THREE ESSAYS ON MICROFOUNDATIONS OF ECONOMICS

A Dissertation

by

GAOSHENG JU

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

August 2011

Major Subject: Economics

THREE ESSAYS ON MICROFOUNDATIONS OF ECONOMICS

A Dissertation

by

GAOSHENG JU

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Approved by:

Co-Chairs of Committee,	Qi Li
	Li Gan
Committee Members,	Leonard Auernheimer
	Rajiv Sarin
	Ximing Wu
Head of Department,	Timothy Gronberg

August 2011

Major Subject: Economics

ABSTRACT

Three Essays on Microfoundations of Economics. (August 2011)

Gaosheng Ju, B.S., Shandong Institute of Economics;

M.S., Peking University

Co–Chairs of Advisory Committee: Dr. Qi Li
Dr. Li Gan

This dissertation, which consists of three essays, studies three applications. Each of them emphasizes the microfoundations of economic models.

The first essay proposes a nonparametric estimation of structural labor supply and exact welfare change under nonconvex piecewise-linear budget sets. Different from previous literature, my method focuses on a nonparametric specification of an indirect utility function. I find that working with the indirect utility function is very useful in simultaneously addressing the labor supply problems with individual heterogeneity, nonconvex budget sets, labor nonparticipation, and measurement errors in working hours that previous literature was unable to. Further, the estimated indirect utility function proves to be convenient and efficient in calculating exact welfare change and deadweight loss under general piecewise-linear budget sets.

In the second essay, I solve the equity premium, risk-free rate, and capital structure puzzles by laying a more solid microfoundation for consumption-based asset pricing models. I argue that the above two asset pricing puzzles arise from the aggregation of hump-shaped life-cycle consumption into per capita consumption, which accounts for the unanimous rejections of Euler equations in the literature. As for the third puzzle, I show that a firm's capital structure can be determined by heterogenous

investors maximizing life-time utility even though the capital structure is irrelevant on the firm side. The endogenously determined leverage generates an even larger equity premium than a fixed one.

The third essay studies the solution concepts of coalition equilibrium. Traditional solution concepts such as Strong Nash Equilibrium, Coalition-proof Nash Equilibrium, Largest Consistent Set, and Coalition Equilibrium violate the fundamental principles of individual rationality. I define a new solution concept, Weak Coalition Equilibrium, which requires each coalitional deviation to be within-coalition self-enforceable and cross-coalition self-enforceable. The cross-coalition self-enforceability endows coalitions with farsightedness. Weak Coalition Equilibrium is a generalization of Coalition-proof Nash Equilibrium and a refinement of the concept Nash Equilibrium. It exists under a weak condition. Most importantly, it is in line with the principle of individual rationality.

To my parents and my wife, Li Yuan.

TABLE OF CONTENTS

CHAPTER Page

LIST OF TABLES

TABLE Page

LIST OF FIGURES

CHAPTER I

INTRODUCTION

The term microfoundations typically appears in the literature of macroeconomics. It refers to microeconomic analysis of the behavior of individual agents that underpins a macroeconomic theory (Barro (1993)). However, the importance of microfoundations is not limited within macroeconomics. In this dissertation, I show that a multitude of other economic applications such as welfare analysis, finance and coalition analysis need a solid microfoundation, too.

The contribution of my dissertation is not about methodology. I will not offer a general theory or method on how to lay solid microfoundations for different economic applications. It seems impossible. On the contrary, I will present three essays, each of which addresses its own issues by building models with solid microfoundations. Therefore, the dissertation has a multiple of contributions spread among three essays.

This first essay contributes to the literature in both the estimation of structural labor supply and the calculation of exact welfare effects. It proposes a nonparametric method to estimate labor supply with nonconvex piecewise-linear budget sets. Different from previous literature such as Blomquist and Newey (2002) and Soest, Das, and Gong (2002), our method focuses on a nonparametric specification of an indirect utility function. I find that working with the indirect utility function is very useful in simultaneously addressing the labor supply problems with individual heterogeneity, nonconvex budget sets, labor nonparticipation, and measurement errors in working hours that previous literature was unable to.

Further, the estimated indirect utility function proves to be convenient and effi-

The journal model is *Econometrica*.

cient in calculating exact welfare change and deadweight loss under general piecewise-linear budget sets. Previous welfare calculation in a nonparametric framework relies on Vartia (1983)'s numerical method (Hausman and Newey (1995), and Kumar (2008)). I show that Vartia's Method is computationally cumbersome particularly when the budget set is nonconvex. With the estimated indirect utility function, I define a generalized indirect utility that is a function of the entire budget set. I am able to estimate the desired choices under piecewise-linear budget sets before and after the tax reform. Also, the utility attained at the choice points are computable with the ordinary indirect utility function. This generalized indirect function greatly facilitates an ex ante assessment of tax reforms.

This essay includes applications to the 1986 tax reform and 2001 Bush tax cut. I find that the estimates of labor elasticities and the exact measure of deadweight loss are sensitive to the functional specifications of preferences. In addition, convexification of nonconvex budget sets may severely bias the estimation of deadweight loss incurred by the reform of Earned Income Tax Credit (EITC) program.

The microfoundation in the above applications lies in the heterogeneous preferences among individuals. In the literature of labor supply estimation with nonparametric methods, researchers usually estimate conditional mean labor supply. The unobserved heterogeneity in preferences thus is integrated out. Correspondingly, the estimation can be interpreted as the preference of an aggregate individual. However, the estimated aggregate labor supply is not proper for the calculation of welfare change due to income tax reforms, particularly when the budget set is piecewise-linear. For example, the tax reform of EITC enormously changes the shape of the budget sets near zero working hours, but slightly modifies the budget segments on which the choice of an aggregate individual falls. An economist using the estimated mean labor supply may conclude that the welfare effect of the reform is near zero since the choice the aggregate individual is unchanged. However, it is no doubt that welfare effects on those eligible for EITC program are significant. In this essay, I find that the 1986 reform of EITC generates a deadweight loss which was not identified in the literature.

The second essay contributes to the literature by solving the equity premium puzzle, risk-free rate puzzle, and capital structure puzzle. I find that all three puzzles arise from weak microfoundations of traditional models. First, two asset pricing puzzles are able to be resolved if the hump-shaped life-cycle consumption curve is applied in an asset pricing model. In my overlapping-generations model which incorporates life cycle features, the calibrated equity premium and risk-free rate are consistent with the observed U.S. historical data. The degree of risk aversion used for the calibration is 1.2, which is plausible and much lower than those in the literature. The hump-shaped consumption curve is the key to resolving the two asset pricing puzzles. Aggregating the consumption across different generations into per capita consumption leads to the non-orthogonality between the stochastic pricing kernel and the market excess return rate, which is the key problem of asset pricing models. Second, the capital structure is resolved as the investors' behaviors are taken into account. Modigliani and Miller (MM) theorem (1958) states that the capital structure of a firm is irrelevant for the firm to maximize its value. However, the assumption of value maximization is not justified in the literature and lacks microfoundations. I address the determination of capital structure in an asset pricing framework. I show that investors who are endowed with life-cycle features and maximize their life-time utility can collectively determine the capital structure by optimizing their asset portfolios. The calibrated debt-to-capital ratio is 1/3, which is close to the observation in U.S. market.

What distinguishes my solutions from those in the previous literature is that I address the capital structure puzzle and asset pricing puzzles in one common framework. First, the determination of capital structure is related to the prices of assets because stock and bond owners make investment decisions based on asset returns. Second, asset prices which represent owners' claims of future payoffs are no doubt mostly determined by the capital structure of firms which determines the future payoffs of assets. Therefore, it seems natural to address three puzzles together in one model.

In the third essay, I study the solution concepts of coalition equilibrium. Traditional solution concepts such as Strong Nash Equilibrium, Coalition-proof Nash Equilibrium, Largest Consistent Set, and Coalition Equilibrium violate the fundamental principles of individual rationality. The violation reveals the lack of microfoundations in the study of group behavior. I define a new solution concept, Weak Coalition Equilibrium, which requires each coalitional deviation to be within-coalition self-enforceable and cross-coalition self-enforceable. The cross-coalition self-enforceability endows coalitions with farsightedness. Weak Coalition Equilibrium is a generalization of Coalition-proof Nash Equilibrium and a refinement of the concept Nash Equilibrium. It exists under a weak condition. Most importantly, it is in line with the principle of individual rationality.

CHAPTER II

NONPARAMETRIC ESTIMATION OF STRUCTURAL LABOR SUPPLY AND EXACT WELFARE CHANGE UNDER NONCONVEX PIECEWISE-LINEAR BUDGET SETS

2.1. Introduction

This essay contributes to the literature in both the estimation of structural labor supply and the calculation of exact welfare effects. It proposes a nonparametric method to estimate labor supply with nonconvex piecewise-linear budget sets. Different from previous literature such as Blomquist and Newey (2002) and Soest, Das, and Gong (2002), my method focuses on a nonparametric specification of indirect utility function. I find that working with the indirect utility function is very useful in simultaneously addressing the labor supply problems with individual heterogeneity, nonconvex budget sets, labor nonparticipation, and measurement errors in working hours that previous literature was unable to.

Further, the estimated indirect utility function proves to be convenient and efficient in calculating exact welfare change and deadweight loss under general piecewise-linear budget sets. Previous welfare calculation in a nonparametric framework relies on Vartia's (1983) numerical method (see Hausman and Newey (1995) and Kumar (2008)). I show that Vartia's Method is computationally cumbersome particularly when the budget set is nonconvex. With the estimated indirect utility function, I define a generalized indirect utility that is a function of the entire budget set. This function facilitates an ex ante assessment of tax reforms.

I apply the method to estimate the labor supply for married women using the Panel Study of Income Dynamics (PSID) data of 1983 and 2000. I find the estimates of labor elasticities are sensitive to the functional specifications of preferences. Using 1983 PSID data, the labor supply elasticity based on the preferred third order polynomials is only half of the elasticity based on the second order polynomials. In welfare calculations, I find that the estimates of the exact deadweight loss are sensitive to the functional specifications of preferences. In addition, convexification of nonconvex budget sets may severely bias the estimation of deadweight loss incurred by the reform of Earned Income Tax Credit (EITC) program.

It is of considerable interest to estimate the labor supply and exact welfare effects with general piecewise-linear budget sets. The progressive (federal and state) income taxes, payroll taxes, and various transfer programs such as Earned Income Tax Credits (EITC) create nonconvex piecewise-linear budget constraints. Each tax reform in the U.S. history abolished or introduced some tax rules. The tax reforms changed each individual's nonlinear budget set to another nonlinear budget set. It is desirable to precisely predict the labor choice on the budget sets and accurately evaluate the welfare change and deadweight loss before the complex reforms are implemented. Accounting for the complete form of the piecewise-linear budget constraints is necessary. The nonlinearity of tax systems results in a simultaneity problem between labor choices and net wages. A large body of work including Burtless and Hausman (1978), Hausman (1985), Kumar (2008), and among others, found evidence of downward bias in wage elasticity estimates if the nonlinearity of the budget sets is ignored. An estimate of welfare change based on the biased labor supply is, of course, misleading.

