated across income centiles in Figures 1 and 2.[9]

[8]From maximum likelihood parameter estimates, the probability of working (LFP), Pr[work = 1], for each non-working family member is constructed. The elasticity of LFP for person i is then constructed as: $e_{LFP,i} = \frac{\partial Pr_i[work=1]}{\partial x_i} \cdot \frac{x_t}{Pr_t[\widehat{work}=1]}$, where x_t is person i's wage, his/her spouse's wage, or non-labor income. Uncompensated own and cross wage elasticities are reported in the table; compensated elasticities are derived by subtracting the income elasticity (see Killingsworth, 1983, pp. 106–7).

[9]In order to better illustrate the differences in elasticities across the income distribution, we divide families into centiles, q1–q100. We then create rolling samples of 25 centiles each, starting with families in q1–q25, then q2–q26, etc., ending with q76–q100. This results in 76 samples for which estimates of elasticities are produced. This methodology is used to produce Figures 1–4.

TABLE 4

Estimated Utility Function Parameters and Labor Supply Elasticities and Simulated Change in Family Welfare Resulting from the 2001 Tax Reform, by Education of Husband, Families in 2000

Entire Sample	Entire Sample	Quartile 1	Quartile 2	Quartile 3	Quartile 4
u1	29.832	43.560	42.155	27.930	16.612
(marginal utility of husband's hrs)	[28.823; 30.673]	[37.558; 48.637]	[40.053; 44.784]	[25.895; 30.293]	[14.979; 18.119]
u2	22.232	20.459	30.141	20.839	11.336
(marginal utility of wife's hrs)	[21.461; 23.064]	[16.665; 23.371]	[28.19; 32.942]	[19.125; 22.74]	[10.132; 12.512]
u3	5.905	13.385	31.234	21.055	2.499
(marginal utility of income)	[5.628; 6.146]	[10.833; 15.64]	[28.311; 34.398]	[18.711; 24.169]	[2.183; 2.79]
If both are working					
Husband own wage elasticity	0.291	0.212	0.096	−0.078	0.065
	[0.279; 0.302]	[0.158; 0.258]	[0.069; 0.125]	[−0.105; −0.052]	[0.05; 0.081]
Husband cross wage elasticity	−0.064	0.216	−0.426	−0.379	−0.146
	[−0.073; −0.057]	[0.147; 0.271]	[−0.457; −0.396]	[−0.403; −0.357]	[−0.16; −0.131]
Husband income elasticity	−0.038	−0.111	−0.083	−0.056	−0.039
	[−0.04; −0.036]	[−0.126; −0.098]	[−0.095; −0.067]	[−0.063; −0.05]	[−0.044; −0.035]
If only husband is working					
Husband own wage elasticity	0.349	0.300	−0.135	−0.319	0.013
	[0.335; 0.364]	[0.237; 0.356]	[−0.165; −0.104]	[−0.355; −0.28]	[−0.01; 0.044]
Husband income elasticity	−0.071	−0.144	−0.192	−0.197	−0.113
	[−0.083; −0.061]	[−0.18; −0.118]	[−0.214; −0.172]	[−0.229; −0.164]	[−0.152; −0.087]
If both are working					
Wife own wage elasticity	0.735	2.006	0.674	0.210	0.386
	[0.713; 0.758]	[1.781; 2.194]	[0.6; 0.746]	[0.174; 0.243]	[0.364; 0.414]
Wife cross wage elasticity	−0.229	0.262	−0.817	−0.843	−0.407
	[−0.246; −0.214]	[0.105; 0.402]	[−0.878; −0.771]	[−0.895; −0.799]	[−0.447; −0.377]
Wife income elasticity	−0.071	−0.117	−0.134	−0.105	−0.077
	[−0.076; −0.066]	[−0.143; −0.093]	[−0.16; −0.108]	[−0.124; −0.087]	[−0.086; −0.071]

Table 4 continued on next page

TABLE 4 (*continued*)

