This PDF is a selection from a published volume from the National Bureau of Economic Research

Volume Title: Reforming the Welfare State: Recovery and Beyond in Sweden

Volume Author/Editor: Richard B. Freeman, Birgitta Swedenborg and Robert Topel, editors

Volume Publisher: The University of Chicago Press

Volume ISBN: 0-226-26192-1

Volume URL: http://www.nber.org/books/free05-1

Conference Dates: September 30-October 1, 2005

Publication Date: February 2010

Chapter Title: Labor Supply, Tax Base and Public Policy in Sweden

Chapter Author: Thomas Aronsson, James R. Walker

Chapter URL: http://www.nber.org/chapters/c5361

Chapter pages in book: (127 - 158)

4

Labor Supply, Tax Base, and Public Policy in Sweden

Thomas Aronsson and James R. Walker

4.1 Introduction

Sweden has long been at the forefront of creating innovative social insurance programs. For nearly as long, these programs have undergone intense scrutiny of effects on economic and social behavior. In terms of effects on hours of work, the consensus of opinion is that male work hours are relatively unresponsive to changes in the tax and transfer system, whereas female work hours appear to be (at least) slightly more responsive to economic incentives. Yet, this seeming nonresponsiveness of labor supply to tax and program incentives is surprising and unsettling. If incentive effects are so important, as many economists believe, why are the effects so elusive to recover empirically? Should one infer that there are *no* labor supply effects from Sweden's high marginal tax rates and generous (by U.S. standards) program benefits?

To address these and other questions, we continue the analysis of Aronsson and Walker (1997). Our earlier study gave a broad overview of incentives generated by the Swedish tax, transfer, and social insurance systems, as well as an overview of labor supply behavior, with an emphasis on studies based on Swedish data. We also related the incentive structure to existing empirical results by focusing attention on the tax and benefit reforms that took place in Sweden during the 1980s and early 1990s.

In this chapter, we consider labor supply trends and literature since our

first study. Additional explanations emerge as possible reasons for the low responsiveness of work hours to taxes and transfers. We summarize this literature, and we review the related literature on why Swedes (and Europeans, in general) appear to supply much less market work than Americans do. Are differences due to higher tax rates in Europe and Sweden than in the United States, or are they due to the functioning of the labor market?

The last decade has witnessed an expansion from hours of work to consider the incentive effects on other dimensions of labor supply. Recent research considers the relationship between the tax system and before-tax income. Pretax income reflects hours of work, and it also reflects effort, occupational choice, wage formation, savings, and to some extent, tax avoidance (depending on the income measure used). Therefore, we combine the study of tax base determination with results from traditional labor supply studies in order to get a better understanding of the relationships between economic incentives and behavior. We also briefly address relationships, on the one hand between the tax system, and on the other, the union wage formation and tax avoidance. The incentives associated with unionized labor market are also relevant from perspectives other than just wage formation—for instance, for understanding international differences with regard to work hours.

The shift in focus from hours of work to broader measures of labor supply is complemented by a shift in interest from the effects of the income tax to the effects of social insurance, and notably for Sweden, of sickness insurance and pension problems. We review this literature and consider Sweden's major pension reform of 1999. The consequences of demographic changes (particularly in connection with pay-as-you-go social security systems) have become increasingly important. The reform of 1999 is the latest example of Sweden's innovative design of social insurance programs.

In summary, we will address the following questions:

- Why are the estimated effects of taxes and transfers on hours of work typically modest? In the light of empirical evidence for Sweden, we will discuss whether social norms and quantitative constraints contribute to this outcome, as well as augment the standard labor supply model with tax avoidance behavior.
- Why are the estimated effects of marginal tax rates on before-tax income typically larger than the corresponding effects on work hours? Are issues such as tax avoidance and wage formation important here?
- What factors determine cross-country differences with respect to work hours—in particular, differences between Sweden, other European countries, and the United States?
- How does the labor supply respond to health insurance and pension systems? The effects of the pension system are particularly interesting to analyze because of (a) the major reform implemented in 1999 and (b) the aging of the population.

Although estimated labor supply elasticities on hours of work are low, as we argued in Aronsson and Walker (1997), behavioral responses to income tax and social insurance programs appear in other dimensions of labor supply. So, our answer to the second question posed in the first paragraph is no—one cannot safely assume that labor supply does not respond to economic incentives. We agree that large responses do not appear in (reported) hours of work, but for reasons presented next, we believe substantial responses appear in other dimensions. As the Organization for Economic Cooperation and Development (OECD 2005) notes in a recent report, in light of these responses, Sweden faces important challenges regarding the structure of its social insurance programs and the tax programs necessary for their finance.

Before turning to this evidence, we first consider recent trends in labor supply.

4.2   Trends in Labor Supply

Over the last fifteen years, traditional measures of labor supply have been stable. Figure 4.1 presents labor force participation rates for three age groups by gender for Sweden and the United States for 1975 to 2004. Whereas labor force participation among Swedish youth equaled or exceeded that in the United States, since 1990, labor force participation in Sweden is 10 to 15 percentage points less than in the United States. Participation rates among prime-age men are nearly identical over this period in the United States and Sweden. Among the oldest age group of men, the secular pattern of participation is quite similar across the two countries. Participation rates for this group are higher in Sweden and have been for thirty years.

Participation rates between the older two age groups of women vary between Sweden and the United States. The role of Sweden's work-related social insurance programs produce high rates of participation, as labor market participation in Sweden is substantially higher for women aged twenty-five and higher. Also evident is the severe recession of the early 1990s, which was particularly harsh on the middle age group (aged twenty-five to fifty-four); participation rates fell sharply (approximately 10 percentage points) and had not yet fully recovered. Participation for women aged fifty-five to sixty-four increased in both countries; however, Sweden's recession depressed participation rates for this group but less so than for the younger age group.

Figure 4.2 presents the average hours of work per week (measured among those working) for three age groups by gender for the United States and Sweden for 1975 through 2004.[1] Among Swedish males, hours of work declined for all three age groups, with hours falling most sharply for the youngest

---

1. These are "usual weekly hours" as reported in the Current Population Survey (CPS) and "weekly hours of work" in the Labor Force Survey (AKU).
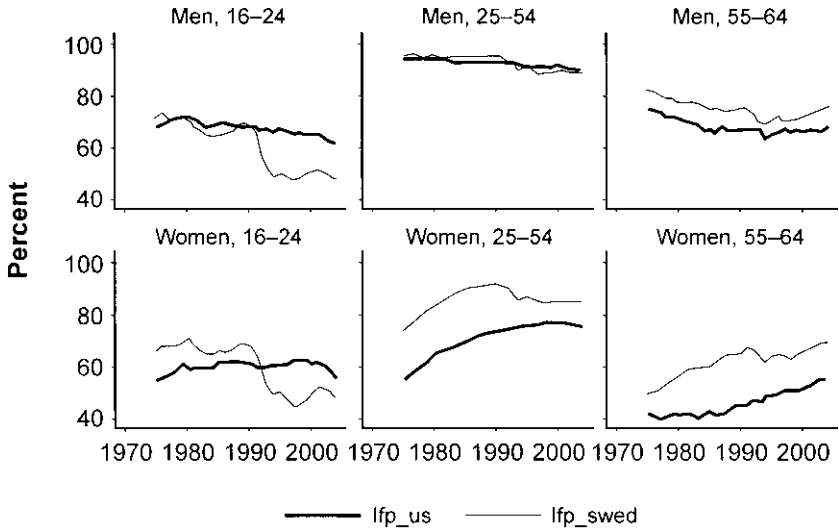
Fig. 4.1   Labor force participation by age and gender, United States and Sweden: 1975 to 2004

Note: Graphs by sex and age group.
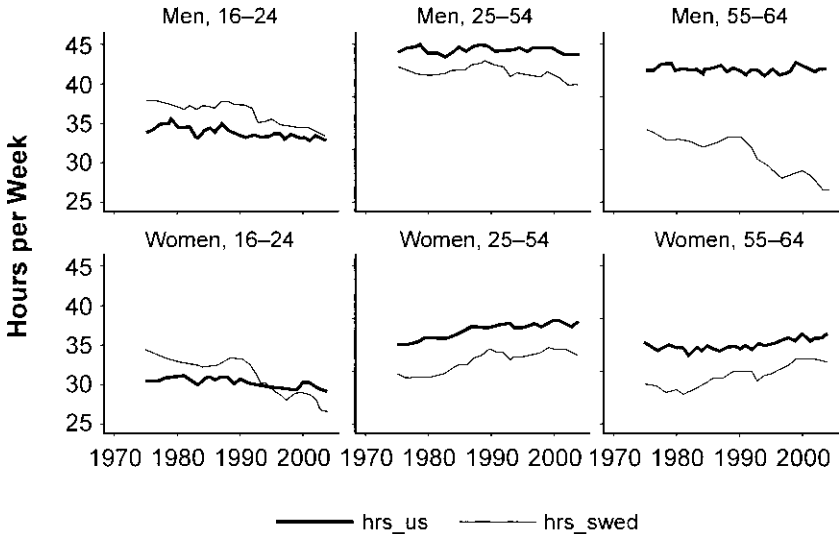
Fig. 4.2   Average hours worked per week by age and gender, United States and Sweden: 1975 to 2004

Note: Graphs by sex and age group.

group. Hours of work among prime-age workers (twenty-five to fifty-four) and mature workers (fifty-five to sixty-four) show little variation in the United States. However, in Sweden, hours of work fell during the recession in the early 1990s. Among women, hours of work declined for the youngest age group in Sweden and in the United States. For prime-age women, hours of work increased slightly in both countries. For mature women, average hours in Sweden increased and narrowed the difference with U.S. mature women.

The sector of employment changed substantially for Swedish women during the 1990s. Figure 4.3 reports the sector of employment from 1970 through 2002. Again, the recession in the early 1990s is visible, as the share of women in the local government sector declined precipitously at that time and never fully recovered. Employment in the local government sector stagnated, and women's employment increased in the second half of the 1990s as employment in the private sector increased.