Another important factor in estimating the labor supply is the specification of the labor supply function. Many researchers prefer a simple (typically linear) form of labor supply primarily for its manageability (see Burtless and Hausman (1978); Hausman (1981b, 1985); Triest (1990); Fullerton and Gan (2004) among others). A simple linear specification of labor supply has the advantages of being able to obtain closed-form indirect utility function by solving a differential equation. Further, the closed-form direct utility function can be derived from the indirect utility function by solving an optimization problem. Both the direct and indirect utility functions are vital to calculations of welfare change and the desired working hour under a nonconvex budget set. However, these simple linear models are shown to have resulted in biased estimates in labor elasticity. When measuring deadweight loss, a misspecified model is especially troublesome since the "second order" properties of the labor supply curve is critical (Hausman (1981a)). Moreover, Brown and Walker (1989) argued that the unobserved heterogeneity in preferences should enter into the basic labor supply[1] function in a nonlinear form. Otherwise, the Slutsky matrix becomes asymmetric, violating the axioms of rationality. Unfortunately, Brown and Walker's work has not drawn much attention in the literature.

In the literature, because of the potential specification bias, several attempts have been made in nonparametric specifications of labor supply under piecewise-linear budget sets. Blomquist and Newey (2002) proposed a nonparametric method that treats the labor supply as a function of the entire budget set.[2] They showed that, if the budget set and preference are convex, the conditional mean hour of working given the budget set is simply the sum of unconditional expected working hours on all segments and kink points. However, their technique does not work in a general context. In the presence of a nonconvex budget set, the desired working hour is discontinuous with respect to unobserved heterogeneity. Thus, some parts of the budget set will not contribute to the conditional mean working hour. It is impossible

[1]Blomquist (1988) called the labor supply functions that would be generated by linear budget constraints the basic labor supply functions, and the labor supply functions generated by nonlinear budget constraints the mongrel labor supply functions.

[2]Liang (2009) extended Blomquist and Newey's (2002) method to handle cases with labor nonparticipation.

to correctly determine these non-contributing parts without a global algorithm. Soest, Das, and Gong (2002) presented another nonparametric method in which the direct utility function is approximated with a series expansion and the budget set is replaced by a finite number of points. The unresolved challenge for Soest et al. (2002) is how to deal with the measurement error in the working hour which enters into the direct utility function in a nonlinear fashion. The measurement error in the working hour, as Soest et al (2002) pointed out, can have a detrimental effect on the estimates of policy relevant parameters.

This essay presents a new method to nonparametrically estimate the labor supply and the exact welfare measures. Instead of focusing on working hours or direct utility, I focus on the nonparametrically specified indirect utility function, and obtain the labor supply function via Roy's Identity. In particular, the desired labor supply is a function of three arguments: the entire budget set, the ordinary indirect utility function, and the heterogeneity of preference for working. Contrast to Blomquist and Newey (2002), I maintain the heterogeneity of preferences in an indirect utility function with a flexible functional form. The importance of unobserved heterogeneity is widely recognized in the labor supply literature. Also, Fullerton and Gan (2004) argued that accounting for the unobserved heterogeneity is crucial in the welfare calculation if the budget is piecewise-linear. Ignoring the unobserved heterogeneity may lead to biased estimation of the exact welfare change. In a traditional nonparametric regression model, the unobserved heterogeneity in preferences is integrated out, as in both Hausman and Newey (1995) and Blomquist and Newey (2002). The regression component in the nonparametric model thus is explained as the labor supply of an average individual. However, the aggregation leads to biased estimates of welfare effects if Hausman and Newey's (1995) two-step method were applied in a context with piecewise-linear budget sets. In my model, the functional of the ordinary indirect utility and the distribution of unobserved heterogeneity can be estimated using a simulated likelihood method.

When indirect utility function is specified, the desired labor supply is easily calculated. For every individual, his/her basic labor supply is easily derived from the indirect utility function via Roy's identity. Then the basic labor supply function can be used to locally determine whether the choice falls on a particular segment or a kink point. If the budget set is nonconvex, the local analysis may predict that more than one segment or kink point are possible to attain the maximum utility. I then compare the indirect utility derived from these segments or kink points to infer the global optimal choice.

A key technique involved in the global algorithm is how to derive the indirect utility at a kink point. My simple idea is to work out the supporting line that is tangent with an indifference curve at this kink point (see Blundell, et al., (1988)[3]). I design an efficient numerical algorithm to solve the tangent line. The algorithm converges to the root at an exponential rate. Once the supporting line is available, the utility at the kink point can be calculated by substituting the slope and intercept of the tangent line into the ordinary indirect utility function. I call the maximum indirect utility derived among all the segments and kink points a generalized indirect utility.

In terms of welfare calculation, a nonparametric specification of the indirect utility function proves to be convenient and efficient. Hausman and Newey's (1995) is the first to examine the exact welfare changes in a nonparametric framework. Their approach is a two-step method. In the first step, they applied nonparametric regression models to the estimation of ordinary demand curves. Then, in the second

[3]Blundell, et al. searched for the supporting line by minimizing the indirect utility function. They applied a grid search method.

step, Vatia's (1983) numerical methodology was employed to approximate the welfare measures. However, their method requires a linear budget set, which impedes its applications to a context with complex taxes and tax reforms. Kumar (2008) proposed an ex post evaluation method using panel data based on Blomquist and Newey's (2002) and Hausman and Newey's (1995) approaches. But, his evaluation requires the choice information before and after a tax reform. It is feasible to extend Hausman and Newey's (1995) welfare evaluation method to deal with general piecewise-linear budget sets (see Appendix A). Unfortunately, the extended method is computationally demanding, particularly when the budget sets are nonconvex. Different from Hausman and Newey (1995) and Kumar (2008), my method of estimation of exact welfare changes is based on the estimated indirect utility function. I treat the generalized indirect utility as a function of the entire budget set. The generalized indirect utility given a piecewise-linear budget set can be efficiently computed using the above global algorithm. As such, it is efficient to numerically compute the compensating variation (CV) and equivalent variation (EV) based on my tool, the generalized indirect utility function.

Finally, I apply my method to study the 1986 tax reform and the 2001 Bush tax cut. I find that estimation is sensitive to the labor supply specification. The labor supply elasticity is only half in a model with a third order approximation (preferred model) than that of the second order approximation. The welfare change and deadweight loss are calculated using the ordinary indirect utility functional and the distribution of unobserved heterogeneity estimated in the structural labor supply. A simulation based method (see Fullerton and Gan, (2004)) is employed to account for the differential welfare effects on heterogeneous individuals.

2.2. An Economic Framework

I consider a static partial equilibrium labor supply model with heterogeneous individuals. Assume that the i-th individual is endowed with a stochastic one-dimension preference ϵ_i. The individual maximizes utility with respect to choices about leisure and other consumption goods g_i (numeraire). Her desired hour of work is denoted to be π_i^*,[4] so $-\pi_i^*$ is her leisure. Let y_i^n denote the after-tax nonlabor income, and w_i^* the latent real gross wage which is exogenously determined in the competitive labor market for the i-th individual. Let $u(\pi_i, g_i, \epsilon_i)$ and $v(w_i^*, y_i^n, \epsilon_i)$ be the random direct utility function and the random indirect utility function respectively. The individual's problem without a labor income tax is as follows:

$$
\begin{aligned}
v(w_i^*, y_i^n, \epsilon_i) = \underset{g_i, \pi_i^*}{max} \;\; & u(\pi_i^*, g_i, \epsilon_i) \\
s.t. \;\; & g_i - w_i^* \pi_i^* = y_i^n, \\
& \pi_i^* \geq 0, g_i \geq 0,
\end{aligned}
\tag{2.1}
$$

where the price of g is normalized to be 1. According to the Roy's identity, the desired hours of working π_i^* can be expressed as

$$
\pi_i^* = \pi(w_i^*, y_i^n, \epsilon_i) = \frac{\partial v(w_i^*, y_i^n, \epsilon_i)/\partial w_i^*}{\partial v(w_i^*, y_i^n, \epsilon_i)/\partial y_i^n}. \tag{2.2}
$$

Typically, the random term ϵ_i enters the desired working hours $\pi(w_i^*, y_i^n, \epsilon_i)$ and indirect utility $v(w_i^*, y_i^n, \epsilon_i)$ in a nonlinear manner, as is known to be important from Brown and Walker (1989). Note that this is a deterministic optimization problem for the particular individual. However, the randomness across individuals make it a random utility model for an econometrician.

---

[4] I use $*$ to emphasize that the variable is known to the individual.

2.2.1. Piecewise-linear Budgets

Under a graduated tax system and various transfer programs, an individual's budget set becomes piecewise linear. It may be convex or nonconvex. Given an individual's before-tax wage rate w_i^*, the after-tax non-labor income y_i^n, and the Tax system T, I can construct the budget set $B_i^* = B(w_i^*, y_i^n, T)$. Let a tax bracket be represented by $\{t_j; Y_{j-1}, Y_j\}$, where t_j is the marginal tax rate for a person whose before-tax income lies within the interval $[Y_{j-1}, Y_j]$. Information about $\{t_j; Y_{j-1}, Y_j\}$ can often be found from tax tables. Note that the relevant budget set is based on after-tax income. Let the end points of the segment in a budget set that corresponds to bracket $\{Y_{j-1}, Y_j\}$ be $\{y_{j-1}^a, y_j^a\}$, where y^a refers to after-tax income. A complete characterization of budget segments requires information on working hours that correspond to the set $[y_{j-1}^a, y_j^a]$, and I denote these hours as $[H_{j-1}, H_j]$. To calculate the location of each budget segment, I start with the first budget segment and proceed through all budget segments. Besides the before-tax wage rate w_i^*, another critical piece of information necessary is Y_i^n, the non-labor income this person may have. Let y_i^n be after-tax non-labor income, where the tax is calculated as if the person had no labor income. Then labor income pushes the person into successively higher tax brackets. I summarize information on budget segments in Table (I).[5] Throughout, the numbering of budget segments is individual-specific.

One interesting observation from Table (I) is that non-labor income affects the location of the budget segments for each individual, since the end points of a budget segment are functions of Y_i^n or y_i^n:

[5]Virtual income is defined as the intercept of the line that extends budget segment j to the zero-hours axis. It is a function of non-labor income and the tax system: $y_{i,j}^v = Y_i^n(2 - t_1 - t_j) - Y_{i,j}(1 - t_j) + \sum_{k=1}^{j}(1 - t_k)(Y_{i,k} - Y_{i,k-1})$. Virtual incomes do not depend on before-tax wage rate. As the before-tax wage rate changes, all the budget segments of one person rotate around their corresponding virtual income points.

Table I.: Summary of Budget Segments

	Budget segment 1	Budget segment $j > 1$
After-tax income	$y_i^a = y_i^n + w_i^*(1 - t_1)h_i^*$	$y_i^a = y_{i,j-1}^a + w_i^*(1 - t_j)(h_i^* - H_{i,j-1})$
Kink points for income	$y_0^a = y_i^n$	$y_{i,j}^a = y_{i,j-1}^a + w_i^*(1 - t_j)(H_{i,j} - H_{i,j-1})$
Kink points for hours	$H_{i,0} = 0$, $H_{i,1} = (Y_{i,1} - Y_i^n)/w_i^*$	$H_{i,j} = (Y_{i,j} - Y_i^n)/w_i^*$
Virtual income	$y_{i,1}^v = y_i^n$	$y_{i,j}^v = y_{i,j-1}^a - w_i^*(1 - t_j)H_{i,j-1}$

$$
\begin{aligned}
H_{i,j} &= (Y_{i,j} - Y_i^n)/w_i^*, \\
y_{i,j}^a &= y_i^n + \sum_{k=2}^{j}(1 - t_k)(Y_{i,k} - Y_{i,k-1}).
\end{aligned} \tag{2.3}
$$

2.2.2. Desired Working Hours under Piecewise-Linear Budgets

It is well known in the literature that a person's locally optimal working hour may be at a kink point or on the interior of a segment, in the framework of piecewise-linear budget constraints. Define

$$
S_{i,j} = \begin{cases} 1 & \text{if the choice is on the interior of segment } j, \\ 0 & \text{otherwise;} \end{cases} \tag{2.4}
$$

$$
K_{i,j} = \begin{cases} 1 & \text{if the choice is at the } j\text{-th kink }, \\ 0 & \text{otherwise.} \end{cases} \tag{2.5}
$$

The conditions determining the values of $S_{i,j}$ and $K_{i,j}$ require knowledge of the basic labor supply function (2.2).The necessary conditions for $S_j = 1$ or $K_j = 1$ are:

$$
\begin{aligned}
&S_{i,j} = 1 \quad \text{if } H_{i,j-1} < \pi(w_{i,j}^*, y_{i,j}^v, \epsilon_i) < H_{i,j}, \\
&K_{i,j} = 1 \quad \text{if } \pi(w_{i,j+1}^*, y_{i,j+1}^v, \epsilon_i) \le H_{i,j} \le \pi(w_{i,j}^*, y_{i,j}^v, \epsilon_i),
\end{aligned} \tag{2.6}
$$

where $w_{i,j}^* = w_i^*(1 - t_j)$.