Entire Sample	Entire Sample	Quartile 1	Quartile 2	Quartile 3	Quartile 4
If only wife is working					
Wife own wage elasticity	1.003	1.810	0.421	0.008	0.455
	[0.946; 1.048]	[1.611; 2.007]	[0.36; 0.489]	[–0.054; 0.075]	[0.377; 0.548]
Wife income elasticity	–0.215	–0.121	–0.562	–0.680	–0.517
	[–0.244; –0.187]	[–0.163; –0.081]	[–0.649; –0.479]	[–0.806; –0.585]	[–0.66; –0.385]
Husband LFP own wage elasticity	0.324	0.690	0.605	0.366	0.099
	[0.303; 0.347]	[0.603; 0.798]	[0.524; 0.723]	[0.27; 0.555]	[0.078; 0.134]
Husband LFP cross wage elasticity	0.075	0.024	–0.462	–0.164	–0.038
	[0.062; 0.084]	[–0.06; 0.096]	[–0.538; –0.407]	[–0.236; –0.121]	[–0.064; –0.018]
Husband LFP income elasticity	–0.214	–0.621	–1.797	–1.770	–0.996
	[–0.273; –0.165]	[–0.753; –0.525]	[–2.23; –1.508]	[–2.885; –1.16]	[–1.376; –0.71]
Wife LFP own wage elasticity	0.843	1.396	1.239	1.010	0.413
	[0.798; 0.889]	[1.263; 1.562]	[1.08; 1.478]	[0.802; 1.443]	[0.357; 0.498]
Wife LFP cross wage elasticity	0.013	0.478	–0.775	–0.941	–0.470
	[–0.019; 0.035]	[0.377; 0.568]	[–0.897; –0.69]	[–1.205; –0.776]	[–0.569; –0.396]
Wife LFP income elasticity	–0.251	0.103	–0.877	–2.080	–1.217
	[–0.325; –0.192]	[–0.022; 0.229]	[–1.478; –0.479]	[–3.652; –1.16]	[–1.721; –0.895]
Husband change on hours worked	4.829	7.380	2.513	1.410	1.377
	[4.702; 4.981]	[7.067; 7.666]	[2.258; 2.734]	[1.258; 1.572]	[1.254; 1.492]
Wife change on hours worked	3.546	4.838	1.818	1.774	2.505
	[3.48; 3.612]	[4.608; 5.084]	[1.624; 1.982]	[1.604; 1.905]	[2.362; 2.635]
dV(change in family welfare)	426.135	734.0	1504.7	828.9	207.7
	[407.62; 442.19]	[642.95; 833.52]	[1356.03; 1662.62]	[742.87; 945.21]	[184.72; 227.43]
Dollar equivalent	$3,828.13	$2,682.00	$2,553.78	$2,055.10	$4,195.38
	[3749.0; 3920.8]	[2508.9; 2923.8]	[2456.5; 2652.7]	[1976.6; 2136.2]	[4070.6; 4332.5]
Observations	24,338	6,085	6,084	6,085	6,084

Note: Estimates are averages across all families in each group. The total change on welfare for households where husband or/and wife are not working is estimating using their probability of working and the predicted hours worked. Unearned income elasticities are evaluated at actual unearned income. Bootstrapped 95% confidence intervals, produced by 300 repetitions, are in brackets. Although the correlation parameter is constrained to be between (–1, 1), the observed results are far from binding. Consequently, the warnings of Andrews (2000) are not relevant in this case. No other constraints are imposed on the parameter estimates.

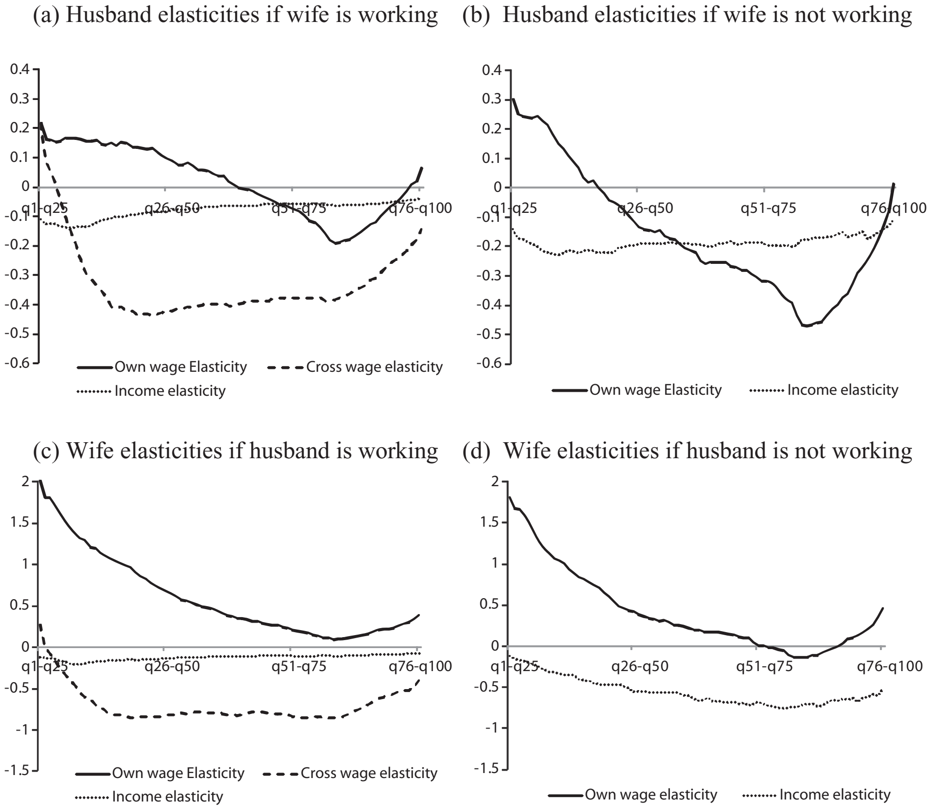

Figure 1. Inter-Quartile Estimates of Husband and Wife Wage Elasticities

Notes: Generated using parameter estimates found in Table 2; see Table 3 and footnote 8.

