Journal of Public Economics 122 (2015) 94–109

Contents lists available at ScienceDirect

Journal of Public Economics

journal homepage: www.elsevier.com/locate/jpube

Optimal disability insurance with unobservable skill heterogeneity☆

Kyung-woo Lee

School of Management and Economics, Handong Global University, 558 Handongro, Pohang, Kyungbuk 791-708, Republic of Korea

ARTICLE INFO

Article history:
Received 5 May 2013
Received in revised form 30 August 2014
Accepted 11 November 2014
Available online 22 November 2014

Keywords:
Disability insurance
Income taxation
Social security

ABSTRACT

This paper studies the U.S. income tax and disability insurance systems in a unified framework with relevant private information and a general social welfare function. To that end, I develop a dynamic model in which agents differ in labor productivity and disability status, which are both private information. The optimal allocation can exhibit negative marginal labor income tax rates for high-skilled individuals but zero for low-skilled individuals, which is contrary to standard results. Also, I find that a reform to the optimal system in the calibrated model would significantly improve the U.S. system. The welfare gains amount to a 3.48 percent increase in consumption with a utilitarian social welfare function, which might be larger with a non-utilitarian function. Better insurance for people in bad states and reduced intratemporal distortions account for almost all the welfare gains. However, intertemporal distortions do not appear to play a major role.

1. Introduction

The Social Security Disability Insurance (SSDI) program is one of the largest social insurance programs in the United States. In 2013, this program provided nearly 11 million people with benefits worth over 140 billion dollars.[1] Because SSDI has grown so much in size and importance, its effectiveness in providing insurance to people with disabilities must be assessed and the possibilities for improving the program analyzed. To address these questions adequately, SSDI must be scrutinized in a dynamic model with uncertain disability shocks and other relevant life-cycle components, and the optimal system should be similarly scrutinized. However, no such analysis has been attempted in the literature on disability insurance, to the best of my knowledge. Moreover, disability insurance has typically been studied in isolation from other tax and social insurance programs[2]; these analyses might also be insufficient or even misleading if economic shocks that are relevant to other programs are systematically intertwined with disability risk. Thus, it would be desirable to incorporate relevant shocks to the analysis of disability insurance.

Motivated by these issues in the literature, this paper provides a comprehensive welfare analysis of the U.S. income tax and disability insurance systems in a dynamic model with private information for both labor productivity and disability. Individuals are either high-skilled or low-skilled throughout their lives, but they can become disabled permanently during any period. However, both their skill types and their disability status are only privately observed. In this economy, the government should provide adequate insurance for people in bad states (low-skilled or disabled) and incentivize those in good states (high-skilled or able). The incentive problem can be quite complex because the two sources of private information interact with one another throughout their life cycles. In addition, the social welfare function is allowed to be non-utilitarian. Thus, the government should take into account that one type might be favored in redistribution.

This paper tackles the two issues in the literature discussed above. First, disability insurance is unified with non-linear income taxation, which addresses unobserved differences in ability in a single model with the general social welfare function. In particular, I ask whether and how these model features change the standard results in either the literature on non-linear income taxation or the literature on disability insurance. I show that the interaction between the two sources of private information and the general social welfare function can significantly alter the standard optimal allocation rules. Second, this paper quantifies the welfare gains that result when the optimal system of income tax and disability insurance is adopted (relative, of course, to the U.S. system) and investigates the factors that account for such welfare gains. For this purpose, I calibrate the model to the U.S. economy, taking advantage of data sources such as the Panel Study of Income Dynamics (PSID). To the best of my knowledge, this is the first paper that quantitatively evaluates the welfare effects of the U.S. SSDI program relative to the optimal system in a rigorously calibrated dynamic model with privately observed disability shocks.

The findings of this paper are as follows. Theoretically, this paper presents three characteristics of the optimal allocation that mark a

☆ I thank the co-editor, Pierre Pestieau, and two anonymous referees of this journal for their valuable comments and suggestions. I am especially indebted to Stefania Albanesi, Bernard Salanié and Pierre Yared for their support, encouragement and helpful discussions. I am also grateful to John Donaldson, Chulyoung Kim, Young Sik Kim, Narayana Kocherlakota, Wojciech Kopczuk, Jaromir Nosal, Stephanie Schmitt-Grohé, Jon Steinsson, Martin Uribe, Michael Woodford, and participants in various seminars and conferences for their comments and discussions.

E-mail address: kwlee@handong.edu.

[1] Source: http://www.ssa.gov/oact/STATS/. Although dependents of disabled workers also receive SSDI benefits, most beneficiaries are disabled workers (81.4% in 2013).

[2] See Diamond and Mirrlees (1978) and Golosov and Tsyvinski (2006), for example.

http://dx.doi.org/10.1016/j.jpubeco.2014.11.004
0047-2727/

sharp departure from the standard principles found in the literature. First, high-skilled agents can face non-standard distortions in their consumption and labor; otherwise, low-skilled but able agents would claim to be high-skilled but disabled. When some of the incentive compatibility (IC) conditions for the "double deviations" are binding, the planner should reduce the consumption for high-skilled but disabled agents such that this state does not look attractive to low-skilled agents. In addition, the planner should raise the capital income tax rates for the high-skilled to prevent excessive saving for self-insurance against disability, which is motivated by the low disability consumption. These distortions represent a utility loss to the high-skilled. Thus, the high-skilled should be additionally incentivized by a marginal subsidy to labor income. I show that the three-way distortions are not just possible in theory but are actually observed in the optimal allocation of the model that is calibrated to the U.S. economy.