If the budget set is convex, condition (2.6) is necessary and sufficient and only one of $S_{i,j}$ or $K_{i,j}$ is 1. Thus, the desired working hour for i-th agent is

$$h_i^* = \sum_{j=1}^{J(i)} S_{i,j}\pi(w_{i,j}^*, y_{i,j}^v, \epsilon_i) + \sum_{j=0}^{J(i)} K_{i,j}H_{i,j}, \tag{2.7}$$

where $J(i)$ is the maximum number of budget segments for the i-th individual. However, if the budget set is not globally convex, more than one of $S_{i,j}$ or $K_{i,j}$ may be 1. A global algorithm is required to determine the desired working hour. The optimal working hour should offer the maximum utility among all the segments and kink points. Define the maximum utility among all the segments and kink points as

$$\begin{aligned} u_i^{max} =& max\Big\{ max\{v(w_{i,j}^*, y_{i,j}^v, \epsilon_i) \mid S_{i,j} = 1, j = 1, \ldots, J(i)\}, \\ & max\{u(H_{i,j}, y_{i,j}^a, \epsilon_i) \mid K_{i,j} = 1, j = 0, \ldots, J(i)\}\Big\}, \end{aligned}$$

where $w_{i,j}^* = w_i^*(1 - t_j)$. If only one segment or one kink point attains u_i^{max}, the desired working hour for i-th individual is

$$\begin{aligned} h_i^* =& \sum_{j=1}^{J(i)} S_{i,j} 1\left(v(w_{i,j}^*, y_{i,j}^v, \epsilon_i) = u_i^{max}\right) \pi(w_{i,j}^*, y_{i,j}^v, \epsilon_i) \\ &+ \sum_{j=0}^{J(i)} K_{i,j} 1\left(u(H_{i,j}, y_{i,j}^a, \epsilon_i) = u_i^{max}\right) H_{i,j}. \end{aligned} \tag{2.8}$$

In a rare case, more than one choice attain u_i^{max}. I assume that the agent's desired working hour is the one closest to the observed working hour h_i. [6]

Overall, the analysis under piecewise-linear budgets is reduced to several local optimization problems under linear budgets. Thus, all the textbook theories under linear budgets can be taken full advantages in my framework.

[6] If more than one h_i^* are identically far away from h_i, I randomly designate one of them as the desired choice.

2.2.3. Kink Points and Indirect Utility

The desired working hours depend on the knowledge of the direct utility function $u(\pi_i, g_i, \epsilon_i)$, the indirect utility function $v(w_i^*, y_i^n, \epsilon_i)$, and the labor supply function $\pi(w_i^*, y_i^n, \epsilon_i)$. As the budget set is convex and the preference is strictly convex, the knowledge of the labor supply is adequate to derive the desired working hours, since only one segment or kink point is chosen and the process of utility comparison is not invoked. However, as the budget set is nonconvex, the labor supply function and utility functions are needed. The problem is the complexity involved in the transformation among these three functions. For example, it is simple to derive the labor supply function from the indirect utility function via Roy's identity while the opposite transformation is much more difficult. It involves solving a differential equation whose closed-form solution is hard to find or does not exist at all.

I find that it is convenient to derive the mongrel labor supply from the indirect utility function in the framework of piecewise-linear budgets and strict convex preferences, since the direct utility at a convex kink point is the same as the indirect utility if I can work out the supporting budget line that is tangent with an indifference curve at this kink point. Note that it if of no need to derive the utility at a concave kink point, since it is never a local optimum choice when the preference is strictly convex.

Consider the j-th kink point A, as in Figure (1). The working hour and after-tax income at this kink point are $(H_{i,j}, y_{i,j}^a)$. The indifference curve at this kink point is represented by AC. Note that AC is not tangent with any existing budget segments. Now let a solid line AB, representing a budget line, be tangent to the indifference curve AC at the kink point A. The slope of this new budget line AB is in the interval $[w_{i,j+1}^*, w_{i,j}^*]$[7]. That is, this new budget line AB represents a marginal tax rate that

[7] If $w_{i,j+1}^* > w_{i,j}^*$, the kink is a concave point and therefore it will never be chosen if the

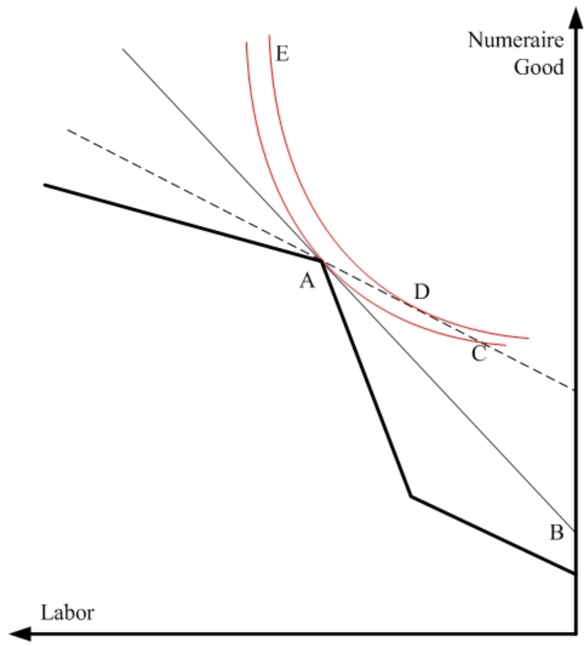

Fig. 1.: The Tangent Line at the Convex Kink Point

is between those of the segment j and segment j+1.

Draw an arbitrary budget line with a slope $\lambda \in [w^*_{i,j+1}, w^*_{i,j}]$ and let it pass through the kink point A, denoted as the dashed line AC in the graph. Its corresponding virtual income is $y^v_\lambda = y^a_{i,j} - H_{i,j}\lambda$. The indifference curve that is tangent with the dashed budget line AC is DE, tangent at point D in the graph. The optimal working hours on this new budget line is $\pi(\lambda, y^v_\lambda, \epsilon_i)$. Its distance with the kink point A is

$$d^2(\lambda) = [\pi(\lambda, y^v_\lambda, \epsilon_i) - H_{i,j}]^2.$$

The purpose is to find the slope of AB. Note that the desired slope (or after-tax wage rate) $\bar{\lambda}$ of the budget line AB is the root such that it satisfies $d^2(\bar{\lambda}) = 0$. Therefore, my problem of finding the slope of AB becomes the problem of obtaining the root of

preference is strictly convex.

the equation $d^2(\bar{\lambda}) = 0$.

Unfortunately, as pointed out in Blundell et. al (1988), the traditional root-finding methods of grid search may encounter the problem of slow convergence. Here I follow a more efficient and effective root-finding binary search method.

Again, consider the solid line AB that tangent to the indifference curve AC at the kink point A. Suppose the dashed line AC has a smaller slope than that of the tangent line. In light of the convexity of the preference, the line AC must intersect the indifference curve to the right of the kink point A. Evidently, the choice falls to the right of A (on the interior of the segment AC) if an individual's budget line were the dashed line AC. Similarly, the choice falls to the left of point A if an individual's budget line had a slope larger than that of the solid line AB. Let us define

$$D(\lambda) = (-1)^{1(\pi(\lambda, y^v_\lambda, \epsilon_i) \leq H_{i,j})} d^2(\lambda), \lambda \in [w^*_{i,j+1}, w^*_{i,j}], \tag{2.9}$$

where $1(\cdot)$ is an indicator function. The function $D(\lambda)$ has a favorable property. If $\lambda < \bar{\lambda}$, I have $D(\lambda) < 0$; If $\lambda > \bar{\lambda}$, I have $D(\lambda) > 0$. The unique root $\bar{\lambda}$ of $D(\lambda)$ can be solved by applying a binary search algorithm as follows:

(1) Initialize an interval $[a, b] = [w^*_{i,j+1}, w^*_{i,j}]$. Thus, the root must be inside the interval.

(2) Divide the interval into two halves $[a, \frac{a+b}{2}]$ and $[\frac{a+b}{2}, b]$. If $D(\frac{a+b}{2}) < 0$, assign a to be $\frac{a+b}{2}$. If $D(\frac{a+b}{2}) > 0$, assign b to be $\frac{a+b}{2}$.

(3) Repeat step (2) till the length of the interval is less than a given precision.

It is well known in the literature that this binary search method converges to the zero point at an exponential rate, much faster than the grid search method. This fast algorithm makes the global algorithm very efficient.

Once the root is available, the direct utility at the j-th kink point is the same as the indirect utility when wage rate is $\bar{\lambda}$:

$$u(H_{i,j}, y^a_{i,j}, \epsilon_i) = v(\bar{\lambda}, y^a_{i,j} - H_{i,j}\bar{\lambda}, \epsilon_i). \tag{2.10}$$

Therefore, the explicit form of the direct utility function is not needed and only the indirect utility function is needed to [8] solve for the labor supply.

The above-mentioned binary search algorithm may be invoked a multiple of times for each observation. The weak law of revealed preference is useful in avoiding unnecessary calculations. If I infer that a segment or a kink point is a candidate for labor choice from a local analysis, and if another kink point is below the supporting line which passes through the segment or the first kink point, I can safely rule out the second kink point from the candidate list.

2.3. The Econometric Model

2.3.1. Specification

My model focuses on the indirect utility function, specified as a polynomial in its three arguments: the net wage w, virtual income y, and heterogeneity in the preference ϵ:

$$v(w, y, \epsilon) = \sum_{r+s+t \in \{0,1,\ldots,K\}} \alpha(r, s, t) w^r y^s \epsilon^t, \tag{2.11}$$

where K is the order of the polynomial and the parameter $\alpha(r, s, t)$ represents the coefficient of the term $w^r y^s \epsilon^t$. As K is allowed to be arbitrarily large, the polynomial can approximate any function of w, y, and ϵ to any degree of accuracy. In this sense, my estimation is nonparametrically flexible. If I fix K, the model reduces to a

[8] In appendix A, I show that only the labor supply function is needed to derive the desired labor supply if Vartia's method is applied. However, it is computationally challenging.

parametric one. In practice however, only small values of K can be used because of the size of the sample. In my applications, a polynomial of order 2 or 3 is enough.

Some restrictions should be imposed on the indirect utility function. First, the ordinal utility function is only identified up to a monotonic transformation. So I impose normalization restrictions: $\alpha(0,0,t) = 0$ and $\alpha(0,1,0) = 1$. Second, the indirect utility function should be monotonically increasing in terms of income. I restrict that $\frac{\partial v}{\partial y} > 0$. This restriction can be imposed through a data-driven method. For example, if I apply the method of simulated annealing to search for the true parameter, the optimization method will avoid the parameters which violate the restrictions. Third, following Blomquist and Newey (2002), I restrict that $\pi(w, y, \epsilon)$ be strictly increasing in ϵ. Thus, I can explain ϵ as the heterogeneity in preference for labor.