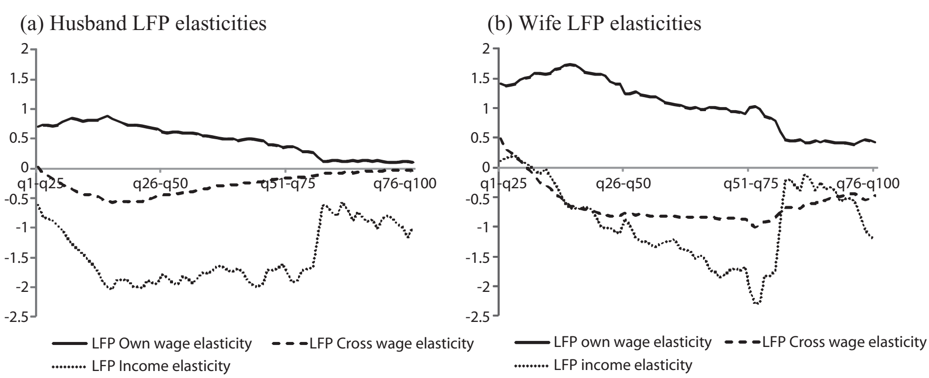

Figure 2. Inter-Quartile Estimates of Labor Force Participation Elasticities

Notes: Generated using parameter estimates found in Table 2; see Table 3 and footnote 8.

Generally, there is significant variation in both husband and wife labor supply elasticities across family income quartiles. The implication is that estimating an average response masks important variation in labor supply responses, thus masking variation in assessment of welfare impact.

The largest own wage elasticities across both margins, for both husbands and wives, can be found among the low income households (around the first quartile) while the highest cross wage elasticities can be observed for middle income households. Recall that ignoring budget constraint segment choice is expected to bias labor supply elasticities in a negative direction (Keane, 2010). While this could explain the estimation of negative own wage hours elasticities for husbands across the middle quartiles, these are not inconsistent with estimates reported by Kaiser *et al.* (1992) for Germany; and Ransom (1987), MaCurdy *et al.* (1990), and Pencavel (2002) using U.S. data. Note that the labor supply of high income men is less responsive to wage changes than labor supply of low income men. This is consistent with estimates presented for married men by education in Meghir and Phillips (2008)—more educated men are less responsive than less educated men; and for affluent men by Moffitt and Wilhelm (2000)—they find labor supply response is lower among affluent men relative to less affluent men.

Among all families, wives' own wages elasticities, across both margins, are much higher than husbands' elasticities, indicating that wives' labor supply is more responsive to changes in their own wages. These estimates for wives' own labor supply elasticities are mostly within the range reported in the literature using U.S. data.[10]

Across income quartiles, wives' hours and labor force participation decisions are more responsive to changes in their husbands' wages than husbands' hours are to changes in their wives' wages. In addition, except for households in the lowest income distribution, husbands' hours are more responsive to changes in their wives' wages than to changes in their own wage. Similarly, wives' hours are generally more responsive to changes in their husbands' wages than to changes in their own wages. The opposite is true, however, for labor force participation.

The estimated negative cross-wage elasticities along both margins (for both husbands and wives), except among those in the lowest quartile, indicate that husbands and wives view their leisure time as substitutes; this is consistent with cross-elasticities estimated in Heim (2009) and Ransom (1987). Both husbands and wives present the expected low and negative income elasticity, although wives are slightly more responsive to changes in non-labor income than their husbands. In general, labor supply elasticities estimated along the extensive margin are larger than those estimated by Heim (2009), but are similar to those estimated by Pencavel (1998).[11]

On average, we estimate that husbands increased their labor supply between 1.38 and 7.38 hours per week in response to the 2001 U.S. tax reform, and wives also increased their labor supply—between 1.77 and 4.84 hours per week. Family

[10]For example, the range of estimates found in Cogan (1981), Hausman (1981), Triest (1990), Ransom (1987), Hotchkiss *et al.* (1997), and Blau and Kahn (2005) is 0.12 to 0.97. Also see Killingsworth (1983, p. 107).

[11]Heim (2009) reports own-wage extensive margin labor supply elasticities for married women between 0.07 and 0.18. Pencavel (1998) reports extensive margin elasticities in the range of 0.77 to 1.83.

members from the lowest quartile of families increased their labor supply the most, while the smallest increase in hours came from husbands in the highest quartile and wives in the middle quartiles.