Second, in contrast to the distortions for the high-skilled, the allocation for low-skilled agents can exhibit zero marginal labor income tax rates if the welfare weight on the low-skilled is sufficiently low. In this case, their lifetime utility becomes sufficiently low that high-skilled agents find no reason to deviate. Therefore, the planner does not need to distort the allocation for the low-skilled to make it less attractive to the high-skilled. Consequently, the labor allocation for the low-skilled can be left undistorted.

In the quantitative welfare analysis, I find that the reform to the optimal system yields large welfare gains. With the utilitarian social welfare function, the welfare gains amount to a 3.48 percent increase in consumption. These welfare gains are then decomposed into improvements to the following aspects of allocation: i) 3.7% of the total gains are caused by reducing intertemporal distortions, ii) 54.5% are due to correcting excessive intratemporal distortions and the redistribution from able to disabled agents, and iii) 41.8% are the result of the redistribution from high-skilled to low-skilled agents. Thus, the lion's share of the welfare gains is attributed to better insurance for agents in bad states and the improvement in intratemporal allocation. This result suggests that the U.S. income tax and disability insurance system does not insure people well against adverse shocks to their ability, despite creating severe intratemporal distortions.

In the case of the non-utilitarian social welfare function, the optimal policy might achieve even larger welfare gains when either skill type is particularly favored in the redistribution. For example, when the social welfare function cares only about low-skilled agents, adopting the optimal system would be equivalent to raising consumption by 9.64% and the redistribution between the skill types would account for 84.4% of the welfare gains. Intuitively, in a society that favors one skill type over another, the redistribution to the favored type matters more for social welfare and can therefore lead to substantial welfare gains.

1.1. Related literature

This paper studies the optimal design of disability insurance and income tax with multiple sources of private information. Thus, it is related to the large volume of literature on non-linear optimal taxation following Mirrlees (1971) (see Albanesi, 2008, for a review). This study also draws from the growing literature on the New Dynamic Public Finance (Golosov et al., 2006; Kocherlakota, 2010). However, my paper is more directly related to papers such as those of Diamond and Mirrlees (1978) and Golosov and Tsyvinski (2006), which study the optimal disability insurance when disability is private information. In particular, I extend the model developed in Golosov and Tsyvinski (2006) by allowing privately observed skill heterogeneity and find that non-standard distortions can emerge for the top skill type.

As with my paper, certain papers examine models with multiple sources of private information. Judd and Su (2006) generate the distortion for the top skill type, but only numerically. In addition, Beaudry et al. (2009) study a redistribution problem when productivity is unobserved in both market and non-market activities. Weinzierl (2011) finds that if an agent's relative position on the skill distribution scale varies with age, there might be distortions at the top. Nevertheless, my paper is most closely related to that of Lozachmeur (2006), which is a two-period version of this model. In addition to generalizing his model, my paper adopts a general social welfare function and conducts a rigorous quantitative welfare analysis. It is notable that in the industrial organization literature, Rochet and Choné (1998) and Armstrong (1996), among others, analyze a non-linear pricing problem with multi-dimensional unobserved characteristics.

Several papers conduct quantitative welfare analyses on tax or social insurance programs with a dynamic Mirrlees approach. Weinzierl (2011) calculates the welfare gains from age-dependent taxes and Hosseini (2010) investigates the welfare effect of a mandatory Social Security program when the annuity market has the adverse selection problem. My paper is similar to that of Huggett and Parra (2010) in terms of quantitative methodology because they engage in an exercise similar to that employed in this study to evaluate the performance of the U.S. social insurance system. Without disability shocks, their paper can be interpreted as concentrating on redistribution across different productivity types. Finally, Farhi and Werning (2012) assess the welfare gains from achieving constrained efficiency in intertemporal allocation in a general-equilibrium model with private information. My paper also utilizes their approach in decomposing the welfare gains from the optimal reform.

1.2. Structure of the paper

The remainder of this paper is organized as follows. Section 2 introduces the model and the social welfare function. Section 3 characterizes the optimal allocation of this model and describes how its key features can modify the standard rules for optimal allocation. Section 4 approximates the U.S. income tax and the SSDI program and finds the equilibrium allocation. In Section 5, I introduce a number of partial reforms that can be used to decompose the welfare gains from the optimal system. In Section 6, I calibrate the model and quantify the welfare gains from the reforms introduced in Section 5 and also investigate the sources of the welfare gains through a decomposition analysis. Finally, Section 7 concludes. The Appendix provides proofs for the main results.

2. The model

2.1. Agents' preferences and production technology

The economy is populated by agents who live for $S+1$ periods from age 0 to S. All agents consume every period but only able agents under the mandatory retirement age z can work. This model nests an economy without a mandatory retirement age as a special case with $z = S+1$.

When the agents work, they produce output by the production function $y = wl$, where y and w denote output and labor productivity, respectively. I assume that $w = 0$ for retirees and disabled agents and $w > 0$ for able agents. Throughout this paper, I assume that *information on ability is private*, i.e., both w and l are observed only privately, whereas output y is publicly observed. Thus, if a government or a planner assigns some y to agents, it should incentivize them appropriately.

In every period, agents evaluate their well-being by the utility function $u(c) - v(l)$, where $u(c)$ and $v(l)$ are the consumption utility and labor disutility, respectively, given consumption c and labor effort l. I make standard assumptions that $u' > 0$, $u'' < 0$, $v' > 0$, $v'' > 0$ and normalize $v(0) = 0$.

2.2. Privately observed shocks

There are two sources of private information on labor productivity: productivity type and disability status. In other words, only agents know whether they are high-skilled (type H) or low-skilled (type L) and whether they are able or disabled. To describe the evolution of