Finally, figure 4.4—which is based on the OECD (2005) country report for Sweden—reports by gender the population share not at work from 1990 through 2003. Not at work includes individuals not in the labor force (e.g., those retired or in school full-time), the unemployed, and employed individuals absent from work. The last group includes individuals on disability insurance, sickness insurance, vacation, and temporary layoff.[2] We combine these different groups of nonworkers to avoid arbitrary labels of the different subpopulations.[3] The striking feature of this graph is the level of the curves—slightly more than 1 in 2.5 women and roughly 1 in 3 Swedish men are not actively working. And consistent with the trends in figures 4.1 and 4.2, the share of inactives increased during the recession of the early 1990s and stabilized at this higher level into the new century. The OECD (2005) country report for Sweden notes that Sweden has one of the highest levels of labor market inactivity levels (rivaled only by the other Nordic countries). This is an ominous trend for Sweden. Arguably, this increasing proportion of nonworkers provides long-run evidence of the work disincentives within the Swedish system that offers relatively generous (again by U.S. standards) income support with high rates of income taxation, which lessen the need of and the returns to market work.

4.3   Hours of Work: Some Evidence for Sweden and Extensions

To provide a starting point for the analyses, consider the simple labor-leisure model that forms the basis of empirical work on labor supply, where

2. The OECD (2005, 60) report notes that Sweden leads the OECD in the average number of days lost per year because of sickness. And Sweden has a relatively large share of potential workers who do not participate in the labor market because of medical reasons, especially among older age groups. Palme and Svensson (2002) report as well that sickness insurance and disability insurance systems serve as unofficial forms of unemployment insurance.

3. Thus, we want to avoid distinctions between unemployed and discouraged workers, neither of which is employed, and only the unemployed are considered in the labor force.

Numbers in 1000's   Women
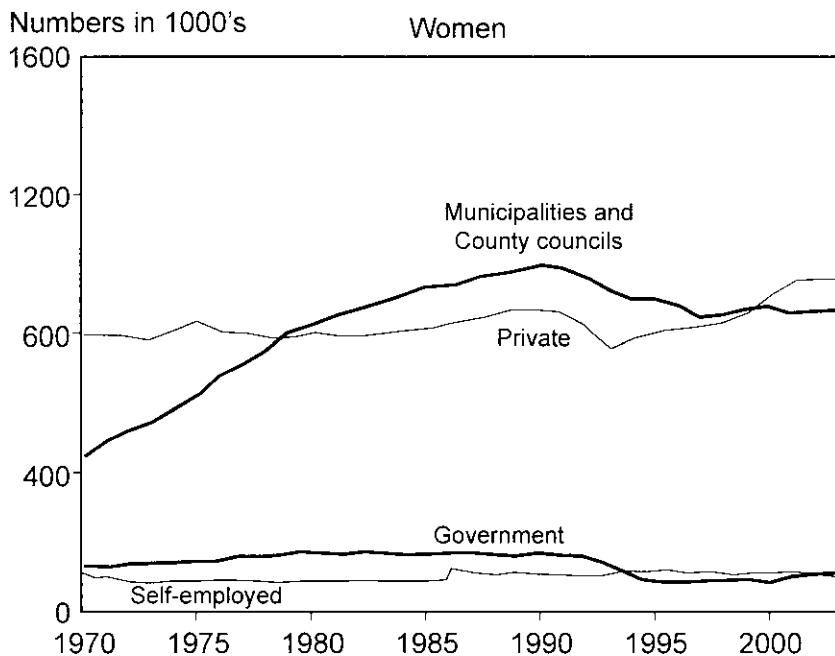
Fig. 4.3   Number of women employed, aged sixteen to sixty-four, by sector: 1970 to 2003

Source: Labor force surveys, Statistics Sweden.

the individual chooses the hours of work and the consumption of a single good to maximize utility subject to the budget constraint; that is:

$$\underset{h,c}{\text{Max}} \; u(c, h; z),$$

subject to $wh + y - T(wh, y_{\text{tax}}) - c = 0$, in which c is consumption, w is the gross wage rate, h is the hours of work, y is the nonlabor income, z is a vector of observable characteristics, y_{tax} (which is part of y) is the taxable nonlabor income, and $T(\cdot)$ is the tax payment net of transfers. The outcome of this problem can be written as a labor supply function:

(1) $h = g(w_n, y_n; z),$

where $w_n = w(1 - T')$ is the marginal wage rate, and $T' = \partial T(\cdot) / \partial(wh)$ is the marginal effect of the tax and transfer system, while y_n typically is referred to as the virtual nonlabor income (which we obtain by linearizing the budget constraint around the optimum point).

The model serves to define the basic labor supply responses to changes in the budget constraint. The income effect, $\partial h / \partial y_n$, is the change in hours of work in response to an increase in the virtual nonlabor income. If leisure is a normal good, the income effect is negative (i.e., an increase in income induces
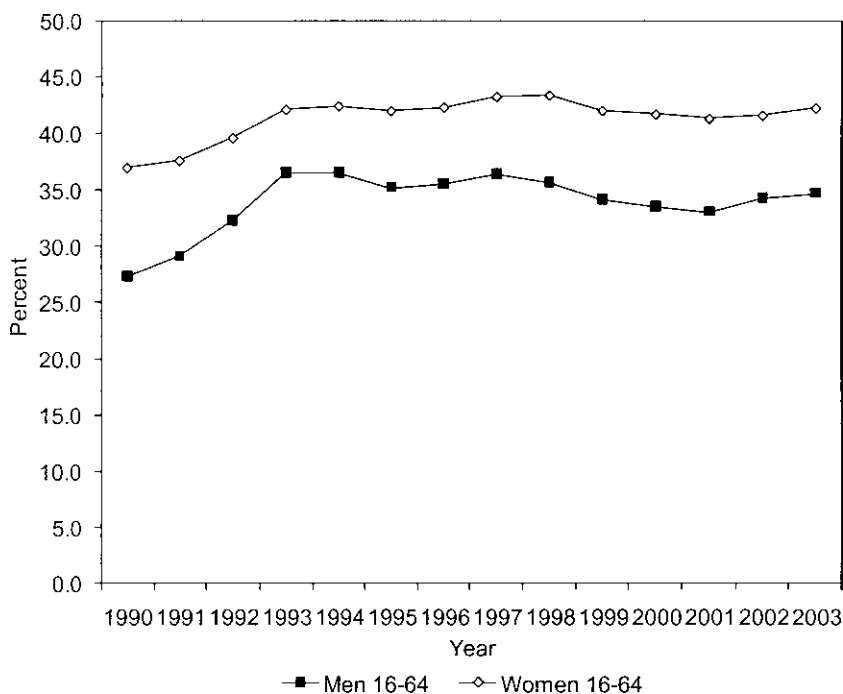
Fig. 4.4   Percent of the population, aged sixteen to sixty-four, not at work: 1980 to 2003

Source: OECD country report for Sweden (2005).

the individual to consume more leisure and thus to work less). The compensated wage effect, $\partial h/\partial w_n|_u$, measures the change in hours to an increase in the marginal wage rate, holding constant the consumer's utility. This effect captures the price or substitution effect. By the law of demand, the compensated wage effect is nonnegative. The compensated effect is important, because the welfare cost of taxes is proportional to the compensated wage effect. Finally, the uncompensated wage effect, $\partial h/\partial w_n$, summarizes the (net) change in hours with an increase in the marginal wage rate (which includes both the substitution and income effects).[4]

Blundell and MaCurdy (1999) summarize a large number of studies presenting estimates of the uncompensated labor supply elasticity with respect to the marginal wage rate (henceforth called wage elasticity) and the income elasticity. Estimates of the wage elasticity for men are typically small; most studies cover the interval from slightly negative point estimates to point

4. Readers interested in more thorough reviews of different labor supply models, as well as the econometric methods typically associated with them, are referred to Blundell and MaCurdy (1999).

estimates around 0.15. The estimates of the income elasticity are typically negative and relatively small in absolute value, although the variation between studies regarding the income elasticity appears to be greater than the corresponding variation with respect to the wage elasticity. For women, the picture is somewhat different; both the wage and the income elasticities are typically estimated to be larger in absolute value than the corresponding elasticities for men, suggesting that women's hours of work are generally more responsive to economic incentives than are male hours of work.[5] The higher wage elasticity for women may also reflect that women work fewer hours on average than men. Studies based on Swedish data yield a similar conclusion (see Aronsson and Walker 1997).[6] The appendix gives a brief summary of empirical results.

4.3.1   Responsiveness to Taxation and Programs

Clearly, one cannot infer the consequences of major reforms in the tax and transfer systems just by looking at the elasticity measures previously discussed. In this subsection, therefore, we briefly review some attempts to measure the labor supply responses to changes in the tax and transfer systems—again with a focus on results based on Swedish data. Aronsson and Palme (1998) estimate a household labor supply model using data from 1980. The model then is used to compute the desired hours of work for the husband and the wife, respectively, under three tax and transfer systems: 1980, 1989, and 1991, respectively (where the changes refer to the income tax, the value added tax, the housing allowance, and the child allowance).[7] Their results show an average 2.1 percent increase in the desired hours of work for the husband and a 0.7 percent increase for the wife due to the switch from the 1980 to the 1989 system. The main explanation for the seemingly small response in the wife's hours of work is a negative cross-wage effect (from the

5. Until recently, it was common to consider women as secondary workers, with men as primary workers. The greater labor supply responsiveness of women often was attributed to their movement in and out of the labor market.

6. See also the more recent studies of male labor supply by Blomquist, Eklöf, and Newey (2001) and Blomquist and Newey (2002) and the study of household labor supply by Flood, Hansen, and Wahlberg (2004). The own-wage elasticity of women's labor supply presented by Flood, Hansen, and Wahlberg is lower than the estimates presented in many earlier studies on labor supply based on Swedish data.