The differences in estimated labor supply response across income quartiles, for both men and women, and for both intensive and extensive margin estimates, indicate that behavior does differ across family characteristics. The implication is that any analysis of tax (or other) policy on the welfare of families should take into account how families will respond differently to incentives and constraints. The lower marginal utilities of leisure and income among the highest income families will further exacerbate the difference in behavior in the evaluation of differences in the welfare effect. The lower additional utility that the highest quartile families receive from additional units of leisure is consistent with a preference structure that would lead to family members working longer hours. In addition, given a higher level of income to start with, an additional dollar of income does not yield as much utility to high income families.

The purpose of the estimation exercise is to obtain wage and non-labor income labor supply elasticities with which to perform the simulation and calculate the impact on family welfare of a specific tax reform. The simulated impacts on family welfare of the 2001 U.S. tax reform are reported at the bottom of Table 4. In spite of the apparent much larger welfare gains (dV) among the second and third quartile families, these families' marginal utilities of income are also higher, making the dollar-equivalent value of the utility gain less than what is experienced by families in the first and fourth quartiles.[12] The next section is devoted to exploring alternative ways to compare the welfare gains across income quartiles and discusses policy implications of such comparisons.

7. Assessing Family Welfare Gains and Policy Implications

7.1. *Welfare Gains by Family Income Quartile*

Figure 3 (panel (a)) plots the pre-tax annual income (the dashed line) along with the dollar equivalent welfare gain accruing to families across income quartiles; welfare gains are calculated separately for each family based on their characteristics. All families experienced welfare gains from the 2001 U.S. tax reform, however households in the lower and higher income brackets seem to have benefited the most, in terms of dollar equivalent gains. This could be reflecting the larger declines in marginal tax rates among the lowest and highest income families; recall that the 2001 U.S. tax reform reduced the marginal tax rate for those in the lowest tax bracket by 5 percentage points and reduced the marginal tax bracket for those in the highest tax bracket by 4.6 percentage points, whereas the marginal tax rates for those in the middle three tax brackets were reduced by only 3 percentage points. In addition, welfare gains among the lowest income families (who are likely

[12]The average dollar equivalent welfare gains of the highest quartile and the first and second quartiles are significantly different (based on the reported 95% confidence intervals in Table 4) from the average dollar equivalent welfare gain experienced by families in the third quartile.

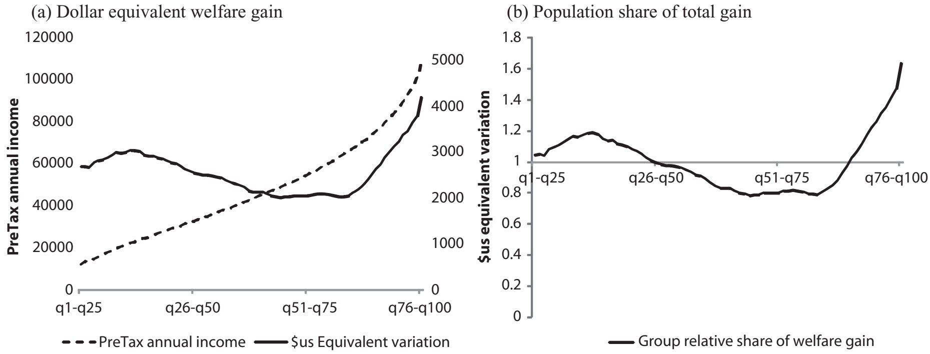

Figure 3. Dollar Equivalent Welfare Gain and Pre-Tax Annual Income, Inter-Quartile Comparison

Notes: Generated using parameter estimates found in Table 2; see note in Table 3 and footnote 8 for details.

to have more children) were likely boosted by the generous tax credits included for children. The direct impact of children on family welfare gain is evaluated in the next section.

The 2001 U.S. tax reform has drawn a fair amount of criticism for the perceived greater gains going to high income families (see, e.g., Kamin and Shapiro, 2004). Tax reform proponents counter that this outcome is perfectly appropriate given that the higher income families pay a higher proportion of the total tax bill (Bartlett, 2005). In reality, as shown in Figure 3 (panel (a)), although the highest income families did accrue the highest dollar equivalent welfare gain, the lowest income families also experienced a welfare gain that exceeded that of middle-income families.

Panel (b) of Figure 3 is a scaled version of the dollar equivalent welfare gains depicted in panel (a). It plots the share of the total welfare gain accruing to families as a group in each income centile. This is constructed by dividing the proportion of the total welfare gain accruing to each centile by its population share.[13] A value less than one means that the families in the centile are receiving less than their population share of the additional welfare gain; a value greater than one means the families in the centile are getting more than their population share. The figure indicates that the highest income families and those in the bottom quartile received more than their population shares of the total welfare gain from the 2001 tax reform, whereas families at the median and third quartile received less than their population shares.