Assume that the heterogeneity of preferences ϵ_i is expressed as

$$\epsilon_i = z_i\gamma + \eta_i, \eta_i \sim N(0, \sigma_\eta^2), \tag{2.12}$$

where z_i is a set of demographic characteristics, and $z_i\gamma$ and η_i can be interpreted as observed and unobserved taste heterogeneity in preference. Following the literature, I assume that the unobserved heterogeneity is independent with the covariates, including the demographic characteristics z_i, the gross wage rate w_i^*, and the before-tax non-labor income Y_i^n. Notice that the unobserved heterogeneity and the net wage rate are correlated.

I consider a labor supply function of the form

$$h_i = h_i^* + u_i = h(B_i^*, v, \epsilon_i) + u_i \tag{2.13}$$

where u_i is an optimization error[9] or the measurement error of the observed labor

[9] Hausman (1981b) explained the difference between the observed and desired working hours as a result of unexpected layoffs, short time, overtime, or the worker's poor health

supply. The desired working hour is a function of three arguments: the entire budget set B_i^*, the ordinary indirect utility functional v, and the heterogeneity ϵ_i. Taking the complete budget set as one argument of the labor supply function is in the spirit of Blomquist and Newey (2002). Contrast to Blomquist and Newey (2002), the curse of dimensionality is not a problem in my model, because the nonparametric specification associates with the ordinary indirect utility functional v rather than the (mongrel) labor supply h. Nonconvex budget sets are handled with the aid of the global algorithm discussed in the previous subsection. My specification is convenient for explicitly considering the measurement error in the working hour. Soest, Das and Gong (2002) argued that the measurement error in hours worked can have a detrimental effect on the estimates of the policy effects. Unfortunately, their framework is unable to taken account of the measurement error because it enters into their model in a nonlinear fashion. Following Hausman (1981b), I assume that the optimization/measurement error is zero when the individual is not willing to work, i.e.,

$$P\{u_i = 0 | h_i^* = 0\} = 1. \tag{2.14}$$

Also, I assume that it is normally distributed, i.e.,

$$u_i | (h_i^*, h_i^* > 0) \sim N(0, \sigma_u^2). \tag{2.15}$$

Note that the current framework can simultaneously address several important issues in the literature such as (i) sample selection due to labor nonparticipation (Heckman 1979), (ii) fixed costs of working (Cogan 1981), and (iii) measurement errors of wage rates (Blomquist 1996). I have not done all of them.

---

together with measurement error.

2.3.2. Estimation

If I fix the polynomial order K in (2.11), all the parameters in this model can be estimated by the method of maximum likelihood (ML). Let $X_i = (z_i, w_i^*, Y_i^n)$. The likelihood of working hour when the individual works is

$$
\begin{aligned}
f(h_i|X_i) &= f(h_i|X_i, h_i^* > 0)P(h_i^* > 0|X_i) \\
&= \int_{h_i^*>0} f(h_i|\eta, X_i, h_i^* > 0)dF(\eta|X_i, h_i^* > 0)P(h_i^* > 0|X_i) \\
&= \int_{h_i^*>0} f(h_i|\eta, X_i, h_i^* > 0)dF(\eta|X_i) \\
&= \int_{h_i^*>0} f_u\left[h_i - h_i^*(B_i^*, v, z_i\gamma + \eta)|\eta, X_i, h_i^* > 0\right]dF(\eta|X_i) \\
&= \int_{-\infty}^{+\infty} f_u\left[h_i - h_i^*(B_i^*, v, z_i\gamma + \eta)|\eta, X_i, h_i^* > 0\right]1(h_i^* > 0)dF(\eta|X_i) \quad (2.16)
\end{aligned}
$$

where $f(\cdot)$, $f_u(\cdot)$ and $F(\cdot)$ denote the probability density functions of h_i and u, and the cumulative density function of η, respectively. The first equality in equation (2.16) arises from the fact that $\mathtt{P}\{h_i^* > 0 \mid h_i > 0\} = 1$. Notice that $u_i|(\eta, X_i, h_i^* > 0) \sim N(0, \sigma_u^2)$. Let $\phi(\cdot)$ denote the density function of the standard normal density. Then $f_u[u_i|(\eta, X_i, h_i^* > 0)] = \frac{1}{\sigma_u}\phi(\frac{u_i}{\sigma_u})$. Therefore, I have

$$
f(h_i|X_i) = \int_{-\infty}^{+\infty} \frac{1}{\sigma_u}\phi(\frac{h_i - h_i^*(B_i^*, v, z_i\gamma + \eta)}{\sigma_u})1(h_i^* > 0)dF(\eta|X_i). \quad (2.17)
$$

The indicator function $1(h_{i,j}^* > 0)$ is not differentiable with respect to parameters. I can smooth it by a differentiable kernel function.

$$
k(h^*) = \begin{cases} 0, & h^* \leq 0; \\ \frac{1}{2}\left[1 - cos(\frac{\pi h^*}{\delta})\right], & 0 < h^* < \delta; \\ 1, & h^* >= \delta. \end{cases} \quad (2.18)
$$

I apply the method of simulated maximum likelihood (SML). Draw R error terms from the distribution of η. The corresponding simulated likelihood is

$$\hat{f}(h_i|X_i) = \frac{1}{R}\sum_{j=1}^{R}\frac{1}{\sigma_u}\phi(\frac{h_i - h_i^*(B_i^*, v, z_i\gamma + \eta_j)}{\sigma_u})1(h_{i,j}^* > 0). \tag{2.19}$$

Evidently, the sample mean $\hat{f}(h_i|X_i)$ is an unbiased and consistent estimator of $f(h_i|X_i)$.

The probability that the individual chooses not to work is $p(h_i = 0|X_i) = p(h_i^* \leq 0|X_i) + p(h_i^* > 0, h_i^* + u_i \leq 0|X_i)$. Note that $p(h_i^* > 0, h_i^* + u_i \leq 0|X_i) = E\left[\Phi(\frac{-h_i^*}{\sigma_u})1(h_i^* > 0)|X_i\right]$, where $\Phi(\cdot)$ is the cumulative density function of a standard normal distribution. The corresponding simulated maximum likelihood is

$$\hat{p}(h_i = 0|X_i) = \frac{1}{R}\sum_{j=1}^{R}(1 - 1(h_{i,j}^* > 0)) + \frac{1}{R}\sum_{j=1}^{R}\Phi(\frac{-h_{i,j}^*}{\sigma_u})1(h_{i,j}^* > 0), \tag{2.20}$$

where $h_{i,j}^* = h_i^*(B_i^*, v, z_i\gamma + \eta_j)$.

Combining (2.19) and (2.20) yields the simulated log-likelihood:

$$\mathscr{L} = \sum_{i=1}^{n} 1(h_i > 0)ln[\hat{f}(h_i|X_i)] + 1(h_i = 0)ln[\hat{p}(h_i = 0|X_i)], \tag{2.21}$$

where the unknown parameters are $\{\alpha(r,s,t) \mid r+s+t \in \{0,1,\ldots,K\}\}$, γ, σ_η and σ_u. Note that some parameters are restricted.

2.4. Measure of Welfare Change and Deadweight Loss

2.4.1. Welfare Calculation Based on the Generalized Indirect Utility Function

The utility that the i-th agent attains depends on five factors: the gross wage rate w_i^*, the after-tax nonlabor income y_i^n, the tax system T, the functional of the basic indirect utility v, and the index of the heterogeneity in preferences ϵ_i. Let us introduce $V(w_i^*, y_i^n, T, v, \epsilon_i)$ to denote the generalized indirect utility function. Similar to the labor supply function, the generalized indirect utility function takes the entire budget set as on argument. The first three arguments are all the factors used to generate the budget set. The mapping from the arguments to the generalized indirect utility is seemingly complex. But it is numerically computable and the computation is efficient.

As the tax system changes from $T^0 = \{t_j^0; Y_{j-1}^0, Y_j^0\}$ to $T^1 = \{t_j^1; Y_{j-1}^1, Y_j^1\}$, the welfare change can be measured in terms of either CV or EV. The CV and EV are formally defined as:

$$V(w_i^*, y_i^n(T^0), T^0, v, \epsilon_i) = V(w_i^*, y_i^n(T^1) + CV_i, T^1, v, \epsilon_i) \qquad (2.22)$$

$$V(w_i^*, y_i^n(T^1), T^1, v, \epsilon_i) = V(w_i^*, y_i^n(T^0) - EV_i, T^0, v, \epsilon_i) \qquad (2.23)$$

where $y_i^n(T^i)$ is the after-tax non-labor income under tax system T^i $(i = 0, 1)$. Note that the generalized indirect utility function is monotonically increasing with respect to the second argument. The CV (similarly EV) can be numerically solved by a binary search as follows:

(1) Compute the maximum indirect utility V_i attained under the tax system T^i (i=1,2).

(2) If $V_0 > V_1$, let a be 0 and choose a positive b such that $V(w_i^*, y_i^n(T^1)+b, T^1, v, \epsilon_i) > V_0$; If $V_0 < V_1$, let b be 0 and choose a negative a such that $V(w_i^*, y_i^n(T^1) + a, T^1, v, \epsilon_i) < V_0$. Thus, the CV_i must be inside the interval $[a, b]$.

(3) Divide the interval into two halves $[a, \frac{a+b}{2}]$ and $[\frac{a+b}{2}, b]$. If $V(w_i^*, y_i^n(T^1)+\frac{a+b}{2}, T^1, v, \epsilon_i) > V_0$, assign b to be $\frac{a+b}{2}$. If $V(w_i^*, y_i^n(T^1) + \frac{a+b}{2}, T^1, v, \epsilon_i) < V_0$, assign a to be $\frac{a+b}{2}$.

(4) Repeat step (3) till the length of the interval is less than a given precision.

The binary search algorithm converges to the desired CV_i at an exponential rate. The efficiency of this numerical computation is useful particularly when a simulation-based method is employed.

The sample mean CV (or EV) can be calculated by employing a simulation-based method (see Fullerton and Gan, 2004). Let $CV_i(\epsilon_{i,j})$ (or $EV_i(\epsilon_{i,j})$) denote the CV (or EV) for an individual with a heterogeneity index $\epsilon_{i,j}$. Then the sample mean values of CV and EV are

$$\overline{CV} = \frac{1}{nR} \sum_{i=1}^{n} \sum_{j=1}^{R} CV_i(\epsilon_{i,j}), \tag{2.24}$$

and

$$\overline{EV} = \frac{1}{nR} \sum_{i=1}^{n} \sum_{j=1}^{R} EV_i(\epsilon_{i,j}), \tag{2.25}$$

where n is the number of observations and R the number of random draws.

Based on the measure of CV or EV, I can define the deadweight loss. The tax reform affects the tax revenues from non-labor income and labor income. After a tax change, the utility-maximizing agent's choices of working hours may change from a segment or a kink point to a new segment or a new kink point. If the agent's before-tax-change choice is on the j^0-th segment or $(j^0 - 1)$-th kink point, i.e. $h_i^{*0} \in [H^0_{i,j^0-1}, H^0_{i,j^0})$, and the after-tax-change choice is on the j^1-th segment or $(j^1 - 1)$-th kink point, i.e. $h_i^{*1} \in [H^1_{i,j^1-1}, H^1_{i,j^1})$, the change of tax revenue from labor income ΔR^l can be expressed as

$$\begin{aligned} \Delta R^l = & \left[\sum_{k=1}^{j^1-1} (H^1_{i,k} - H^1_{i,k-1}) w_i^* t_k^1 + (h_i^{*1} - H^1_{i,j^1-1}) w_i^* t_{j^1} \right] \\ & - \left[\sum_{k=1}^{j^0-1} (H^0_{i,k} - H^0_{i,k-1}) w_i^* t_k^0 + (h_i^{*0} - H^0_{i,j^0-1}) w_i^* t_{j^0} \right]. \end{aligned} \tag{2.26}$$

Let ΔR^n denote the changes of tax revenue from non-labor income. Then the deadweight loss calculated based on CV and EV are $CV_i - (\Delta R^l + \Delta R^n)$ and $(\Delta R^l + \Delta R^n) - EV_i$ respectively.