7. The most important changes refer to the income tax system. In Sweden, the income tax system contains two parts: a nonlinear national income tax and a proportional local income tax. (The local tax rate varies across local jurisdictions, and the average rate is slightly above 31 percent.) In the early 1980s, the national income tax was highly progressive with many tax brackets; for instance, in 1980, the tax rates of the national income tax schedule ranged from 0 to 58 percent, depending on the level of taxable income. In addition, labor income and capital income were taxed jointly, and capital losses were deductible from the labor income when computing the taxable income. Together with the local tax rate, an average earner may have faced a marginal income tax rate of around 50 percent, and the top rate exceeded 80 percent. The tax reform implemented during 1983 to 1985 meant splitting the national income tax into two parts: a basic part and a supplemental part. Nearly all income earners paid the basic tax, and this tax schedule had two brackets (with tax rates equal to 5 and 17 percent, respectively,

increase in the husband's marginal wage rate), accompanied by a relatively large income effect from the transfer system. The corresponding changes in desired work hours due to the shift from the 1989 to the 1991 system are 3.1 percent for the husband and 2.8 percent for the wife, respectively. Blomquist, Eklöf, and Newey (2001) use data from 1973, 1981, and 1991 to estimate the labor supply of married men aged twenty to sixty. The estimated model is used to simulate how the desired hours of work respond to the change in the tax and transfer systems between 1980 and 1991 (where the change refers to the income tax, the value added tax, the payroll tax, the child allowance, and the housing allowance). The average response in the desired hours of work is estimated to be 2.2 percent. (If evaluated in isolation, the changes in the marginal income tax rates cause the desired hours of work to increase by approximately 4 percent.) Our interpretation is that the estimated effects are relatively small, considering the magnitude of the reforms. We shall return to the 1991 reform in section 4.6, where we more fully explore the relationship between the marginal tax rate and the before-tax income.

Other types of policy-reform simulations based on Swedish data refer to combinations of deductions and transfer payments. This is interesting, because certain transfer payments may give rise to important marginal effects of relevance for the labor supply decision. Flood, Hansen, and Wahlberg (2004) estimate a household labor supply model using data from 1993 and 1999. Their model then is used to analyze the labor supply responses among married men and women to a substantial increase in the basic deduction, in combination with reduced housing allowances and welfare benefits (by 25 percent each). The results show very small (almost negligible) effects on the desired hours of work, except for low-income households (where the labor supply increase is substantial for women).

4.3.2   Does the Standard Model Capture All Relevant Effects?

As many studies on labor supply refer to time periods with increases in women's labor force participation, it is not surprising to find that women's hours of work are generally more responsive to economic incentives (measured in terms of the wage and income elasticities) than are the hours of work

---

in 1989). The supplemental tax, which contained higher tax rates, only was paid by those with sufficiently high income. Capital losses could still be deduced from the tax base for the basic tax, although it could not be deducted from the tax base associated with the supplemental tax (implying that the latter tax base was broader). The top rate for the national income tax was 42 percent in 1989. In 1991, an additional (and major) tax reform was implemented, which meant introducing a simpler national income tax system with two tax brackets for the labor income tax; the national tax rate was set to 0 for taxable incomes below 170,000 SEK (i.e., individuals with lower levels of income only paid the local income tax) and to 20 percent for higher incomes. It also implied a broader tax base by including some previously fringe benefits, as well as separate taxation of capital income (including capital gains) at 30 percent. By analogy, 30 percent of net capital losses were made deductible from the labor income tax payment. Another part of the reform was to increase the value added tax. For a more detailed discussion, see Aronsson and Walker (1997).

supplied by men. In addition, the labor supply results for women seem to vary more across studies than do the corresponding results for men, indicating (perhaps) that it is more difficult to reach a consensus regarding women's labor supply; more recent research—for example, Flood, Hansen, and Wahlberg (2004)—suggests a relatively low responsiveness of women's labor supply to economic incentives. The results from studies on male labor supply are surprisingly stable over time and across countries, in the sense that the estimates of the wage and income elasticities typically are found to be relatively small in absolute value. How can we explain this seemingly stable result? Does it mean there are no substantive responses to taxes and/or transfers?

To answer the second question, we first draw on Slemrod's (1992) behavioral hierarchy of response. This perspective recognizes that decisions involving transactions are the most responsive, followed by a variety of financial, accounting, and avoidance decisions. Real decisions involving savings, labor supply, and investment are the least responsive. In section 4.6, we consider some of the other responses to assess other possible effects of the tax. To answer the first question of why the real labor supply effects are so small, one possibility, of course, is that the standard labor supply model is correct, in which case the explanation has to do with the preferences for the trade-off between consumption and leisure. However, alongside this (unlikely) explanation, other studies try (at least in part) to attribute this finding to the character of the standard labor supply model by arguing that it neglects vital aspects of either the preferences or the choice set. We will discuss some of these approaches next by extending the standard model in order to capture quantitative constraints, social norms, and tax avoidance, respectively.

4.3.3   Quantitative Constraints

By quantitative constraints, we normally mean demand-oriented constraints, such that individuals facing them are off their labor supply curves. The presence of such constraints, at least in principle, can explain why the hours of work appear to be relatively unresponsive to small changes in the marginal wage rate and the virtual nonlabor income, respectively. Using Swedish data on prime-age married men from 1980, Sacklén (1996) addresses the issue of quantitative constraints. His study makes use of a question posed to the participants in the Level of Living Survey, where the respondents were asked to say whether they were satisfied with their work hours or if they wanted to increase or decrease their hours of work, given a corresponding change in their income. Those who claimed to be rationed in this sense were treated as if they were off their labor supply curves. This means that the labor supply of those who report rationing is treated as a latent variable. Therefore, a basic hypothesis is that once quantitative constraints are recognized and properly addressed, we may expect to find that the underlying (latent) labor supply is more responsive to economic incentives than in the standard model. However, the results do not show strong support for this hypothesis.

4.3.4   Social Norms

The standard labor supply model assumes atomistic behavior—that individuals make their labor supply decisions in isolation from one another. There is a growing literature that merges economic and sociological perspectives. One class of models incorporates social norms by augmenting individual preferences to depend on actions (and/or preferences) of others. Thus, utility declines as the individual's behavior deviates from the average choice within the group. Individual actions determine their own consumption but with possible spillover effects on others operating through the group choice. Moreover, multiple equilibria are possible and likely, as both Lindbeck, Nyberg, and Weibull (1999) and Brock and Durlauf (2001a, 2001b) show. Small changes of individual behavior operating through the social norm can magnify, leading to large changes in observed outcomes—at least over a longer time perspective. Thus, social multipliers may arise, much like the old investment multiplier of Keynesian economics in which the total effect of investment is larger than the individual marginal effect.

To exemplify, consider a slight modification of the utility function previously used, $u(c, h, \overline{h}; z)$, where $\overline{h}$ is interpretable as the average hours of work in the reference group with which the individual compares himself or herself. This variable may reflect a utility loss of deviating from a social norm regarding work hours (which here is assumed to be reflected in the average hours of work). In addition, to simplify the calculations, suppose that the utility function is such that the resulting labor supply function takes the form

$$(2) \qquad\qquad h = \alpha + \beta w_n + \gamma y_n + \delta \overline{h} + \nu z.$$

The parameters β and γ reflect the influence on hours of work of a change in the marginal wage rate and the virtual nonlabor income, respectively, conditional on $\overline{h}$. In the present context, these are only partial effects; a social multiplier exists, in the sense that if the marginal wage rate and the virtual nonlabor income facing all individuals would change (e.g., if the tax system is subject to reforms), so would the average hours of work in the reference group. Therefore, the total effect of a proportional increase in the marginal wage rate is $\beta / (1 - \delta)$, where $\delta \in [0,1)$, and $1 / (1 - \delta)$ is the social multiplier. If δ is large, we may expect substantial indirect effects via the presence of interdependent behavior. Therefore, in the standard labor supply model, if at least part of the effect of $\overline{h}$ is hidden in the constant, we may underestimate the total response to a change in the tax system.

The study of interdependent behavior in the context of labor supply models has a rich history.[8] Using Swedish data referring to prime-age married men, a variant of the model previously set out has been estimated by

---

8. See, for example, Kapteyn (1977), Alessie and Kapteyn (1991), and Woittiez and Kapteyn (1998). See also Manski (1993) for a thorough analysis of identification and Brock and Durlauf (2001a, 2001b) for a partial solution.

Aronsson, Blomquist, and Sacklén (1999). The results show that interdependent behavior is present and that it affects the labor supply. In addition, the effects are estimated to be sizeable; in terms of equation (2), the point estimate of δ is around 0.67. The partial wage and income elasticities evaluated at the mean of the data, $\beta \overline{w}_n / \overline{h}$ and $\gamma \overline{y}_n / \overline{h}$, are equal to 0.061 and –0.017, respectively, which resemble (or fall slightly short of) those found in earlier studies, which do not address interdependent behavior. The corresponding estimates of the full elasticities, $\beta \overline{w}_n / [(1 - \delta)\overline{h}]$ and $\gamma \overline{y}_n / [(1 - \delta)\overline{h}]$, are 0.187 and –0.052, respectively. Taken at face value, these results suggest that the effects of interdependent behavior are substantial and that the traditional labor supply model may imply considerable underestimation of the effects of taxation.[9] However, it is necessary to be careful in the interpretations, as we note the construction of social reference groups is arbitrary. The Monte Carlo simulations carried out by Aronsson, Blomquist, and Sacklén (1999) suggest that this is a potential problem; it may lead to the overestimation of δ, which is the parameter of main interest.

Although not formally addressed in the previous example, the time dimension is possibly very important here. As a consequence, there is no formal distinction between short run and long run. In the short run, social norms are likely to constrain behavior—suggesting moderate behavioral responses to policy—whereas social norms may evolve gradually in the long run as a result of the behavior of all individuals. If this argument is correct, then the long-run behavioral effects—which also incorporate the effects via changes in the social norm—will exceed those that follow in the short or medium run.