However, an alternative perspective erases the perspective of dramatic gains among the highest income families. Figure 4 plots the dollar equivalent welfare gain across centiles as a percentage of pre-reform, pre-tax annual family income.

[13]The group's share of the welfare gain relative to its population share is calculated as: $\frac{N_i dV_i}{\sum N_i dV_i} \div \frac{N_i}{\sum N_i} = dV_i \frac{\sum N_i}{\sum N_i dV_i}$, where N_i is the number of families in education group i, and dV_t is the welfare gain of the representative family in education group i. This transformation shows that the group's share relative to its population share (plotted in Figure 3, panel (b)) is merely a scaled version of the welfare gain plotted in Figure 3, panel (a) (dV_t).

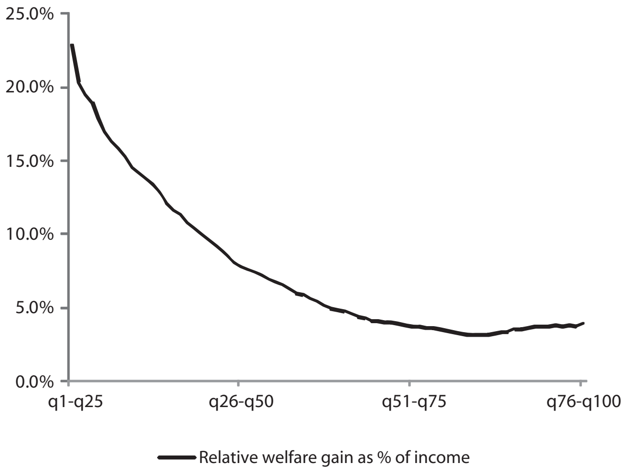

Figure 4. Dollar Equivalent Welfare Gain as a Share of Pre-Tax Annual Income

Notes: Generated using parameter estimates found in Table 2; see note in Table 3 and footnote 8 for details.

In relative terms, the welfare gains decline sharply as families move up the income distribution, with the welfare gains among the highest income families representing only 5 percent of their pre-reform income. The welfare gains among the lowest income families were roughly 20 percent of their pre-reform income. This illustrates the point made by Bourguignon (2011) who argues that an assessment of the total social welfare impact of a tax reform should take into account individuals' pre-tax reform status.

Considering these results, the 2001 U.S. tax reform can be considered to have had a disproportionate impact across income groups, with the largest benefits accruing to the lowest and highest income groups. In relative (to family pre-tax income) terms, however, families in the lowest end of the distribution benefited the most from the tax reform. The implication of these results is that a repeal of the 2001 U.S. tax reform, as was being considered by the U.S. Congress late in 2010, would have at least two predictable effects: (1) all families' welfare would decline; (2) high income families' welfare would decline the most; and (3) low income families could experience a loss in welfare similar to that of high income families. The potential repeal of the U.S. tax reform was averted in late 2010, but for only two years, when it is again set to expire.

7.2. *Welfare Gain and Number of Children*

One of the most popular features of the 2001 U.S. tax reform was the expansion of the child tax credit. The per-child credit was scheduled to increase gradually from $500 to $1000 by the year 2010. The Jobs and Growth Tax Relief Reconciliation Act of 2003 advanced families the $1000 credit per child for 2003 and 2004, and then was supposed to revert back to the original phase-in. Then, the Working Families Tax Relief Act of 2004 extended the $1000 credit per child for each year

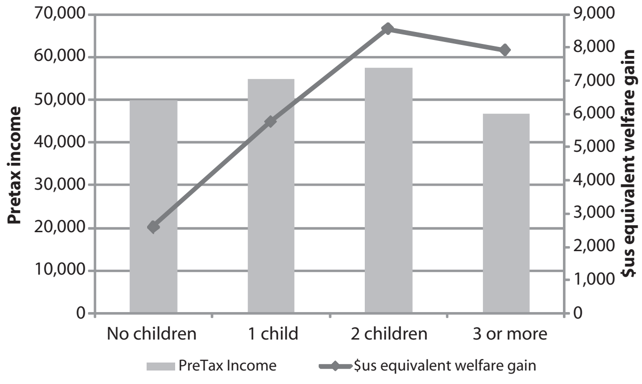

Figure 5. Dollar Equivalent Welfare Gain and Pre-Tax Annual Income, Families Grouped by Number of Children

Notes: Generated using parameter estimates found in Appendix B; see note in table for details.

until 2010. This is an additional fixed dollar amount accruing to families with children. However, the credit is limited for families with income over certain thresholds. The analysis of welfare gains by income quartiles suggests that the child tax credit boosted the welfare gains of lower-income families, who have more children on average (among families with children) and who benefit fully from the credit (incomes below the cut-off level).[14] This section evaluates welfare gains of the 2001 U.S. tax reform across families with different numbers of children.