Similarly, the sample mean of deadweight loss is calculated using the simulation-based method. Let $DWL_i(\epsilon_{i,j})$ denote the deadweight loss of an individual with a heterogeneity index $\epsilon_{i,j}$. Then the sample mean of DWL is

$$\overline{DWL} = \frac{1}{nR} \sum_{i=1}^{n} \sum_{j=1}^{R} DWL_i(\epsilon_{i,j}). \tag{2.27}$$

2.4.2. Vartia's Method and Welfare Calculation under Piecewise-Linear Budgets

When the labor supply function is flexibly specified, Vartia (1983)'s method is typically used in the literature (see Hausman and Newey 1995, and Kumar 2008) to estimate the exact welfare changes.

Vartia's method is useful if the budget constraint is linear. Suppose a policy changes an individual's net wage from w^0 to w^1. What is the compensating variation? The key insight of Vartia's method for this problem is moving along the indifference curve by compensating the income in response to the change of the wage. Vartia's (1983) idea is as follows. Let $\pi(w, y)$ and $v(w, y)$ denote the market ordinary labor supply and its indirect utility function respectively. On the same indifference curve, the net wage rate w and the compensated non-labor income y are subject to the following differential equation:

$$\frac{dv(w(t), y(t))}{dt} = \frac{\partial v(w(t), y(t))}{\partial w} \frac{dw(t)}{dt} + \frac{\partial v(w(t), y(t))}{\partial y} \frac{dy(t)}{dt} = 0. \tag{2.28}$$

Combining with the Roy's identity $\pi(w, y) = \frac{\partial v(w,y)/\partial w}{\partial v(w,y)/\partial y}$, Vartia obtained a differential equation

$$\frac{dy(t)}{dt} = -\pi(w(t), y(t)) \frac{dw(t)}{dt} \tag{2.29}$$

with the initial conditions $w(0) = w^0$ and $y(0) = y^0$. The compensated income $y(1)$ at the new net wage rate $w(1) = w^1$ can be solved numerically to any degree of accuracy. A simple numerical method is to recursively apply the following formula:

$$y(t_k) - y(t_{k-1}) = -\pi(w(t_{k-1}), y(t_{k-1}))(w(t_k) - w(t_{k-1})),$$

where $t_k = k/K$ (k=1,2,...,K). Note that the functionals $w(\cdot)$ and $\pi(\cdot,\cdot)$ are known.

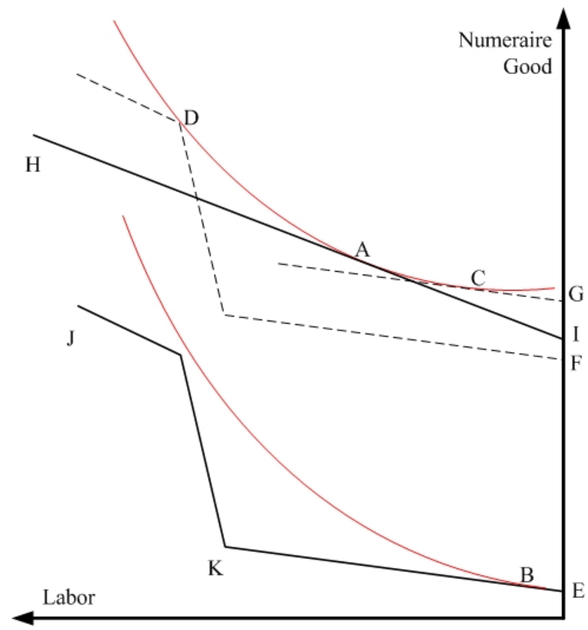

Fig. 2.: The Calculation of CV under Piecewise-linear Budgets

However, in the presence of piecewise-linear budget sets, Vartia's method is not applicable. Consider the calculation of welfare change as shown in Figure (2). Before the tax reform, the budget line is HI and the choice is at point A. After tax reforms, the budget line becomes JKE and the new choice is point B. If the nonlinearity of tax reform is ignored, the calculated CV based on Vartia's method is equal to EG. However, the individual attains the same utility as that at point A by choosing point D on the compensated budget line DF. Obviously, the calculation bias of CV is FG. Therefore, it is a must to modify Vartia's method to account for the nonlinearity unless an alternative methodology is available.

It is complicated to extend Vartia's method to a general context with piecewise-linear budget sets and complex tax reforms. The first and foremost challenge is to determine the touching point between the old indifference curve and the compensated after-tax budget set. Note that the slope of the supporting line which is tangent to the old indifference curve at the above touching point may not be equal to the slope of the supporting line which passes through the post-reform choice point. Thus, Vartia's method is not applicable in a general context. In the appendix A, I extend Vartia's method to handle general piecewise-linear budget sets. Unfortunately, the updated Vartia's method is computationally demanding particularly when the budget is nonconvex. The second challenge is how to determine the desired working hours given a piecewise-linear budget set. If the budget set and the preference are convex, it is easy to calculate the unique choice. However, if the budget set is convex, more than one local optimum choice are predicted. Without closed-form utility functions, Vartia's method has to be applied a multiple times to select the global optimal choice. Evidently, it is computationally demanding to estimate the structure labor supply function.

Contrasting the above two methods, I conclude that, if the budget set is piecewise-linear, working with the indirect utility function is more convenient and efficient than with the labor supply function. This is particularly true if the budget set is nonconvex.

2.5. Application One: The 1986 Tax Reform

The first application of my estimation methodology is the 1986 Tax Reform. I consider the welfare change and the deadweight loss incurred by the modifications of the federal income tax and the EITC program due to the Tax Reform Act of 1986. I use the tax rules in tax year 1987 as the new tax regime. Income taxation in the United States is a complex system. Following Hausman (1981b) and especially Triest (1990), I focus on four types of income taxes in tax year 1983—the federal income tax, state income

Table II.: U.S. Federal Income Tax Schedules in 1983 and 1987

Tax Schedule in 1983		Tax Schedule in 1987	
Income (Dollars)	Marginal Tax Rate	Income (Dollars)	Marginal Tax Rate
\$0-\$2,100	0.11	\$0-\$3,000	0.11
\$2,100-\$4,200	0.13	\$3,000-\$28,000	0.15
\$4,200-\$8,500	0.15	\$28,000-\$45,000	0.28
\$8,500-\$12,600	0.17	\$45,000-\$90,000	0.35
\$12,600-\$16,800	0.19	\$90,000+	0.385
\$16,800-\$21,200	0.23		
\$21,200-\$26,500	0.26		
\$26,500-\$31,800	0.30		
\$31,800-\$42,400	0.35		
\$42,400-\$56,600	0.40		
\$56,600-\$82,200	0.44		
\$82,200-\$105,600	0.48		
\$105,600+	0.50		

tax, payroll tax and EITC program.

2.5.1. The U.S. Income Tax System in 1983 and 1987

Table (II) presents the federal income tax schedules for married couples filing joint returns. In 1983, the federal income tax consisted of 13 brackets.[10] The Tax Reform Act of 1986 greatly simplified the tax code. In 1987, the number of tax brackets decreased to 5. This reform generally reduced the tax burden of an average family, with the maximum marginal tax rates lowering from 50% to 38.5%.

The federal income tax entitled married couples to claim some exemptions and deductions. In 1983, the standard exemption for a married couple was $\$2,000$ plus $\$1,000$ per dependent. If the couple used the standard deduction, an amount of \$3400 was allowed to be subtracted from their taxable income (TI). If they instead itemized the deductions, I follow Triest (1990)'s method to assign an amount equal to

[10] The zero-bracket is family-specific. It does not appear in Table (II).

the average itemized deduction (excluding the state tax payments deduction) within their adjusted gross income class in Individual Income Tax Returns (by Statistics of Income Division, Internal Revenue Service). In 1983, a special deduction was allowed for the married couples when both work. The deduction was the minimum of $\$3,000$ and 10 percent of the earned income of the spouse with the smaller earnings. In 1987, the standard exemption for a married couple and each of their dependents increased to \$3800 and \$1900 respectively. Also, the standard deduction rose to \$3760.

The income tax varied among states.[11] In 1983, ten states imposed no or very limited income tax. Thirty three states imposed a progressive tax which is structurally similar to the federal income tax. Still eight states imposed a flat tax. Among them, Illinois, Michigan and Pennsylvania imposed a flat rate on taxable income, Indiana on AGI, yet Nebraska, Rhode Island and Vermont on the federal income tax liability. The state Massachusetts imposed on non-labor income (interest, dividend and net capital gains) and earned income each a different flat tax rate.

If married couples itemize deductions, the state tax payments can be subtracted from TI when computing federal income tax liability. In sixteen states, federal income tax payments are allowed to be deducted from TI when computing state income tax liability. The one-way and two-way deducting considerably reduced the effective marginal tax rates. Following Triest (1990), I assume that couples who itemized deductions on their federal returns also itemized on their state returns, and claimed the same amount of deductions. The deductions complicate the combination of the federal income tax schedule and the state income tax schedule.

Both the federal income tax and the state income taxes are progressive. Their combination produces a progressive tax schedule and convex budgets for each indi-

[11] The information on the state tax systems is available in the 1982-83 edition of Significant Features of Fiscal Federalism.

Table III.: EITC Tables in 1983 and 1987

EITC in 1983		
Earned income (x)	Stage	Credit
$0-$5,000	phase in	10%*x
$5,000-$6,000	plateau	$500
$6,000-$10,000	phase out	$500-12.5%*(x-6,000)
$10,000+	no credit	$ 0

EITC in 1987		
Earned income (x)	Stage	Credit
$0-$6,080	phase in	14%*x
$6,080-$6,920	plateau	$851.2
$6,920-$15,432	phase out	$851.2-10%*(x-6,920)
$15,432+	no credit	$ 0

vidual. However, the regressive properties in the federal payroll tax and the EITC program create nonconvex regions in the budgets.

In 1983, workers paid 6.7 percent of their earned income (up to $\$35,700$) for the social security tax. The upper limit of the payroll tax will generate a concave kink point in the budgets. Note that the payroll tax is imposed on an individual worker rather than a family. Thus, I calculate the payroll taxes of two spouses separately.

Table (III) presents the structure of the EITC (1983 and 1987) for couples filing jointly with at least one dependent child. The credits were characterized by three stages. In the phase-in stage, married couples(or single heads of households) collected a tax credit equal to a fixed percent of the household earned income (up to an upper limit) from the federal government. In the plateau stage, no more credits can be collected. In the last stage, the tax credit phased out at a fixed rate until it reached zero. The upper limit to the phase-out of EITC creates a concave point in the budgets.

Compared to the EITC in 1983, the EITC in 1987 expanded both intensively and extensively. In 1987, households collected more tax credits than in 1983 from each dollar of earned income in the phase-in stage. Also, the width of the phase-in stage was increased. Moreover, in 1987, the tax credits began to phase out at a larger level of earned income and at a slower rate than in 1983.