4.3.5   Tax Avoidance and the Choice of Assets

Clearly, many economists would expect the highly progressive income tax characterizing Sweden in the 1970s and 1980s to have caused serious disincentive effects in terms of work hours. Yet, these expected behavioral responses were not found in the empirical work based on data for that time period (at least not for men). Can asset trade, which is carried out for the purpose of avoiding taxation, explain the seemingly moderate influence taxes have on the labor supply? If it can, then the tax tables may exaggerate the effective degree of tax progression for certain groups of wage earners. We will formalize our argument by using a model developed by Agell and Persson (2000), which only requires a slight extension of the reference model previously set out. Consider an economy with two types of assets: a tax-exempt asset and a taxable asset, the returns on which now constitute the nonlabor income part of the reference model. Let e denote the initial endowment, x the tax-exempt asset, and d the taxable asset. Both assets are

---

9. Some normative implications of social norms and social interaction in the labor market are analyzed by Aronsson and Sjögren (2009).

risk free, and we give the rates of return r for the taxable asset and ρ for the tax-exempt asset. By using the wealth constraint, $d = e - x$, the optimization problem of a typical individual can be written as

$$\underset{h,c,x}{\text{Max}}\, u(c, h; z),$$

subject to $wh + \rho x + r(e - x) - T[wh + r(e - x)] - c = 0$, $x \geq 0$.

We assume that $T(\bullet)$ is such that $T'(\bullet) > 0$ and $T''(\bullet) > 0$ for $wh + r(e - x) \geq 0$. The tax system described here at least in part resembles the Swedish system before the 1983 tax reform, where negative asset income was fully deductible from labor income when calculating the taxable income. Examination of the first-order conditions suggests that when the pretax wage rate is sufficiently high, $x > 0$, whereas $x = 0$ for lower pretax wage rates. If the individual decides not to use the tax-exempt asset, then the labor supply takes the same form as in the reference model previously described. On the other hand, if the individual decides to place part of his or her endowment in the tax-exempt asset, the hours of work choice will obey the following condition:

$$(3) \qquad\qquad u_c(c, h; z)w\frac{\rho}{r} + u_h(c, h; z) = 0.$$

The labor supply implicit in equation (3) takes the form $h = h(w\rho / r, \hat{y}, z)$, where $\hat{y}$ denotes the virtual nonlabor income. This result has a very interesting implication; conditional on the ratio between the rates of return, ρ / r, a small tax reform will only give rise to an income effect. In other words, there is a distinction in terms of labor supply between tax avoiders and those who do not use the possibility to adjust their marginal income tax rate via the tax-exempt asset. For avoiders who lower their effective marginal income tax rate via avoidance behavior, the statutory marginal income tax rates exaggerate the distortions imposed by the income tax system. Therefore, avoiders in a sense will create their own labor supply incentives, suggesting that the work disincentives associated with the tax system may not have been as great as they may have appeared from the tax tables in the late 1970s and early 1980s.

Empirical research based on the Swedish tax system from the 1980s suggests that people responded to the incentives associated with tax avoidance; individuals with high labor income were found to be more inclined to have tax-favored assets than those with lower labor income (Edin, Englund, and Ekman 1995). Similarly, a decline in indebtedness seems to have taken place after the 1991 tax reform was implemented (Agell, Englund, and Södersten 1998). These results are interpretable as supporting the idea that individuals choose their portfolios in order to avoid taxation, which in turn may have implications for their labor supply behavior. In addition and perhaps even more important, if individuals respond to taxation via their portfolios, we

may expect taxation to have a larger effect on taxable income than on hours of work.

4.4 Sickness Insurance and Family Policy

A distinguishing feature of the Swedish social insurance system is its two-part structure—nearly every program offers some minimal or guaranteed benefit and another component that replaces a fraction of earnings. For example, the sickness insurance, parental insurance, and unemployment insurance systems all have this feature. The work-conditioned benefits are distinctly more generous than the guaranteed benefits and offer a strong incentive for employment. Yet, the existence of the work-conditioned benefits requires that we distinguish between contract or paid hours of employment and active hours of work—hours spent working to produce a good or service. The 2005 OECD country report on Sweden does this, and the results are striking. Workers in Sweden were on the job and actively working an average 35.4 weeks per year, while the average for Europe was 40.7 weeks. The number of weeks away from work due to holidays varied little across countries in Europe. Hence, the primary difference in annual hours worked in Sweden versus the European average was in weeks absent for other reasons. In Sweden, the average number of weeks absent (excluding holidays) was more than twice the average for Europe (9.8 versus 4.5). We next explore some of the features of the social insurance programs that may give rise to this difference.

4.4.1 Sickness Insurance

Recent reforms that require increased medical documentation of illnesses and that require employers to pay sickness benefits for the first two weeks of absence were intended to increase the monitoring of usage and thereby to tighten control over benefit costs. Nevertheless, use of sickness insurance benefits continues to remain high in Sweden.

Using the conceptual framework presented in section 4.3, it is straight-forward to explain the effect of sickness insurance benefits on labor supply. As we mentioned, demand-side considerations (e.g., setup costs and returns to coordination) may dictate the number of hours per period that must be supplied. Jobs can be seen as a tied sale of bundled attributes of which hours of work is only one of many attributes describing the working conditions and the nature of the job. Workers select the job that gives them the highest utility. Because it is cheaper to do so, employers have an incentive to offer job packages that workers prefer. Nevertheless, technology limits the choices available, and some workers may select jobs with required hours of work that are greater than desired (as defined by the labor supply function in equation [1]). Access to sickness insurance benefits permits workers to adjust their hours of work to more nearly equate desired and demanded hours. The

qualifying period and less-than-full replacement rate means the adjustment will be less than complete. In this interpretation, there is nothing stochastic or unexpected about the use of sickness insurance—the worker accepts the job demanding a fixed number of hours, knowing he or she can adjust the level of effort downward through the use of sickness insurance.

Building on the same structure, we can view the labor supply function in equation (1) as implicitly defining the (unrestricted) supply price of labor. Inverting equation (1) gives $w_n = g^{-1}(h, y_n; z_n)$, where w_n is the worker's asking price (net of taxes and transfers) to supply h units of labor at a given level of effort or intensity. We can think that the effort required on the job may vary over the business cycle, with more effort requested in expansions and less in economic contractions. Once again, sickness insurance benefits permit the worker to moderate the demands over the cycle. This framework implies that we should see a procyclical use of sickness benefits—the number of sick days should increase when demand for labor is high and should decline when demand for labor is low. We reported in Aronsson and Walker (1997) that from 1967 (when the three-day qualifying period was reduced to a single day) through the early 1990s, usage of sickness cash benefits has been strongly procyclical, with the simple correlation of unemployment and sick days as high as –0.7.

Henrekson and Persson (2004) investigate the responsiveness of sick days to the major reforms of the sickness insurance system from 1968 to 2002. The National Insurance Board (RFV) reports sick days funded by the national sickness insurance system. These register data undercount the true number of sick days during the 1990s, as occupation sickness insurance plans offered additional coverage that partially offset the cost of the qualifying period and that covered earnings above the basic amount threshold of the national system. Henrekson and Persson's aggregate time-series data do not permit parameterizing the reforms in terms of their direct monetary incentives. Instead, the authors code the reforms by time-varying indicator variables and limit their analysis to only the major reforms. They adopt a distributed lag specification, with the number of sick days as the dependent variable, and they include controls for the age composition of the working population, the gender mix, and the unemployment rate as regressors. They find that sick days and unemployment are negatively related. Indeed, they find the 1991 reform, which increased the length of the qualifying period and reduced benefits, caused sick days to decline by 20 percent. The careful and thorough analysis provides compelling evidence of the link between the structure and generosity of the sickness insurance plan and the hours of work.

4.4.2   Social Norms and Transfer Programs

Recall our earlier discussion of social norms on hours of work, where we noted that social norms may moderate behavior in the short run, whereas

over the longer run, they may generate larger effects. This is so for two reasons. First, the social multiplier may take time to work as individuals learn about reference group behavior. And second, social norms may evolve over time, providing an internal dynamic for changing behavior. Lindbeck and his coauthors (Lindbeck 1995; Lindbeck, Nyberg, and Weibull 1999, 2003) consider economies in which agents either work or receive a transfer payment and consume full leisure. Program participation entails a psychic cost (stigma), which is a decreasing function of the population living off benefits. As is common in models with social interaction, multiple equilibria exist. If the social norm for work is strong, with high psychic costs for those who deviate, then a good equilibrium with high program benefits and low free riding can be supported. Another bad equilibrium entails low benefits and high free riding. Lindbeck, Nyberg, and Weibull (1999) show that these extreme equilibria are stable compared to intermediate ones with moderate levels of program benefits and free riding. Shelling (1971) was the first to see that such economies may experience tipping, in which a small (exogenous) perturbation can shift the economy from one equilibrium to another.

Such extreme behavior has not been observed in Sweden, though its possibility is daunting. Consider the type of dynamic response in which the social norm depends on past program usage. An increased usage of program benefits today reduces the stigma of living off benefits tomorrow. Higher free riding tomorrow further lowers the cost. The tipping phenomenon is more gradual, but the system gravitates from the good equilibrium to the bad one.

The OECD (2005) reports that on an average day in Sweden, 14 percent of the working-age population is on sick leave or disability insurance. Figure 4.4 shows that the proportion of the population not at work increased dramatically during the early 1990s and remained at the higher level into the early 2000s. Disability claims show increases among women and those of younger ages. In our 1997 paper, we discussed intergenerational differences in the use of sick leave, where the results suggest that young people's usage of sick leave may be less related to medical need than to work preferences. All these figures point to a weakening of collective work ethic and to the decline of a social norm against free riding on social insurance programs. The OECD (2005) recognized this shift and called for a "change in the culture of sickness and disability" to one of "mutual obligation" to get the sick and disabled back to work as quickly as possible. If norms erode too much and a tipping point is reached, even more radical reforms will be needed.