The same methodology to estimate welfare gains across income quartiles was used to obtain parameter estimates for family groups differentiated by the number of children: no children, one child, two children, and three or more children. All of the estimation results and sample means are found in Appendix B. As one might expect, families with no children are older, relatively more educated, and have higher income levels. Husbands in childless families work less and wives in childless families work more, on average, than families with children. In the families with three or more children, wives work the least number of hours on average, and husbands work about the same.

Figure 5 presents a welfare gain assessment analogous to the assessment by income quartile presented in Figure 3, panel (a). In Figure 5, families are grouped by the number of children, and the dollar equivalent welfare gain is presented along with the average pre-tax family income by family type. Figure 5 illustrates how generous the child tax credit is, yielding an average welfare gain three times greater for families with three or more children, relative to families with no children. And because families with more children are more likely to come from the lower end of the income distribution, welfare gain as a share of pre-tax family

[14]The Child and Dependent Care Tax Credit was also increased over this time period, but since the CPS does not report child care expenditures, the specific impact of that provision cannot be incorporated into the present analysis. Since this tax credit is underutilized by low-income families (Forry and Anderson, 2006), its omission is expected to mostly influence the welfare estimates among higher income families.

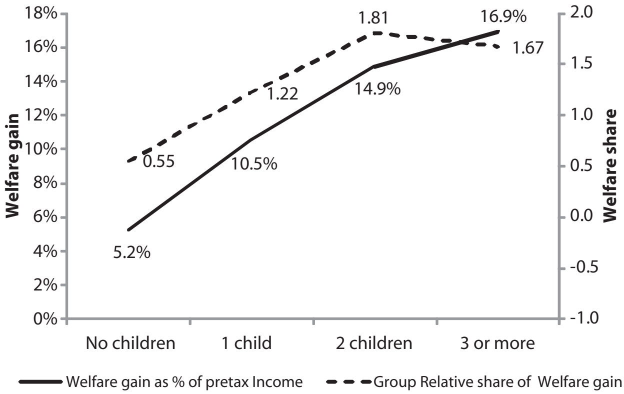

Figure 6. Dollar Equivalent Welfare Gain as a Share of Pre-Tax Annual Income and Dollar Equivalent Welfare Gain Share Relative to Population Share, Families by Number of Children

Notes: Generated using parameter estimates found in Appendix B; see note in table for details.

income (Figure 6) monotonically increases with the number of children. Families with three or more children experienced a welfare gain representing roughly 17 percent of their pre-tax family income, compared with only a 5 percent relative gain by families with no children. Figure 6 also shows that families with at least one child all experienced welfare gains exceeding their population share, whereas families with no children received less than their population share of the total welfare gain.

The implication of this analysis by family size is that a complete repeal of the 2001 U.S. tax reform would lower family welfare the most among families with more children. However, it is not clear how the child tax credits would be affected by repeal, or whether they would be phased out among higher income families only. If this is the case, then welfare gains among the largest families would most likely have not been affected by the reversion that was proposed in 2010.

7.3. *Caveats*

The purpose of this paper is to illustrate how to take the assessment of the welfare impact of a change in tax policy down to the micro level of the family, that the welfare impact can be much more heterogeneous than suggested by mere assessment at the mean, and that the assessment of welfare impact can look quite different than the impact on net income alone. Because all estimations are subject to error and the product of many assumptions (all of which were discussed earlier), the actual absolute values of welfare (or dollar equivalent welfare) are to be digested with caution. However, the use of simulation methods, which hold everything about the family constant except the tax code parameters and their induced behavioral changes, should nonetheless produce accurate conclusions and relative comparisons for the questions posed.

Be that as it may, there are other potential behavioral implications of tax reform that are not considered here. For example, Kniesner and Ziliak (2002) estimate the welfare cost of the U.S. Economic Recovery Act of 1981 and the U.S. Tax Reform Act of 1986 as measured by their impact on consumption stability. They estimate that these two tax reforms enhanced consumption stability among low-income households, but generated substantial welfare costs for most middle- and upper-income families as a result of a reduction in their consumption stability by 50 percent. The implication for the welfare change estimates presented in this paper is that the estimates for low-income families are likely lower and those for higher income families higher than would result if the model accounted for risk-aversion and changes in consumption stability.

8. Concluding Remarks

There are multiple ways to assess the welfare impact of a change in tax policy. The impact of tax reform is often evaluated in terms of efficiency, or excess burden (see, e.g., Eissa *et al.*, 2008; Fehr and Jess, 2007). The analysis in this paper defines welfare as family utility and offers an assessment of how family utility is affected by a specific tax reform—the 2001 U.S. tax reform. By estimating the parameters of the family's utility function, the impact of the tax code changes on labor supply can be estimated and the overall impact on family welfare can be calculated in terms of dollar equivalent utility. The analysis is a microsimulation in which the tax code is not specified structurally. Rather, family welfare is evaluated under two different tax code structures, holding everything about that family constant, except the tax code (and resulting labor supply change in response to changes in the tax code). This is accomplished through the use of a publicly available tax calculator for the U.S., developed by the National Bureau of Economic Analysis called TAXSIM.