2.5.2. Construction of Exact Budget Sets

The budgets used in the literature typically are rough approximates of the actual ones. A common problem in combining tax rules is that the possibility of mutual deductions of payments is not precisely taken account of when the federal tax schedule and the state income tax schedule are combined together. Though the approximation error in the budgets might cause only minor problems in estimating the labor supply elasticities (see Hausman 1981b and Triest 1990), it may severely affect the the estimates of welfare measures. It is impossible to derive an exact budget set for each individual, due to lack of information on detailed household characteristics and the inability to incorporate all the tax systems and welfare programs. However, I can construct the most precise budget sets based on all the information at hand.

Let $T_f = \{t_j^f; Y_{j-1}^f, Y_j^f\}$ and $T_s = \{t_j^s; Y_{j-1}^s, Y_j^s\}$ denote the federal and state income tax schedules respectively. The first tax brackets $\{t_1^f; Y_0, Y_1\}$ and $\{t_1^s; Y_0, Y_1\}$ can be zero brackets which include exemptions and deductions (standard or itemized deductions[12]). Let TI_f and TI_s denote the effective taxable incomes for the federal or state income tax. If sample members indicated that they used the standard deduction to compute the federal tax liabilities, then TI_f is equal to the total income I. Similarly, I have $TI_s = I$ if married couples chose standard deduction or the federal income tax payments were not deductible as they computed the state tax liability. Let TL_f and TL_s denote federal and state tax liabilities. If TL_s were deductible,

[12]Over here, the itemized deduction excludes the federal or state income tax payments.

then $TI_f = I - TL_s$; If TL_f were deductible, then $TI_s = I - TL_f$.

The combined tax schedule of the federal and state tax can be defined recursively. Let $T_{fs} = \{t_j^{fs}; Y_{j-1}^{fs}, Y_j^{fs}\}$ denote the combined tax schedule. Suppose that the $(j-1)$-th bracket has been defined, the j-th bracket will be defined if I know the marginal tax rate t_j^{fs} and the right end of the j-th tax bracket, Y_j^{fs}. I consider three scenarios.

- Case 1) Both TL_f and TL_s are deductible:

  Suppose that the effective taxable incomes of the federal and state income taxes are TI_f and TI_s when the total income I equals Y_{j-1}^{fs}. Assume that the relevant federal and state marginal tax rates are t_f and t_s corresponding to an infinitesimal increase of I, ΔI. Then I have

$$\Delta TL_f = (\Delta I - \Delta TL_s)t_f \tag{2.30}$$

  and

$$\Delta TL_s = (\Delta I - \Delta TL_f)t_s, \tag{2.31}$$

  where ΔTL_f and ΔTL_s are infinitesimal increases of TL_f and TL_s. The solutions of the equations (2.30) and (2.31) are

$$\Delta TL_f = \frac{t_f - t_f t_s}{1 - t_f t_s}\Delta I, \tag{2.32}$$

  and

$$\Delta TL_s = \frac{t_s - t_f t_s}{1 - t_f t_s}\Delta I. \tag{2.33}$$

  Therefore, the marginal tax rate at the j-th bracket of the combined tax schedule, t_j^{fs}, equals $\frac{t_f + f_s - 2t_f t_s}{1 - t_f t_s}$. The corresponding changes of the effective taxable incomes will be

$$\Delta TI_f = \Delta I - \Delta TL_s = \frac{1 - t_s}{1 - t_f t_s}\Delta I \tag{2.34}$$

and

$$\Delta TI_s = \Delta I - \Delta TL_f = \frac{1-t_f}{1-t_f t_s}\Delta I. \tag{2.35}$$

As the tax brackets on which TI_f and TI_s fall do not change, t_f and t_s are constant and the changes of TL_f, TL_s, TI_f, and TI_s are proportional to ΔI. Let D_f denote the maximum change of TI_f such that t_f keeps constant. Similarly, I define D_s. Therefore, I have

$$Y_j^{fs} = Y_{j-1}^{fs} + min\{\frac{1-t_f t_s}{1-t_s}D_f, \frac{1-t_f t_s}{1-t_f}D_s\}. \tag{2.36}$$

Given the above procedure, all the brackets of the combined tax schedule $T_{fs} = \{t_j^{fs}; Y_{j-1}^{fs}, Y_j^{fs}\}$ can be defined one by one. Notice that $TI_f = 0$ and $TI_s = 0$ when $I = 0$.

- Case 2) TL_f is not deductible but TL_s is deductible:
  In this scenario, I have

$$\Delta TL_f = (\Delta I - \Delta TL_s)t_f \tag{2.37}$$

and

$$\Delta TL_s = (\Delta I)t_s. \tag{2.38}$$

The solutions of the equations (2.37) and (2.38) are

$$\Delta TL_f = (1-t_s)t_f \Delta I, \tag{2.39}$$

and

$$\Delta TL_s = t_s \Delta I. \tag{2.40}$$

Thus, the new marginal tax rate t_j^{fs} equals $t_f + t_s - t_f t_s$. The effective taxable incomes are linked to the total income by $\Delta TI_f = (1-t_s)\Delta I$ and $\Delta TI_s = \Delta I$.

Therefore,

$$Y_j^{fs} = Y_{j-1}^{fs} + min\{\frac{1}{1-t_s}D_f, D_s\}, \tag{2.41}$$

where D_f and D_s are defined as in case (1).

- Case 3) TL_s is deductible and the state tax is α percent of federal income tax liability:

  The federal and state tax liabilities can be expressed as

$$\Delta TL_f = t_f \Delta TI_f = (\Delta I - \Delta TL_s)t_f, \tag{2.42}$$

$$\Delta TL_s = \Delta TL_f \alpha. \tag{2.43}$$

  So, $\Delta TL_f = \frac{t_f}{1+t_f\alpha}\Delta I$ and $\Delta T_s = \frac{t_f\alpha}{1+t_f\alpha}\Delta I$. Thus, the marginal tax rate t_j^{fs} equals $\frac{t_f(1+\alpha)}{1+t_f\alpha}$. The new brackets will be the federal tax brackets enlarged by a factor of $1 + t_f\alpha$, because $\Delta TI_f = \frac{1}{1+t_f\alpha}\Delta I$.

2.5.3. The Data Used

The data are drawn from Wave XVII of the PSID. I follow Triest (1990)'s observation selection procedure. Observations from the Survey of Economic Opportunity are excluded. I restrict that married wives and their spouses be aged between 25 and 55 and have average hourly earnings between \$1 and \$50. Those who were disabled and those who reported self-employment or farm incomes are eliminated from the sample. This procedure results in 1004 observations.[13]

The characteristic variables describing a wife's observed heterogeneity in preference for labor include the number of children less than 6 years old, the family size, her

[13] Triest (1990)'s data set has 978 observations. Despite this difference, the summary statistics for my data and Triest's data are very close. One possible explanation is that the new version of PSID has fewer missing values.

school years of education, a dummy for college education (equal to 1 if a woman has completed more than 12 years of education), yearly payments of mortgage, a dummy for her bad health status, her age and an extra age[14] (equal to 0 for women less than 35, equal to age-35 for those between 35 and 45, and equal to 10 for those 45 or over). In the appendix, Heckman's sample selection procedure is provided to impute the latent wage rates of nonparticipating wives.

2.5.4. Estimation Results

2.5.4.1. The Structural Labor Supply

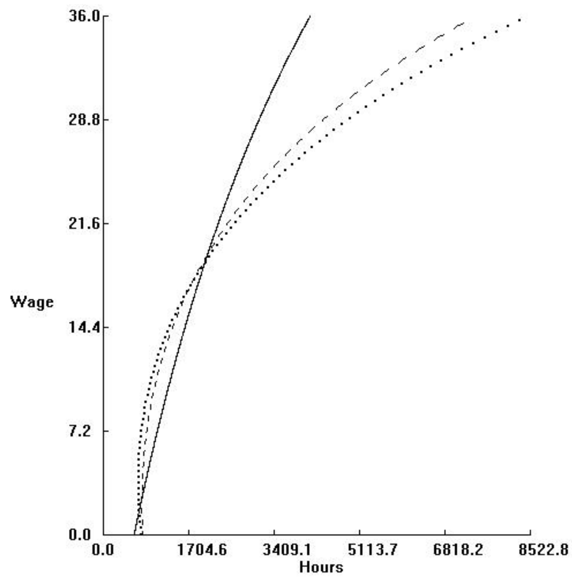

Fig. 3.: Labor Supply Curves

Maximization of the log likelihood functions is performed using the method of simulated annealing.[15] This method statistically guarantees finding a global optimal

[14]It was used by Hausman (1981b) and Triest (1990).

[15]A "C++" software for Windows XP is constructed to compute the problem of maximiz-

solution if it runs in an infinite time. I modify the method of simulated annealing to incorporate the restriction on the indirect utility function. For example, the derivative $\frac{\partial v}{\partial y} > 0$. If the restriction is violated at an iteration, the modified algorithm traces back to the previous iteration and restarts the random search.

I have estimated the model for $K = 2$, 3, and 4. The model with $K = 1$ is not interesting, since its corresponding labor supply is constant. In Figure (3) I present the labor supply curves for a married woman with an average non-labor income and an average heterogeneity in the preference for labor. The solid curve corresponds to the model with $K = 2$. It is very close to a linear labor supply model. However, nonlinearity of the labor supply appears as K increases. The dashed and the dotted curves in the figure correspond to the models with $K = 3$ and 4 respectively.

Likelihood ratio tests suggest that a series expansion of order three is enough to approximate the unknown preferences. The third order is not rejected by the fourth order model at the 5% significance level, but the second order model is rejected. If the Akaike information criterion (AIC) is used, I also arrive at the same conclusion that the model with $k = 3$ is the best. This result is similar to Soest et al (2002) who found that a polynomial series expansion of order two is enough to approximate the direct utility function in their discrete choice model in a study of Dutch labor supply.

Table (IV) presents results of three models with different specifications of the indirect utility functions. I find that only the following eight terms are relevant: w, y, w^2, $w * y$, $w * \epsilon$, w^3, $w^2 * e$, and w^4. Notice that some coefficients of the polynomials are limited to be zeros because of the various restrictions on the indirect utility functions. The coefficients of $w*y$ are associated with the income effects. In all

---

ing likelihood. Traditional matrix platforms such as Gauss and Matlab may be not efficient since my problem involves two levels of optimization. At every iteration of maximizing likelihood, it may be needed to numerically solve the supporting wage rates at the kink points.

three specifications, I identify negative income effects that are statistically significant. Among all the demographic variables, the number of kids (aged less than 6) and the family size significantly affect a women's decision in working hours.

As in most applications of the two error model, the estimated standard deviation of unobserved heterogeneity in preferences is statistically significant. Also, I observe that the standard deviation of optimization/measurement errors has a tendency to decrease as I increase the flexibility of preferences. This is not very surprising since I specify this error term to be orthogonal to the desired working hour.

The wage elasticity is an important variable for policy analysis. Since the labor supply model is nonlinear and the elasticities vary over the sample, it is necessary to define a wage elasticity for the whole population. The elasticities in Table (IV) are computed according to the following formula,[16]

$$\lim_{t \to 1^+} \frac{\sum_{i=1}^n \sum_{j=1}^R h_i^*[B_i^*(t \cdot w_i^*, y_i^n, T), v, \epsilon_j] - \sum_{i=1}^n \sum_{j=1}^R h_i^*[B_i^*(w_i^*, y_i^n, T), v, \epsilon_j]}{(t-1) \sum_{i=1}^n \sum_{j=1}^R h_i^*[B_i^*(w_i^*, y_i^n, T), v, \epsilon_j]}.$$

Put simply, the aggregate elasticity is the percentage change of the total desired working hours if everyone's gross wage rate rises by 1%. This definition takes full account of the tax system and heterogeneity in preferences.