All insurance programs face issues of moral hazard and adverse selection. Lindbeck, Nyberg, and Weibull (2003) note the difficulty of the insurance provider to develop objective criteria to distinguish between the deserving (unlucky) and undeserving (free riders) beneficiaries. When sufficiently strong, social norms may be a cost-efficient way of limiting free riding. Lindbeck, Nyberg, and Weibull (2003) argue that social norms are enforced tac-

itly by individuals who are close to the beneficiaries and who have better information than do program administrators. And enforcement of the social norms may operate through a variety of social channels, unconstrained by the formal rules and procedures that restrict program administrators. A hidden cost of weakening social norms against free riding on social insurance programs, especially sickness and disability insurance, will be increased administrative costs for additional monitoring and enforcement. In addition, stricter enforcement may imply that some of the truly needy will be denied benefits.

The social cost of continuing to allow easy access to generous social insurance programs is large. Getting people back to work has a twofold benefit of increased output and lower tax rates to cover the reduced program expenditure. Harder to measure are the reduced deadweight losses associated with the lower tax rates. We recognize these costs and realize the challenge encompasses the design of social insurance programs and their administration.

4.4.3   Parental Benefits

Historically, Sweden has been one of the leaders in offering benefits connected to childbearing and child rearing. Sweden offers subsidies to defray the medical costs of childbearing and child allowances to offset child-related expenditures necessary for young children and adolescents. As part of its family policy, parental benefits offer either parent subsidies to stay home with the child. It is important to recognize that parental benefits are one component of family policy and have the intended role of helping parents, primarily women, to balance the demands of family and the workplace. Thus, an evaluation of their effects requires a broader perspective than simply their incentive effects on labor supply. Even on labor supply, the effects are complicated and somewhat offsetting. Of course, upon childbirth, parental benefits offset the cost of leaving the market and will increase time away from work. Within the OECD framework previously described, absences while on parental benefits fall within the "absent for other reasons" category. However, guaranteed benefit levels are so low that prospective parents have an incentive to enter the labor market prior to *each* birth.[10] Empirical studies are few,[11] but the net effect for labor supply is likely to be negative, particularly as the entitlement period is substantially shorter than the benefit period.

The recent OECD country report on Sweden is critical of the level of parental benefits in Sweden, labeling the benefits as very generous and questioning their value on child development and welfare. We recognize the likely

10. See Mortensen (1977) for an elegant early analysis of entitlement effects that evaluates the employment effects of unemployment insurance within the United States.

11. See Walker (1996) for an (unsuccessful) attempt to measure the effects of parental benefits on fertility and female employment. Fertility is not completely controllable, and to the extent that births are unplanned, it weakens entitlement effects on employment.

disincentive effects on labor supply, but we are less critical of the current generosity of parental benefits. Households and individuals make decisions within the current and anticipated social insurance programs. For example, private-savings decisions are made with knowledge of the benefit levels of the public pension system, and fertility decisions are made when anticipating the programs and subsidies available to families with children. Reductions in child-related subsidies and programs that are interpreted as signaling the future policies to be even less supportive to families with children may decrease fertility. In light of Sweden's aging population, policies that may lower fertility merit further attention.

4.5   Public Pensions

Sweden retains its position as leader in innovative social insurance programs with its pension reform of 1999 that completely restructured the public pension system. The prereform pension system was a pay-as-you-go (PAYG) defined benefit system that had two components: a basic pension and a supplemental pension. The basic pension offered income security and established the minimum income floor for all workers. The supplement pension augmented the basic pension and targeted payments to replace about 60 percent of the worker's labor market earnings. Under the supplemental pension system, benefits were tied to the worker's highest earnings over fifteen years, and to obtain a full supplement benefit, an individual had to work thirty years. Over time, however, benefits became more generous, life expectancies rose, and birthrates fell. Thus, pensioners had to be supported for more years and were entitled to more generous benefits, with a declining worker base to support them. These changes necessitated either large tax increases or a significant cut in benefits.

4.5.1   The 1999 Reform

The Swedish pension reform of 1999 offers an innovative and in many ways radical solution to these problems.[12] The largest change is that the public pension system switched from a defined benefit to a *notional defined contribution* (NDC) system. In the new system, workers have an individual account, which receives each year a pension contribution equal to 18.5 percent of their pensionable income. Now, instead of guaranteed benefits upon retirement, workers have guaranteed contributions credited each year to their accounts. It is *notional* because workers' accounts are only credited; actual assets are not set aside. Balances in the account earn a rate of return each year equal to annual growth in average wages. Upon retire-

---

12. Like the previous system, the new pension system will be phased in over a number of birth cohorts; specifically, the birth cohorts of 1938 to 1953.

ment (after age sixty-one), the individual's accumulated notional wealth is converted into an annuity, dependent on the life expectancy of the individual's cohort.

A second innovation of the 1999 reform introduced limited individual control over the pension. Contributions equal to 16 percentage points go into the national public pension fund, which is managed by the government. The remaining 2.5 percent contribution is mandatory; however, these funds are under the individual's control. Hence, the government will collect all revenues and will disperse payouts, although the individual funds will be managed by private investment firms that are selected by the individual from an approved list of investment advisors. The last innovation is a significant broadening of the income base entering the public pension system. Besides earnings, the 1999 reforms recognized additional social insurance payments as part of the income base. This includes stipends to students pursuing postsecondary education and parental benefits. Indeed, child-rearing activities earn supplemental contributions, as does income while serving in the military.

The Swedish pension system remains a pay-as-you-go pension system, as tax contributions paid by those currently working are used to finance the pension payments of those currently retired. At an 18.5 percent contribution rate, the system's designers expect the new pension system to deliver pension benefits at about the same level as the old system. An interesting political compromise rests with the 18.5 percent contribution rate. Consensus estimates within the public debate was that a 16 percent contribution would suffice to maintain the (then) existing level of pension benefits. Yet, to ensure income security, a higher contribution rate was accepted, with the stipulation that the remaining funds were under individual control.

The new pension system represents a striking departure from Sweden's long history of guaranteed benefits and horizontal equity. Under the new system, different cohorts will receive different pensions for the same accumulated pension wealth as life expectancies change. Members of the same birth cohort will receive different pensions for the same life expectancy should the individuals retire in different years (and thus have the opportunity to accumulate different levels of pension wealth to annuitize at retirement). Finally, and the most individual specific, the individual accounts will accumulate wealth based on the individual's investment strategy. Individuals adopting more conservative portfolios may accrue less wealth than those who adopt more aggressive portfolios. Government oversight of the fund managers and restrictions on the types of investments permitted reduce this form of variability, and it may be more symbolic than real. Yet, within a social insurance system founded on the notion of horizontal equity in which alikes are treated alike, this source of individual-level diversity is a striking departure from the traditional Swedish social insurance system.

4.5.2 Pensions and Labor Supply

There is a vast empirical literature on the old pension system[13] and very little (because of its newness) on the 1999 pension reform. We note, however, that between 1999 and 2004, labor force participation rates increased among the elderly (fifty-five to sixty-four). Future research will determine whether the work incentives created by the 1999 pension reform increased these rates. We offer some observations that suggest it may have.

The shift from the old defined benefit PAYG system to the notional defined contribution system will induce both short-run and long-run effects on labor supply. In the long run, as Lindbeck (2002) and Feldstein and Leibman (2002) show, a notional or quasifunded benefit system can reduce the distortionary incentive effects caused by a tax on labor to finance benefits. That is, in a PAYG system in which there is *no* relationship between taxes paid and benefits eventually received, the implicit rate of return on tax contribution is minus 100 percent, and the deadweight loss reflects the entire pension tax on labor income. Under an NDC, taxes and benefits are related, as taxes paid are returned as future benefits, with an implicit rate of return equal to the growth rate of the population multiplied by the growth rate of labor income. If the implicit rate of return on pension contributions is close to the rate of return on other forms of savings the individual could make otherwise, the distortion of labor supply is much reduced. Indeed, if the implicit rate of return exactly equals the return on other forms of savings, then the distortion on labor supply is zero. Lindbeck and Persson (2003) show that for reasonable growth rates, the reduction in the tax wedge on labor income could be as much as 10 percentage points. The reduction in marginal tax rates should therefore increase labor supply (via participation, hours, and quality). The reduction of deadweight loss generated by the NDC may potentially improve welfare for both young and old generations.[14]

The short-term labor supply incentive effects are more transparent. The switch from a defined benefit to a defined contribution plan changes the nature of pension accruals. Under the NDC, another year of work will increase pension wealth because of the direct contribution from earnings, the return on accumulated contributions, and if the person is past normal retirement age, the increased annuity values for shorter benefit period. In the defined benefit system, the pension depends on years of employment and average earnings over the highest fifteen years of earnings. For someone with the same level of pension wealth and with fewer than thirty years of employ-

---

13. See the various studies by Mårten Palme; for example, Palme and Svensson (2004).

14. Importantly, the introduction of the defined contribution plan converts entitlements to contributions made rather than to benefits to be paid. Hence, the government is protected from having to guarantee benefits in the presence of increasing life expectancies and below-replacement birth rates.

ment or with earnings above their fifteenth-highest earnings, an additional year of employment thus would increase the annuity paid at retirement by increasing either the replacement rate or the average earnings (or both) used within the benefit formula. However, most workers near retirement age have worked for more than thirty years and thus under the defined benefit system have weak incentives to continue working. Therefore, the reformed Swedish pension plan now offers stronger work incentives, especially to those with persistent labor market attachment.

Yet, one cautionary remark is in order. Palme and Svensson (2002) estimate that approximately 20 percent of men and 27 percent of women exit the labor market via the social insurance programs. Under the 1999 reform, social insurance income is pensionable and thus may exacerbate this tendency. Moreover, recent work on social norms suggests that this problem is even more pernicious.

4.6   Other Aspects of the Tax Base

So far, we mainly have focused on hours of work. From a tax revenue perspective, it is of interest to understand how and why taxation affects before-tax income. In addition, by extending the analysis to before-tax income, several additional mechanisms will appear; for instance, in addition to hours of work, before-tax income also reflects effort, occupational choice, wage formation, and possibly also tax avoidance. By Seldrod's hierarchy of response, we may expect taxable income to respond differently to tax policy changes in comparison with hours of work.