Separate utility function parameters are estimated for families across income quartiles and also across families grouped by the number of children in the family. The results reinforce the importance already identified in the tax literature of accounting for taxpayer heterogeneity when evaluating the impact of tax policy (e.g., Blundell *et al.*, 2007). For all families combined, the overall dollar equivalent welfare gain from the 2001 U.S. tax reform was roughly $3,400. Within quartile, the gain ranged from roughly $2,000 in the third quartile to roughly $2,700 and $4,200 for the lowest and top quartiles, respectively. As a share of pre-tax income, the lowest quartile families exhibited the greatest gains in family welfare; merely having the highest income did not necessarily mean that a family would reap the greatest gain from the 2001 tax reform.

There was also a great deal of variation across families with different numbers of children. Generous child tax credits were an important part of the 2001 U.S. tax reform, and this showed up with families with three or more children experiencing dollar equivalent welfare gains that were three times larger than the gains to families with no children.

The bottom line is that tax reform directly affects the well-being and welfare of individuals and families. Furthermore, the impact varies across individual and family characteristics. This paper illustrates a straightforward microsimulation strategy for quantifying those effects at the level where they are acutely felt.

References

Andrews, Donald W. K., "Inconsistency of the Bootstrap when a Parameter is on the Boundary of the Parameter Space," *Econometrica*, 68, 99–405, 2000.

Apps, Patricia and Ray Rees, *Public Economics and the Household*, Cambridge University Press, 2009.

Bargain, Olivier and Nicolas Moreau, "Is the Collective Model of Labor Supply Useful for Tax Policy Analysis? A Simulation Exercise," CESifo Working Paper No. 1052, October 2003.

Bartlett, Bruce, "Reality over Appearance: The Wealthy Pay *More* in Taxes when Tax Rates Fall. Period," *National Review Online*, December 7, 2005 (http://www.nationalreview.com/nrof_bartlett/bartlett200512070900.asp; accessed October 2, 2007).

Blau, Francine D. and Lawrence M. Kahn, "Changes in the Labor Supply of Married Women: 1980–2000," NBER Working Paper #11230, March 2005.

Blundell, Richard, "Labour Supply and Taxation: A Survey," *Fiscal Studies*, 13(3), 15–40, 1992.

Blundell, Richard, Costas Meghir, Elizabeth Symons, and Ian Walker, "Labour Supply Specification and the Evaluation of Tax Reforms," *Journal of Public Economics*, 36, 23–52, 1988.

Blundell, Richard, Alan Duncan, Julian McCrae, and Costas Meghir, "The Labour Market Impact of the Working Families' Tax Credit," *Fiscal Studies*, 21(1), 75–104, 2000.

Blundell, Richard, Pierre-Andre Chiappori, and Thierry Magnac, "Collective Labour Supply: Heterogeneity and Non-participation," *Review of Economic Studies*, 74, 417–45, 2007.

Bourguignon, Francois, "Status Quo in the Welfare Analysis of Tax Reforms," *The Review of Income and Wealth*, 57(4), 603–21, 2011.

Citizens for Tax Justice, "Year-By-Year Analysis of the Bush Tax Cuts Shows Growing Tilt to the Very Rich," June 12, 2002 (http://www.ctj.org/html/gwb0602.htm; accessed November 16, 2006).

Cogan, John F., "Fixed Costs and Labor Supply," *Econometrica*, 49, 945–63, 1981.

Eissa, Nada, Henrik Jacobsen Kleven, and Claus Thustrup Kreiner, "Evaluation of Four Tax Reforms in the United States: Labor Supply and Welfare Effects for Single Mothers," *Journal of Public Economics*, 92, 795–816, 2008.

Feenberg, Daniel and Elisabeth Coutts, "An Introduction to the TAXSIM Model," *Journal of Policy Analysis and Management*, 12(1), 189–94, 1993.

Fehr, Hans and Heinrich Jess, "Who Benefits from the Reform of Pension Taxation in Germany?" *Fiscal Studies*, 28(1), 73–101, 2007.

Fiorio, Carlo V., "Analyzing Tax-Benefit Reforms Using Non-Parametric Methods," *Fiscal Studies*, 29(4), 499–522, 2008.

Forry, Nicole D. and Elaine A. Anderson, "The Child and Dependent Care Tax Credit: A Policy Analysis," *Marriage and Family Review*, 39(1), 159–76, 2006.

Gale, William G. and Samara R. Potter, "The Bush Tax Cut: One Year Later," Brookings Institute Policy Brief #101, June 2002.