An interesting finding is that the functional specification of preferences has a substantial impact on the estimated elasticities. The estimated wage elasticity decreases from 0.250 to 0.102 as the order of series expansion K rises from 2 to 3. This result suggests that (i) a linear labor supply model is most likely misspecified; and (ii) a misspecified model would generate a substantial bias in labor supply elasticities.

Triest (1990) has estimated wage elasticity using the same data set with a parametrically specified linear labor supply function. His wage elasticity is around 0.25,

[16]This definition is similar to the one used by Soest, Das, and Gong (2002).

Table IV.: Estimates of Labor Supply with Flexible Preferences (1983)

	2nd Order	3rd Order	4th Order
Heterogeneity in Preferences			
σ_η	647.590	745.108	737.638
	(40.021)	(7.716)	(53.079)
constant	909.183	1091.905	1074.850
	(NA)	(25.333)	(31.854)
Number of kids (Age$\leq$ 6)	-331.650	-370.177	-367.529
	(22.429)	(16.376)	(15.181)
Family Size	-128.776	-116.139	-118.335
	(7.960)	(33.213)	(5.696)
Age (35-45)	11.081	3.791	8.488
	(17.788)	(16.601)	(14.306)
Age-45	-22.756	-15.170	-19.181
	(13.381)	(12.879)	(11.300)
Education	18.846	27.870	29.262
	(18.336)	(8.120)	(8.297)
College education	108.910	104.337	102.393
	(92.451)	(84.970)	(78.687)
Log of yearly mortgage payment	10.892	14.609	17.716
	(9.828)	(10.452)	(10.104)
Bad health	-200.262	-217.705	-231.633
	(157.092)	(150.678)	(142.906)
Indirect Utility Function			
w	64.112	77.676	83.689
	(218.173)	(171.781)	(144.371)
w^2	26.356	-14.106	-26.856
	(8.891)	(18.279)	(3.322)
$w * y$	-1.087e-2	-1.166e-2	-1.203e-2
	(3.672e-3)	(6.044e-4)	(3.062e-3)
w^3		1.126	1.639
		(0.721)	(1.055)
$w^2 * \epsilon$		1.969e-3	3.656e-3
		(1.089e-2)	(1.036e-2)
w^4			-3.703e-3
			(3.903e-2)
σ_u	766.124	634.510	616.327
	(73.068)	(103.882)	(71.573)
Uncompensated wage elasticity:	0.250	0.102	0.112
Log likelihood=	-6331.605	-6325.079	-6324.535

Note: (1) The coefficient of y and $w * \epsilon$ are set to be 1;
(2) The number of simulated repetitions R=2000.

Table V.: Estimates of Welfare Changes and Deadweight Loss (K=3;1986)

	Federal income tax	EITC	EITC (Convexified)	Federal income tax+EITC
Average working hours	977.046	977.046	977.325	977.046
before the reform	(48.526)	(50.304)	(47.889)	(57.908)
Average working hours	986.547	976.333	976.568	985.805
after the reform	(55.570)	(50.199)	(47.793)	(63.151)
Labor supply change as % of	0.972	-0.073	-0.078	0.896
working hours before the reform	(1.285)	(0.046)	(0.039)	(1.243)
Average tax revenue	8573.588	8573.588	8575.768	8573.588
before the reform	(191.543)	(181.887)	(191.049)	(213.824)
Average tax revenue	6955.484	8539.245	8541.370	6921.327
after the reform	(187.813)	(182.371)	(191.264)	(207.827)
Tax revenue change as % of	-18.873	-0.401	-0.401	-19.272
old tax revenue	(0.517)	(0.020)	(0.016)	(0.572)
Average CV	-1708.989	-32.806	-35.788	-1741.670
	(40.284)	(1.987)	(3.238)	(41.131)
CV as % of	-19.933	-0.383	-0.417	-20.314
old tax revenue	(0.102)	(0.027)	(0.043)	(0.162)
Average DWL	-90.885	1.537	-1.390	-89.409
	(48.348)	(2.432)	(3.062)	(49.268)
DWL as % of	-1.060	0.018	-0.016	-1.043
old tax revenue	(0.551)	(0.028)	(0.035)	(0.564)
Average AGI	37066.446	37066.446	37069.725	37066.446
before the reform	(454.493)	(438.676)	(449.903)	(510.223)
Average AGI	37275.170	37062.026	37064.967	37270.557
after the reform	(546.279)	(437.779)	(449.000)	(600.068)

Note: (1) Standard errors are in parentheses.
(2) $CV < 0$ means a gain.
(3) Working hours are evaluated for a married woman.
(4) Tax revenues and AGI are evaluated for a family.
(5) The number of simulated repetitions R=2000.

almost same as the wage elasticity calculated here when using second order polynomial. A more flexible third and fourth order polynomials generate much smaller wage elasticity. This is consistent with Blomquist and Newey (2002) who found that non-parametric estimation of wage elasticity is 60% of one based on a linear parametric method.

2.5.4.2. Welfare Change and Deadweight Loss

The welfare effects are obtained by doing simulations based on the structural labor supply function and the generalized indirect utility function. Standard errors are computed by bootstrapping. I repeat the simulations many (100) times with new parameters drawn from their estimated distributions. Standard errors are calculated from the bootstrapped sample.

Table (V) presents the estimated welfare effects due to the change of the federal income tax and/or the expansion of EITC program (The order of series expansion K is equal to 3). My framework enables us to study these reforms separately or as a whole. The reform of the federal income tax significantly reduced the federal government's tax revenue but made working families much better off. An average couple would pay 20% of their before-tax-reform tax revenue to be just as well off after the reform as they were before the reform. The reform of the federal income tax in 1986 associated with a reduction in DWL because it encouraged married women to work or work more hours. However, the changes of working hours are not statistically significant, which is consistent with the estimated relatively low labor elasticity. The expansion of EITC in 1986 also greatly benefited the families with dependent children. However, it produced DWL by discouraging married wives to work if their family income was near the phase-out stage of the EITC before its expansion. The slight decrease of working hours indicates that the overall income effects dominate the overall substitution effects.

Another interesting finding is that convexifying the budgets may bias the estimate of CV downward. In this application, the sign of the DWL of the expansion of EITC reverses if I use convex hulls in lieu of the original nonconvex budget sets. Evidently, an accurate construction of budget sets matters in evaluating welfare effects.

This application shows that a correct specification of preferences also matters in estimating welfare effects. Table (VI) presents the estimated welfare effects from the restricted model ($K = 2$). Contrast to the results in Table (V), the magnitude of the labor supply change is much larger and statistically more significant. The change of federal income tax increases the average married woman's working hour by about 2.81% of that before the reform in the restricted model, while only 0.97% in the more flexible model. The expansion of EITC in 1986 decreases the her working hour by about 0.16% in the restricted model, while only 0.07% in the more flexible model. This difference in the responsiveness of labor supply can be explained by the difference in estimated labor elasticity from two models. The change of the federal income tax in 1986 actually decreases the tax rate and thus increases a married woman's net wage rate. A model with a larger estimated wage elasticity will produce a larger labor supply increases corresponding to the tax reform. The expansion of EITC in 1986 actually decreases the net wage rate just beyond the phase-out stage of the old EITC. A model with larger estimated wage elasticity may produce a larger decreases in labor supply. More importantly, the CVs estimated from two models differ slightly, while the estimated DWLs see a considerable difference. Overall, the restricted model overstates the effects of the reforms.

Table VI.: Estimates of Welfare Changes and Deadweight Loss (K=2;1986)

	Federal income tax	EITC	EITC (Convexified)	Federal income tax+EITC
Average working hours	963.767	963.767	964.321	963.767
before the reform	(33.521)	(34.283)	(37.919)	(31.515)
Average working hours	990.874	962.242	962.850	989.347
after the reform	(34.797)	(34.275)	(37.846)	(34.746)
Labor supply change as % of	2.813	-0.158	-0.153	2.654
working hours before the reform	(0.831)	(0.029)	(0.031)	(0.769)
Average tax revenue	8585.954	8585.954	8589.656	8585.954
before the reform	(123.418)	(128.022)	(138.526)	(124.204)
Average tax revenue	6991.278	8549.911	8553.737	6955.794
after the reform	(115.285)	(128.046)	(138.714)	(119.864)
Tax revenue change as % of	-18.573	-0.420	-0.418	-18.986
old tax revenue	(0.332)	(0.013)	(0.014)	(0.326)
Average CV	-1717.112	-32.977	-36.098	-1749.787
	(26.697)	(0.825)	(0.845)	(27.652)
CV as % of	-19.999	-0.384	-0.420	-20.380
old tax revenue	(0.060)	(0.012)	(0.015)	(0.060)
Average DWL	-122.436	3.128	-0.179	-119.627
	(33.290)	(0.539)	(0.605)	(33.656)
DWL as % of	-1.426	0.036	-0.002	-1.393
old tax revenue	(0.377)	(0.006)	(0.007)	(0.376)
Average AGI	37158.192	37158.192	37164.410	37158.192
before the reform	(282.532)	(293.290)	(327.539)	(284.932)
Average AGI	37468.000	37149.489	37155.739	37459.103
after the reform	(323.485)	(292.823)	(326.809)	(342.463)

Note: (1) Standard errors are in parentheses.
(2) $CV < 0$ means a gain.
(3) Working hours are evaluated for a married woman.
(4) Tax revenues and AGI are evaluated for a family.
(5) The number of simulated repetitions R=2000.

Table VII.: The Bush Tax Cut of 2001

Old tax regime		New tax regime	
Income (Dollars)	Marginal Tax Rate	Income (Dollars)	Marginal Tax Rate
\$0-\$43,850	0.15	\$0-\$6,000	0.10
\$43,850-\$105,950	0.28	\$6,000-\$43,850	0.15
\$105,950-\$161,450	0.31	\$43,850-\$105,950	0.25
\$161,450-\$288,350	0.36	\$105,950-\$161,450	0.28
\$288,350+	0.396	\$161,450-\$288,350	0.33
		\$288,350+	0.35

2.6. Application Two: The 2001 Bush Tax Cut

In the second application of my estimation methodology, I consider the welfare change and the deadweight loss for a married woman due to the Economic Growth and Tax Reconciliation Act of 2001 (the Bush tax cut). Table (VII) presents the tax regimes before and after the Bush cut. I consider only the changes in marginal tax rates. Since the changes are phased in, I use the rates after 2006 when all changes are fully implemented.

Similar to the application in section (2.5), the budget set of a married women is constructed by considering four types of income taxes—the federal and state income taxes, payroll tax, and the EITC program. Compared to the year 1983, the income taxes changed significantly in the year 2000. For a detailed description of the state income tax, see the 2001 state tax handbook. In 2006, workers paid 6.2% of their earned income (up to \$76,200) for the social security tax. All earnings were subject to Medicare's Hospital Insurance tax (1.45%). Table (VIII) shows the EITC parameters in 2000. A family without a child was also eligible for EITC if the family income is very low. The tax credits for a family with one or more children were quite generous.