4.6.1   The Relationship between Marginal Tax Rates and Before-Tax Income

The literature on tax base determination, initiated by Lindsay (1987) and further developed by Feldstein (1995), offers an interesting complement to the traditional study of labor supply, because it allows us to address several aspects of behavior simultaneously. In other words, although the framework for studying tax base determination typically resembles the traditional labor supply model, it does not restrict its attention solely to hours of work. Depending on the definition of income, it also (at least in principle) may capture avoidance, as well as the effects of tax policy on the gross wage rate, effort, occupational choice, and savings. Therefore, part of the difference in results among studies likely is to be due to differences regarding the way in which income is measured—for example, whether the analysis refers to taxable income or a broader income concept.

In their comprehensive study on tax base determination in the United States, Gruber and Saez (2002) make a distinction between two income concepts: taxable income and broad income, where the latter is defined as

the sum of all items that compose total income less capital gains.[15] A basic hypothesis is that taxable income is more sensitive to marginal taxation than is broad income, because changes in taxable income also reflect (some aspects of) tax avoidance. Indeed, this is precisely what Gruber and Saez find. First, the elasticity of taxable income with respect to the net of tax rate (measured as one minus the marginal tax rate) is estimated to be around 0.4 on average.[16] The corresponding elasticity for broad income is 0.07. Second, income effects are very small, implying that the uncompensated and compensated effects on the before-tax income of a change in the net of tax rate are similar. Third, the effect on taxable income of a change in the net of tax rate appears to be strongest in the upper part of the income distribution.

There are several studies based on Swedish data dealing with the influence of marginal tax rates on before-tax income. Let us start by briefly discussing a paper by Ljunge and Ragan (2004). Instead of considering taxable income and broad income, respectively, as in the paper by Gruber and Saez, Ljunge and Ragan focus their attention on the determinants of labor earnings. Although this choice of dependent variable means neglecting some of the possible effects of taxation due to tax avoidance, it nevertheless will reflect the joint effect of taxation on several interesting aspects of behavior. As such, we may expect the behavioral responses to be different from those presented for work hours. Ljunge and Ragan use panel data and focus on individuals aged twenty-five to fifty-five during 1989. Their study period is from 1989 to 1994, meaning that attention is paid to the 1991 tax reform in terms of its effect on labor earnings. The results imply that the compensated elasticity of earnings with respect to the net of tax rate is around 0.33 on average, which is a relatively large response. (Recall that the corresponding elasticity for taxable income in the United States is measured to be around 0.4 by Gruber and Saez.) The uncompensated elasticity of labor earnings with respect to the net of tax rate is estimated to be in the interval of 0.25 to 0.33. In addition, there is considerable variation across groups. For instance, both low-income earners and high-income earners have higher estimated elasticities than income earners in the middle of the distribution. On average, the estimated behavioral response to the 1991 tax reform is an increase in earnings by 10 to 15 percent.

Using panel data for the years 1989 and 1992 and focusing on the group aged twenty-five to sixty in 1989, Hansson (2004) estimates the elasticity of

15. Broad income contains wages, salaries and tips, interest income, dividends, alimony received, business income, total Individual Retirement Account (IRA) distributions, total pensions and annuities, income reported on schedule E, farm income, unemployment income, and other income.

16. This number is an average estimate in comparison with earlier literature. Studies concentrating on the upper part of the income distribution typically estimate a higher number. On the other hand, some earlier studies that do not refer to the United States seem to imply a much weaker relationship between marginal tax rates and taxable income; see, for instance, Aarbu and Thoresen (2001) and Sillamaa and Veall (2001).

taxable labor income with respect to the net of tax rate to be 0.43 or 0.57 on average, depending on the choice of instruments for the net of tax rate. She also finds that the relevant elasticity may differ considerably across groups (with women being more responsive than men in at least one of the models). However, because she does not incorporate a measure of virtual income in the analysis, it is not clear whether the relevant elasticity estimate should be interpreted as a compensated or an uncompensated effect. The results also appear to be somewhat sensitive to the use of the estimation method (the choice of instrument for the marginal tax rate). Selén (2002) estimates a model similar to that of Gruber and Saez using data for the period from 1989 to 1992. He concentrates the analysis to men in the group aged twenty-five to fifty-five in 1989. His results imply a slightly lower estimate of the (compensated) elasticity of taxable income with respect to the net of tax rate, in the interval of 0.2 to 0.4. Finally, Blomquist and Sehlin (2008) estimate how taxable labor earnings as well as the hourly gross wage rate depend on the net of tax rate. By using data for 1981 and 1991, their preferred estimates imply that the elasticity of taxable earnings with respect to the net of tax rate is 0.26 for men and 0.75 for women (again indicating that women respond more to tax policy than do men), whereas the corresponding elasticities for the hourly gross wage rates are 0.20 and 0.33, respectively.[17] We will return to the relationship between gross wage rates and the tax system next.

Although we believe that the study of how marginal tax rates affect before-tax income is interesting in the sense that it provides a broader view of tax responses than the traditional study of labor supply, it is important to be careful when interpreting these results. This is so for at least two reasons. First, this research area is relatively new by comparison—at least when applied to Swedish data. To be useful for purposes of policy evaluations and recommendations, these studies should be supplemented by additional research.[18] Second, and more importantly, the theoretical foundations for the study of before-tax income is not always convincing; the income supply model by Gruber and Saez is only one possible model, and it resembles the

17. Blomquist and Sehlin also recognize that responses to tax policy are likely to be asymmetric; for instance, although a decrease in the marginal tax rate may lead to additional investments in human capital or a move toward jobs with higher wages, an increase in the marginal tax rate is not likely to lead to disinvestment in skills or moves toward jobs with lower pay. Indeed, when reestimating their models using similar data for other time periods, they obtain results that differ substantially from those referred to earlier.

18. As we previously mentioned, most earlier Swedish studies use the 1991 tax reform to identify the effects of marginal taxation on the before-tax income. An exception is Holmlund and Söderström (2008), who use data for the postreform period from 1991 to 2002; in addition, they consider a dynamic regression model, which allows them to distinguish between short-run and long-run effects of a change in the net of tax rate. For men, the estimates of the long-run elasticity of income with respect to the net of tax rate typically are found in the interval of 0.1 to 0.3, whereas the estimates for women are imprecise and statistically insignificant. They also calculate the fiscal consequences of a reform that reduces the top marginal income tax rate by 5 percentage points, and they find that this reform is likely to have only a minor effect on tax revenues.

labor supply model too much to provide a very interesting alternative. For instance, the income supply model itself is silent about important issues such as wage formation, as well as about the opportunities and constraints underlying avoidance behavior. As a consequence, we may not learn much about the mechanisms underlying behavior by studying only the relationships discussed here.

4.6.2   Wages, Taxes, and Tax Progression

One reason as to why marginal income taxation causes a larger response in the tax base than in the hours of work was discussed in section 4.3, where we argued that individuals may adjust to taxation via tax avoidance, which in turn tends to diminish the effects of taxation on hours of work. Alongside this explanation, empirical evidence also suggests that the wage formation system may contribute to explain the apparently large effect on the tax base following the 1991 tax reform. If the pretax wage rates are determined by bargaining between unions and firms, which is common in European labor markets, standard models for wage setting, as well as some of the empirical evidence associated with them, predict that an increase (a decrease) in the marginal tax rate with the average tax rate held constant leads to decreased (increased) pretax wage rates.[19] The intuition is based on a tradeoff (which the trade union is assumed to be facing) between the consumption wage per employed member and the number of employed members: a higher marginal tax rate tends to increase the opportunity cost of wage increases in terms of lost employment. On the other hand, a change in the average tax rate with the marginal tax rate held constant may either increase or decrease the pretax wage rate.

Let us discuss studies based on Swedish data. Holmlund and Kolm (1995) use data from the Swedish income distribution surveys (HINK) which include time series data for different income groups, as well as microlevel panel data.[20] In the panel data regressions, the results suggest that an increase in the marginal tax rate of 10 percentage points will reduce the pretax wage by about 4 to 6 percent for the average worker with the average tax rate held constant, whereas the corresponding number in the time series regression is around –2.5 percent. There is also evidence suggesting that this effect is stronger among the highest income earners. Aronsson, Wikström, and Brännlund (1997) estimate a union wage model based on panel data at the firm level for the Swedish pulp and paper industry. Their results imply that with the average tax rate held constant, the elasticity of the pretax wage rate with respect to the marginal tax rate is –0.5 on average. The corresponding elasticity with respect to the average tax rate (with the marginal tax rate held constant) is estimated to be 0.5.

19. Normative implications of unionized labor markets are analyzed, for example, by Fuest and Huber (1997) and Aronsson and Sjögren (2004a, 2004b).

20. Their model essentially is based on the seminal work by Lockwood and Manning (1993).

Despite arguments for caution,[21] it is tempting to compare the results on tax progression and wage formation with those discussed earlier on the determinants of the hours of work and before-tax income, respectively. Such a comparison provides at least one possible explanation as to why the labor earnings most likely increased more than the hours of work as a result of the 1991 income tax reform. In addition, this argument is further strengthened (from yet another perspective) by Blomquist and Sehlin (2008), who show that the difference between the tax base elasticity and the wage rate elasticity for men previously referred to is very close to the typical labor supply elasticity estimated in earlier studies (indicating that the study of tax base determination gives results that are consistent with the labor supply literature). Interestingly, therefore, their results suggest that the response in labor earnings to a change in the marginal tax rate that is due to the response in the wage rate is more important than the response in the hours of work—at least for men.

4.7   Cross-Country Comparisons

Most earlier studies on work hours, and particularly on how hours of work respond to taxation, transfer payments, and social insurance, are based on within-country microdata. More recently, however, several studies[22] have emerged with the explicit purpose of explaining cross-country differences, with a focus on differences between Europe and the United States. These differences have increased in general since the early 1970s and are now substantial: Americans tend to do much more market work than Europeans do. An important question is whether the differences in work hours between Europe and the United States can be explained by differences in taxation and transfer programs or whether they are (mainly) attributable to other factors, such as labor market institutions and legislation.