Hall, Robert E., "Wages, Income, and Hours of Work in the U.S. Labor Force," in Glen G. Cain and Harold W. Watts (eds), *Income Maintenance and Labor Supply*, University of Chicago Press, Chicago, IL, 102–62, 1973.

Hashemzadeh, Nozar and Wayne Saubert, "The Effects of Bush's Tax Cuts on Income Distribution and Economic Growth in the United States," *Problems and Perspectives in Management*, 3, 111–20, 2004.

Hausman, Jerry A., "The Effect of Taxes on Labor Supply," in H. Aaron and J. Pechman (eds), *How Taxes Affect Economic Behavior*, Brookings, Washington DC, 1981.

Heckman, J. J., "Shadow Prices, Market Wages, and Labor Supply," *Econometrica*, 42, 679–93, 1974.

Heim, Bradley T., "Structural Estimation of Family Labor Supply with Taxes," *Journal of Human Resources*, 44, 350–85, Spring 2009.

Hernstein, Richard J. and Charles Murray, *The Bell Curve: Intelligence and Class Structure in American Life*, Simon and Schuster, Free Press, New York, London, and Toronto, 1994.

Hotchkiss, Julie L., Mary Mathewes Kassis, and Robert E. Moore, "Running Hard and Falling Behind: A Welfare Analysis of Two-Earner Families," *Journal of Population Economics*, 10, 237–50, 1997.

Johnson, David A., Jonathan A. Parker, and Nicholas S. Souleles, "Household Expenditure and the Income Tax Rebates of 2001," *American Economic Review*, 96, 1589–610, 2006.

Kaiser, H., P. Spahn, and U. van Essen, "Income Taxation and the Supply of Labour in West Germany," *Jahrbücher für Nationalökonomie und Statistik*, 209, 87–105, 1992.

Kamin, David and Isaac Shapiro, *Studies Shed New Light on Effects of Administrations' Tax Cuts*, Center on Budget and Policy Priorities, Washington DC, September 13, 2004 (http://www.cbpp.org/8-25-04tax.htm; accessed October 2, 2007).

Keane, Michael P., "Labor Supply and Taxes: A Survey," Working Paper No. 160, University of Technology Sydney, July 2010.
Killingsworth, Mark R., *Labor Supply*, Cambridge University Press, Cambridge, 1983.
Kniesner, Thomas J. and James P. Ziliak, "Tax Reform and Automatic Stabilization," *American Economic Review*, 92, 590–612, June 2002.
Lam, David, "Marriage Markets and Assortative Mating with Household Public Goods: Theoretical Results and Empirical Implications," *Journal of Human Resources*, 23, 462–87, 1988.
MaCurdy, T., D. Green, and H. Paarsch, "Assessing Empirical Approaches for Analyzing Taxes and Labor Supply," *Journal of Human Resources*, 25, 415–90, 1990.
McElroy, Marjorie B., "The Empirical Content of Nash-Bargained Household Behavior," *Journal of Human Resources*, 25, 559–83, 1990.
Meghir, Costas and David Phillips, "Labour Supply and Taxes," IZA Discussion Paper No. 3405, March 2008.
Moffitt, Robert A. and Mark O. Wilhelm, "Taxation and the Labor Supply Decisions of the Affluent," in Joel B. Slemrod (ed.), *Does Atlas Shrug? The Economic Consequences of Taxing the Rich*, Russell Sage, New York, 193–234, 2000.
Muthen, Brengt, "Moments of the Censored and Truncated Bivariate Normal Distribution," *British Journal of Mathematical & Statistical Psychology*, 43(1), 131–43, 1990.
Pencavel, J., "The Market Work Behaviour and Wages of Women: 1975–94," *Journal of Human Resources*, 33, 771–804, 1998.
———, "A Cohort Analysis of the Association between Work Hours and Wages among Men," *Journal of Human Resources*, 37, 251–74, 2002.
Ransom, Michael R., "An Empirical Model of Discrete and Continuous Choice in Family Labor Supply," *Review of Economics and Statistics*, 69, 465–72, 1987.
Triest, Robert K., "The Effect of Income Taxation on Labor Supply in the United States," *Journal of Human Resources*, 25, 491–516, 1990.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Appendix A: First Order Conditions of Utility Maximization Problem, Labor Supply Equations, and Likelihood Function Estimated

Appendix B: Sample Details and Results for Families by Number of Children Categories, Families in 2000

Table B1: Sample Means by Number of Children, Families in 2000

Table B2: Maximum Likelihood Parameter Estimates and Utility Function Parameters, by Number of Children, Families in 2000

Table B3: Estimated Utility Function Parameters and Simulated Change in Hours, Family Income, and Utility Resulting from Effects of 2001 Tax Reform, by Number of Children, Families in 2000

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.