Differences between Europe and the United States with respect to work hours per person have been analyzed, for example, by Prescott (2004) and Olovsson (2004) in the context of numerical general equilibrium models. From our perspective, the study by Olovsson is particularly interesting, as it deals explicitly with a comparison between Sweden and the United States. Olovsson makes a distinction between market work and household production: the main difference between the two countries refers to how the households divide their time between market work and household production (more time is spent on market work and less time in household production in

21. Note that these strong results are not confirmed fully by more recent research based on data for other countries. For instance, by using Danish data, Lockwood, Slök, and Tranaes (2000) find evidence suggesting that the effects of higher tax progression on the pretax wage rate is income dependent. Similarly, Brunello, Parisi, and Sonedda (2002) use Italian data and find the opposite effect in comparison with the Swedish studies previously referred to: tax progression works to increase the pretax wage rate.

22. See, for example, Davis and Henrekson (2004), Prescott (2004), Olovsson (2004), and Alesina, Glaesen, and Sacerdote (2005).

the United States in comparison with Sweden), whereas the total time spent working is about the same in both countries. Olovsson uses an intertemporal model of a competitive economy, where preference and production parameters are set to reflect Olovsson's assessment of estimates in the empirical literature, whereas the policy parameters are chosen either to reflect Sweden or the United States. This enables him to compare how differences in public policy between the two countries affect the time spent in market work and household production, respectively. Adopting a long-run perspective of comparing steady states, Olovsson shows that differences in public policy (and in particular, income taxation) between the two countries can explain the differences with respect to how the households divide their time between market work and household production: high marginal tax rates in Sweden induce individuals to substitute from market to nonmarket production, with the total hours of work virtually equal between the United States and Sweden. His model predicts a gradual reduction in the hours spent in market work in Sweden between 1960 and 1980, the direction of which is consistent with the observed behavior, although it overestimates the hours of market work per person at the beginning of the period (implying that the predicted reduction is greater than the reduction that actually took place).

Although it is likely that differences in tax policy between Sweden (or Europe, in the paper by Prescott) and the United States give rise to differences in the hours of work per person, there are at least three problems with the preceding analyses. First, the expenditure side of the government's budget is dealt with in a somewhat superficial way; neither Olovsson nor Prescott allow public consumption to enter the utility function, and both of them give the tax revenues (net of the useless public consumption) back to the consumers in the form of lump-sum transfers. The latter weakens the income effect so that the substitution effect becomes dominant for the behavioral response to tax policy. Clearly, if the government spending (or part thereof) were not treated as a perfect substitute for private consumption, the income effect would tend to offset the behavioral responses to tax policy—an argument also put forward in a comment by Ljungqvist (2006). Second, the social insurance system gives rise to its own incentives, which are likely to contribute to differences in the hours of work per person. For instance, differences between countries with respect to the generosity of unemployment benefits may imply differences in the incentives among the unemployed to actively search for employment. This aspect is particularly interesting in light of how the difference between Europe and the United States with regards to the hours of work per person refers to differences in terms of employment-to-population ratios. Although employment-to-population ratios are generally lower in Europe than in the United States, the differences appear to be particularly large for the youngest and oldest age groups (see Gordon 2006). Generous unemployment benefits are likely to increase the reservation wage rates. Therefore, if human capital depreciates

with unemployment, it becomes more difficult for the unemployed to find acceptable jobs. This in turn may contribute to reduce the search intensity. (See also Ljungqvist and Sargent [2005], as well as their chapter 6 in this volume.) As it is less beneficial for older individuals than for younger individuals to reinvest, this also suggests that generous unemployment benefits may have a relatively strong effect on the incentives facing the unemployed elderly. It is also consistent with the estimates by Palme and Svensson (2002) that unemployment is an important pathway to early retirement in Sweden. Third, the literature focusing on differences in tax policy between Europe and the United States as an explanation to differences in terms of the hours of work is often silent about the labor market structure—an issue to which we turn next.

4.7.1   Market Structure and the Role of Unions

Is there econometric evidence in favor of the hypothesis that differences (between countries) in economic policy, and particularly in tax policy, are related to differences in hours of market work? The answer is yes; there are studies showing a negative correlation between hours of work per person and the (average) marginal income tax rate. Alesina, Glaeser, and Sacerdote (2005) discuss the possibility that these observed correlations to some extent may be due to a missing variables problem. Their study focuses on differences in labor market institutions between the United States and Europe: European labor markets are often characterized by trade unions—which have a strong influence on wage formation and/or serve as a pressure group behind the economic policy—whereas the influence of trade unions is much weaker in the United States. They use panel data for the OECD countries for the period from 1960 to 1995. If they disregard the differences in labor market characteristics previously discussed (in a way similar to earlier literature in this area), their results show a negative and significant relationship between the marginal tax rate and the hours of work per person, which is consistent with the results previously mentioned. However, by adding a measure of union density (the fraction of union members in the labor force) and a measure of employment protection, this significant result disappears; instead, both the union density and the employment protection variable show negative and significant relationships with hours of work. Therefore, it seems as if unionization and regulations better explain differences between Europe and the United States than does the tax system.

The idea that unions are able to affect work hours (either directly or indirectly via wage formation) becomes more plausible with stronger trade unions. Unions may engage in pressure group activities; due to differences in membership and/or political strength, the expected return from doing so is likely to have been greater in some of the major European countries than in the United States. As such, unions may have had greater opportunity to affect market outcomes, as well as legislation, in Europe.

We do not want to draw strong conclusions from these results. This is so for several reasons. First, studies based on country-level data cannot provide as detailed descriptions of incentive effects as the microstudies discussed before, meaning that direct comparisons are difficult. Second, although union wage setting and measures of regulations may correlate with hours of work per person at the country level, so does the marginal tax rate, according to the results presented by Alesina, Glaeser, and Sacerdote (2005). Therefore, it very well may be the case that the tax and transfer systems, the social insurance system, and the institutional characteristics of the labor market all contribute to explain the differences with regard to work hours, and we are not yet able to establish which aspect is most important.

4.8   Summary and Discussion

Sweden's diverse set of social insurance programs provides a high safety net for its citizens but requires a high rate of taxation. Social insurance programs create their own disincentives for market work. Historically, Sweden's programs have been robust to individual malfeasance, yet there is emerging evidence that malfeasance is on the rise. The large and growing proportion of the near-retirement population on disability and sickness insurance benefits is disturbing and may mask other structural problems in the labor market. Equally worrisome, it seems that the disability and sickness insurance programs are now entitlements rather than insurance.

The high tax rates necessary to fund the social insurance programs also create incentives against market work. The 1991 tax reform reduced marginal tax rates, but tax rates in Sweden remain high compared with non-Nordic European countries and particularly the United States. The econometric literature finds little evidence that high marginal tax rates are reducing hours of work in Sweden. The literature offers a variety of reasons to explain why the distorting effects of high tax rates may not appear in hours of work. However, while taxes may not have very large effects on hours of work, there is emerging evidence that the high marginal tax rates do influence other choices that affect before-tax income. In the absence of any behavioral effects, an increase in the marginal income tax should have no effect on before-tax income. Recent estimates using Swedish data suggest that a 10 percent increase in marginal tax rate reduces before-tax income by 3 to 4 percent.

International comparisons reveal that Swedes work fewer hours in the market than Americans do. The income guarantees offered by the social insurance programs and high marginal tax rates are certainly part of the explanation, though there is far from a consensus on their relative importance. And it is easy to recommend that tax rates be lowered or that social insurance programs be reduced, but doing so would miss the Swedish perspective on the right level of social insurance. A fruitful approach is to

inquire whether the social insurance programs meet their original intent with the least social cost.

Sweden's 1991 tax reform and the 1999 pension reform increased incentives for market work. These reforms addressed politically contentious problems and illustrate Sweden's remarkable ability to find pragmatic public policy solutions. The challenge facing the Swedish welfare system is to keep enough people working enough hours to fund its generous benefits. To preserve its high safety net, Sweden will need to increase self-insurance of small (short-term) risks and may need to increase the monitoring of benefits to limit the free riding that stems from the apparent shift in norms away from market work.

Appendix

Table 4A.1

Study	Description	e_w	e_y
	Results for men		
Blomquist (1983)	LNU 1973, married men aged 25–55	0.08	–0.04
Blomquist and Hansson-Brusewitz (1990)	LNU 1981, married men aged 25–55	(0.08, 0.12)	(–0.13, 0.02)
Flood and MaCurdy (1992)	HUS, married men aged 25–55	(–0.24, 0.2)	(–0.10, 0.04)
Aronsson and Palme (1998)	LNU 1981, married couples aged 25–55	0.12	–0.03
Flood, Hansen, and Wahlberg (2004)	HINK 1993, 1999, married couples, each spouse younger than age 56	0.05	—
Blomquist, Eklöf, and Newey (2001)	LNU 1973, 1981, and 1991, married men aged 25–55	0.075	–0.04
	Results for women		
Blomquist and Hansson-Brusewitz (1990)	LNU 1981, married women aged 25–55	(0.38, 0.79)	(–0.24, –0.03)
Aronsson and Palme (1998)	LNU 1981, married couples aged 25–55	0.44	–0.12
Flood, Hansen, and Wahlberg (2004)	HINK 1993, 1999, married couples, each spouse younger than age 56	0.1	—

Note: The variable e_w measures the elasticity of the labor supply with respect to the own marginal wage rate, whereas e_y is the labor supply elasticity with respect to the virtual nonlabor income. In some of the studies, the authors estimate several variants of the model, and the numbers refer to the intervals for the point estimates. LNU = Level of Living Survey; HUS = Swedish Panel Study on Market and Nonmarket Activities; HINK = Swedish Household Income Survey.

References

Aarbu, K., and T. Thoresen. 2001. Income responses to tax changes: Evidence from the Norwegian tax reform. *National Tax Journal* 54 (June): 319–35.

Agell, J., P. Englund, and J. Södersten. 1998. *Incentives and redistribution in the welfare state: The Swedish tax reform.* Basingstoke, U.K.: Macmillan Press.

Agell, J., and M. Persson. 2000. Tax arbitrage and labor supply. *Journal of Public Economics* 78 (1/2): 3–24.

Alesina, A., E. Glaeser, and B. Sacerdote. 2005. Work and leisure in the U.S. and Europe: Why so different? NBER Working Paper no. 11278. Cambridge, MA: National Bureau of Economic Research, April.

Alessie, R., and A. Kapteyn. 1991. Habit formation, interdependent preferences and demographic effects in the almost ideal demand system. *Economic Journal* 101 (406): 404–19.

Aronsson, T., S. Blomquist, and H. Sacklén. 1999. Identifying interdependent behaviour in an empirical model of labour supply. *Journal of Applied Econometrics* 14 (6): 607–26.

Aronsson, T., and M. Palme. 1998. A decade of tax and benefit reforms in Sweden: Effects on labour supply, welfare and inequality. *Economica* 65 (1): 39–67.

Aronsson, T., and T. Sjögren. 2004a. Efficient taxation, wage bargaining and policy coordination. *Journal of Public Economics* 88 (12): 2711–25.

———. 2004b. Is the optimal labor income tax progressive in a unionized economy? *Scandinavian Journal of Economics* 106 (4): 661–75.

———. 2009. Optimal income taxation and social norms in the labor market. *International Tax and Public Finance,* forthcoming.

Aronsson, T., and J. R. Walker. 1997. The effects of Sweden's welfare state on labor supply incentives. In *The welfare state in transition: Reforming the Swedish model,* ed. R. Freeman, B. Swedenborg, and R. Topel, 203–67. Chicago: University of Chicago Press.

Aronsson, T., M. Wikström, and R. Brännlund. 1997. Wage determination under non-linear taxes: Estimation and an application to panel data. *Oxford Economic Papers* 49 (3): 404–18.

Blomquist, S. 1983. The effect of income taxation on the labor supply of married men in Sweden. *Journal of Public Economics* 22 (2): 169–97.

Blomquist, S., M. Eklöf, and W. Newey. 2001. Tax reform evaluation using non-parametric methods: Sweden 1980–1991. *Journal of Public Economics* 79 (3): 543–68.

Blomquist, S., and U. Hansson-Brusewitz. 1990. The effect of taxes on male and female labor supply in Sweden. *Journal of Human Resources* 25 (3): 317–57.

Blomquist, S., and W. Newey. 2002. Nonparametric estimation with nonlinear budget sets. *Econometrica* 70 (6): 2455–80.

Blomquist, S., and H. Sehlin. 2008. Hourly wage rate and earnings responsiveness to changes in marginal tax rates. Working Paper no. 2008:16. Uppsala University, Department of Economics, November. Available at: http://www.nek.uu.se/Pdf/wp2008_16.pdf.

Blundell, R., and T. MaCurdy. 1999. Labor supply: A review of alternative approaches. In *Handbook of labor economics,* vol. 3A, ed. O. Ashenfelter and D. Card, 1559–695. New York: North-Holland.

Brock, W., and S. Durlauf. 2001a. Discrete choice with social interactions. *Review of Economic Studies* 68 (2): 236–60.

———. 2001b. Interactions-based models. In *Handbook of econometrics,* vol. 5, ed. J. Heckman and E. Leamer, 3297–380. New York: North-Holland.

Brunello, G., M. Parisi, and D. Sonedda. 2002. Labor taxes and wages: Evidence from Italy. CESifo Working Paper no. 715. Munich: Center for Economic Studies and the Information and Forshung Institute for Economic Research at the University of Munich, May.

Davis, S., and M. Henrekson. 2004. Tax effects on work activity, industry mix and shadow economy size: Evidence from rich-country comparisons. NBER Working Paper no. 10509. Cambridge, MA: National Bureau of Economic Research, May.

Edin, P.-A., P. Englund, and E. Ekman. 1995. Avregleringar och hushållens skulder. In *Bankerna under krisen.* Stockholm: Bankkriskommittén.

Feldstein, M. 1995. The effect of marginal tax rates on taxable income: A panel study of the 1986 Tax Reform Act. *Journal of Political Economy* 103 (3): 551–72.

Feldstein, M., and J. Liebman. 2002. Social security. In *Handbook of public economics,* vol 4, ed. A. Auerbach and M. Feldstein, 2245–324. New York: North-Holland.

Flood, L., J. Hansen, and R. Wahlberg. 2004. Household labor supply and welfare participation in Sweden. *Journal of Human Resources* 39 (4): 1008–32.

Flood, L., and T. MaCurdy. 1992. Work disincentive effects of taxes: An empirical analysis of Swedish men. *Carnegie-Rochester Conference Series on Public Policy* 37: 239–77.

Fuest, C., and B. Huber. 1997. Wage bargaining, labor-tax progression, and welfare. *Journal of Economics* 66 (2): 127–50.

Gordon, R. 2006. Issues in the comparison of welfare between Europe and the United States. Unpublished manuscript. Northwestern University, Department of Economics.

Gruber, J., and E. Saez. 2002. The elasticity of taxable income: Evidence and implications. *Journal of Public Economics* 84 (1): 1–32.

Hansson, Å. 2004. Taxpayers' responsiveness to tax changes and implications for the costs of taxation. Working Paper no. 2004:5. Lund University, Department of Economics, February.

Henrekson, M., and M. Persson. 2004. The effects on sick leave of changes in the sickness insurance system. *Journal of Labor Economics* 22 (1): 87–113.

Holmlund, B., and A.-S. Kolm. 1995. Progressive taxation, wage setting and unemployment: Theory and Swedish evidence. *Swedish Economic Policy Review* 2 (2): 424–70.

Holmlund, B., and M. Söderström. 2008. Estimating dynamic income responses to tax reforms: Swedish evidence. IFAU Working Paper no. 28. Uppsala, Sweden: Institute for Labor Market Policy Evaluation.

Kapteyn, A. 1977. A theory of preference formation. PhD diss., Leiden University.

Lindbeck, A. 1995. Welfare state disincentives with endogenous habits and norms. *Scandinavian Journal of Economics* 97 (4): 477–94.

Lindbeck, A. 2002. Pensions and contemporary socioeconomic change. In *Social security and pension reform in Europe,* ed. M. Feldstein and H. Siebert, 19–49. Chicago: University of Chicago Press.

Lindbeck, A., S. Nyberg, and J. Weibull. 1999. Social norms and economic incentives in the welfare state. *Quarterly Journal of Economics* 114 (1): 1–35.

———. 2003. Social norms and welfare state dynamics. *Journal of the European Economic Association* 1 (2/3): 533–42.

Lindbeck, A., and M. Persson. 2003. The gains from pension reform. *Journal of Economic Literature* 41 (1): 74–112.

Lindsay, T. 1987. Individual tax payer response to tax cuts: 1982–1984, with implications for the revenue maximizing tax rate. *Journal of Public Economics* 33 (2): 173–206.

Ljunge, M., and K. Ragan. 2004. Who responded to the tax reform of the century? Working Paper. University of Chicago, Department of Economics.

Ljungqvist, L. 2006. Comment on "Work and leisure in the U.S. and Europe: Why so different?" by A. Alesina, E. Glaeser, and B. Sacerdote. In *NBER macroeconomics annual 2005,* ed. M. Gertler and K. Rogoff, 65–77. Cambridge, MA: MIT Press.

Ljungqvist, L., and T. Sargent. 2005. The European unemployment experience: Uncertainty and heterogeneity. Unpublished manuscript. New York University, Department of Economics. Available at: http://homepages.nyu.edu/~ts43/RESEARCH/europe_employment_1.pdf.

Lockwood, B., and A. Manning. 1993. Wage setting and the tax system: Theory and evidence for the United Kingdom. *Journal of Public Economics* 52 (1): 1–29.

Lockwood, B., T. Slök, and T. Tranaes. 2000. Progressive taxation and wage setting: Some evidence from Denmark. *Scandinavian Journal of Economics* 102 (4): 707–23.

Manski, C. 1993. Identification of endogenous social effects: The reflection problem. *Review of Economic Studies* 60 (3): 531–42.

Mortensen, D. 1977. Unemployment insurance and job search decisions. *Industrial and Labor Relations Review* 30 (4): 505–17.

Olovsson, C. 2004. Why do Europeans work so little? IIES Seminar Paper no. 727. Stockholm University, Institute for International Economic Studies, February.

Organization for Economic Cooperation and Development (OECD). 2005. *OECD economic surveys: Sweden.* Paris: OECD.

Palme, M., and I. Svensson. 2002. Pathways to retirement and retirement incentives in Sweden. Arbetsrapport no. 2002:9. Stockholm: Institute for Future Studies.

———. 2004. Income security programs and retirement in Sweden. In *Social security and retirement around the world: Micro-estimation,* ed. J. Gruber and D. Wise, 579–642. Chicago: University of Chicago Press.

Prescott, E. 2004. Why do Americans work so much more than Europeans? *Federal Reserve Bank of Minneapolis Quarterly Review* 28 (1): 2–13.

Sacklén, H. 1996. Essays on empirical models of labor supply. PhD diss., Uppsala University.

Selén, J. 2002. Taxable income responses to tax changes: A panel analysis of the 1991 Swedish tax reform. FIEF Working Paper no. 177. Stockholm: La Fédération Internationale pour l'Economie Familiale.

Shelling, T. 1971. Dynamic models of segregation. *Journal of Mathematical Sociology* 1:143–86.

Sillamaa, M.-A., and M. Veall. 2001. The effects of marginal tax rates on taxable income: A panel study of the 1988 tax flattening in Canada. *Journal of Public Economics* 80 (3): 341–56.

Slemrod, J. 1992. Do taxes matter? Lessons from the 1980s. *American Economic Review* 82 (2): 250–6.

Walker, J. R. 1996. Parental benefits, employment and fertility dynamics. In *Research in population economics,* vol. 8, ed. T. P. Schultz, 125–72. Greenwich, CT: JAI Press.

Woittiez, I., and A. Kapteyn. 1998. Social interaction and habit formation in a model of female labour supply. *Journal of Public Economics* 70 (2): 185–205.