I use the data set from the PSID of 2001. The data pertain to the year 2000. The

Table VIII.: The EITC Table in 2000

No child		
Earned income (x)	Stage	Credit
\$0-\$4,610	phase in	7.65%*x
\$4,610-\$5,770	plateau	\$352.665
\$5,770-\$10,380	phase out	\$352.665-7.65%*(x-5,770)
\$10,380+	no credit	\$ 0

One child		
Earned income (x)	Stage	Credit
\$0-\$6,920	phase in	34%*x
\$6,920-\$12,690	plateau	\$2352.8
\$12,690-\$27413	phase out	\$2352.8-15.98%*(x-12,690)
\$27,413+	no credit	\$ 0

More than one child		
Earned income (x)	Stage	Credit
\$0-\$9,720	phase in	40%*x
\$9,720-\$12,690	plateau	\$3888
\$12,690-\$31,152	phase out	\$3888-21.06%*(x-12,690)
\$31,152+	no credit	\$ 0

criterion for data exclusion is similar to the above application except that I restrict the hourly earning for either spouse in the range between \$1 and \$70. This leaves us with 1,166 observations. The demographic variables are chosen almost the same as in the above application except that the number of children less than 6 years old is replaced with the age of the youngest child[17]. In addition, the latent wage rates of nonparticipating wives are imputed using Heckman's sample selection procedure (see Appendix B).

The parameter estimates of the structural labor supply are listed in Table (IX). Both the likelihood ratio test and the AIC model selection support the model with $K = 2$. One interesting aspect of the estimates is the small standard deviation of the heterogeneity of preference for labor.

The estimates of welfare effects are listed in Table (X). The model with $K = 2$ and the model with $K = 3$ offer similar estimates of welfare effects. The Bush tax cut reduced the government revenue but encouraged married women to work more hours (not statistically significant). The policy benefited an average family and reduced the deadweight loss.

2.7. Summary

In this essay, I estimate labor supply and exact welfare change under nonparametrically specified preference and nonconvex piecewise-linear budgets. Different from Blomquist and Newey (2002), my estimation method can handle nonconvex budget sets; different from Soest et al, my method can handle measurement errors in working hours. My estimation starts from a indirect utility function from which the derivation of uncompensated labor supply involves a differentiation rather than an integration.

[17]If no child is in a family, the value of this variable is assigned 18.

Table IX.: Estimates of Labor Supply with Flexible Preferences (2000)

	2nd Order	3rd Order	4th Order
σ_η	97.609	80.975	82.351
	(828.415)	(829.424)	(905.074)
constant	22.905	32.379	31.891
Age of the youngest child	66.714	67.153	67.135
	(10.111)	(9.258)	(20.614)
Education	44.539	47.281	52.636
	(24.763)	(11.631)	(26.251)
Log of yearly mortgage payment	89.838	89.949	90.562
	(17.355)	(16.879)	(29.848)
Bad health	-1034.941	-1029.143	-1030.921
	(69.656)	(73.791)	(345.952)
Indirect Utility Function			
w	17.118	17.563	27.695
	(325.583)	(191.432)	(663.198)
w^2	12.969	9.021	4.401
	(6.695)	(5.297)	(62.008)
y^2	5.228e-6	5.228e-6	5.250e-6
	(1.998e-6)	(2.667e-6)	(1.720e-6)
$w * y$	5.132e-4	-2.900e-3	-2.018e-3
	(3.424e-3)	(6.899e-3)	(8.311e-3)
w^3		6.242e-2	9.424e-2
		(0.298)	(2.413)
$w^2 * y$		1.025e-4	5.847e-5
		(1.742e-4)	(6.363e-4)
$w^2 * e$		6.213e-6	1.371e-5
		(7.311e-3)	(2.163e-2)
$w * y^2$		1.907e-8	6.954e-9
		(2.606e-8)	(1.912e-8)
w^4			-3.828e-4
			(2.576e-2)
w^3y			1.379e-6
			(1.575e-5)
w^3e			3.040e-6
			(5.691e-4)
σ_u	912.817	913.057	908.181
	(47.449)	(41.639)	(45.501)
Uncompensated wage elasticity:	0.105	0.127	0.113
Log likelihood=	-8313.233	-8311.479	-8311.250

Note: (1) For identification, The coefficients of y and $w * \epsilon$ are restricted to be 1;
(2) The constant in equation (2.12) and the coefficient of w are not identified simultaneously.

Table X.: Welfare Effects of the 2001 Bush Tax Cut (K=2 and K=3)

	2nd order	3rd order
Average working hours	1383.422	1387.154
before the reform	(30.850)	(46.185)
Average working hours	1386.484	1392.087
after the reform	(30.772)	(47.022)
Labor supply change as % of	0.222	0.354
working hours before the reform	(0.158)	(0.201)
Average tax revenue	20879.581	21010.172
before the reform	(217.583)	(368.588)
Average tax revenue	19915.736	20063.669
after the reform	(209.158)	(356.817)
Tax revenue change as % of	-4.616	-4.505
old tax revenue	(0.089)	(0.117)
Average CV	-1003.940	-1014.139
	(17.809)	(26.544)
CV as % of	-4.808	-4.826
old tax revenue	(0.043)	(0.049)
Average DWL	-40.094	-67.636
	(19.828)	(27.335)
DWL as % of	-0.192	-0.321
old tax revenue	(0.094)	(0.128)
Average AGI	84253.832	84564.388
before the reform	(531.068)	(896.025)
Average AGI	84349.025	84728.338
after the reform	(552.023)	(935.728)

Note: (1) Standard errors are in parentheses.
(2) $CV < 0$ means a gain.
(3) Working hours are evaluated for a married woman.
(4) Tax revenues and AGI are evaluated for a family.
(5) The number of simulated repetitions R=2000.

This facilitates building the flexible preferences into a fully structural model, and also welfare calculation.

I apply my method to study the 1986 tax reform and 2001 Bush tax cut. I estimate the labor supply model using PSID data of 1984 and 2001. My estimates show that a correct specification of preferences and an accurate approximation of the budget sets are vital to the estimation of exact welfare effects, particularly of the DWL. Convexifying the budgets sometimes results in misleading conclusions in the welfare evaluation particularly when the relevant policies are regressive. As such, incorrectly specified preferences may lead to an understate or overstate of the size of DWL.

CHAPTER III

A RESOLUTION TO EQUITY PREMIUM, RISK-FREE RATE, AND CAPITAL STRUCTURE PUZZLES

3.1. Introduction

In this essay, I address the equity premium puzzle, risk-free rate puzzle, and capital structure puzzle by examining a large, overlapping-generations and consumption-based asset pricing model. One novelty is that it explains the asset pricing puzzles using hump-shaped life-cycle consumption. I show that the equity premium puzzle arises primarily from the aggregation of hump-shaped life-cycle consumption. The stochastic discount factor constructed with per capita consumption and utility functions of micro-level agents is typically not orthogonal to the market excess return rate. To force the orthogonality condition of a representative agent model to hold generally involves an artificial specification of preferences which do not reflect the behaviors and risk attitudes of realistic micro-level agents. The risk-free rate is much smaller than that in a representative agent model because a portion of individuals in my model save for their future lives while all the individuals in a representative agent model typically prefers to borrow in a growing economy. Another novelty is that the capital structure is endogenously determined by investors hedging against risks. The endogenously determined leverage generates an even larger equity premium than a fixed leverage.

The hump-shaped life-cycle consumption exhibits a good performance in resolving the asset pricing puzzles. I maintain the traditional rational expectation framework and adopt the simple constant relative risk averse (CRRA) utility function. The baseline model with a degree of risk aversion 1.2 generates an equity premium as high as 4% and a risk-free rate as low as 0.7%. Also, the model performs very well in explaining the formation of capital structure. In the baseline model, the calibrated debt-to-equity ratio is around 0.5. Equivalently, the calibrated debt-to-capital ratio is around 0.33. They are consistent with the empirical results.

The best known empirical failures of consumption-based asset pricing models[1] are the equity premium puzzle and the risk-free rate puzzle. Mehra and Prescott (1985) investigate U.S. data from 1889 to 1978 and find that the mean annual premium of equity return over the riskless rate was around 6% which is too large to be justified by the standard Arrow-Debreu model with a plausible degree of risk aversion. In other words, stocks are not sufficiently riskier than Treasury bills to explain the spread in their returns. This is the well-known equity premium puzzle. In the same paper, Mehra and Prescott point out that the real interest rate of 0.8% is too low to be explicable in their model. Weil (1989) terms the low interest rate as a puzzle because it cannot be justified by his representative agent model with a plausible degree of risk aversion and an arbitrary level of inter-temporal elasticity of substitution. These asset pricing puzzles represent large gaps in our understanding of macroeconomy (see Kockerlakota 1996).

In the past two decades, numerous ideas were presented to generate a high equity premium or low risk-free rate. However, more and more economists agree that it is not sufficient for resolving the asset pricing puzzles to build a model capable of generating reasonable asset return rate. A convincing solution must explain why the Euler equations are rejected in an Econometric test. A large volume of literature, including Hansen and Singleton (1982), Grossman, Melino, and Shiller (1987), Hansen and Jagannathan (1991), Ferson and Constantinides (1991), and Kockerlakota (1996)

[1]See Merton (1971, 1973), Rubinstein (1976), Lucas (1978), Breeden (1979), and Cox, Ingersoll, and Ross (1985).

report the rejections.

Euler equations are associated with four classes of factors including consumption paths, preferences, probability distributions of asset returns, and other elements such as market incompleteness and imperfections. I argue that the consumption path is responsible for the rejections of Euler equations. Most models in the literature are based on a representative agent model with per capita consumption. However, from the literature of life-cycle studies, a more realistic, micro-level consumer's consumption curve is hump-shaped, which is in sharp contrast to the stochastically growing consumption curve of an individual who infinitely lives in a representative agent model. Since the consumption growth is the basic element of pricing kernels, the pricing equations derived from a representative agent model with per capita consumption may not reflect a realistic consumer's investment behavior and risk attitudes.

The existing literature has intensively explored virtually all the factors except for consumption paths. However, puzzles are still unresolved. One line of work in the literature specifies alternative preferences.[2] Despite the fact that it produces fruitful results, altering utility functions is always under suspicion. More importantly, the new preferences designed for a representative agent may not reflect the behaviors and risk attitudes of realistic micro-level agents. Further, this line of research by its nature is unable to answer why the simple CRRA utility function and rational expectation theories fail to explain the asset returns but work pretty well in other areas such as real business cycle. Another line of literature modifies the probability distributions of asset returns. Reitz (1988), Barro (2006) and Barro (2009) argue that

[2]Typically, researchers adopt utility with habit persistence or non-expected utility. For habit persistence, see Abel (1990), Boldrin, christiano and Fisher (2001), Campbell and Cochrane (1999), Constantinides (1990) and Ferson and Constantinides (1991). For non-expected utility theories, see Benartzi and Thaler (1995), Epstein and Zin (1989, 1991) and Weil (1989).

rare disasters explain a large portion of the equity premium. However, the calibrated degree of risk aversion is still too large to be true. Still a third line of literature tries to use market incompleteness and/or market imperfections[3] to explain asset pricing puzzles. Unfortunately, it generally lacks success in explaining the equity premium puzzle (see Kockerlakota 1996). The recent work by Constantinides, Donaldson, and Mehra (2002) and Storesletten, Telmer, and Yaron (2007) find that life-cycle features are useful to generate a high equity premium.[4] But their focus still is on how to generate reasonable asset return rate, but not on how to address the rejections of Euler equations.

Though the choice of consumption path accounts for the key problems of asset pricing models, it does not imply a high equity premium. In my model, the high equity premium is related to an endogenously determined capital structure. It is well known that a fixed leverage is able to generate a large equity premium. The literature includes Abel (1999), Barro (2006), Benninga and Protopapadakis (1990), Chan and Kogan (2002), Constantinides (1990), Constantinides, Donaldson, and Mehra (2002), Ebrahim and Mathur (2001), Jermann (1998), etc. An endogenous leverage is able to produce an even larger equity premium. In my model, the firm borrows more as the expected system growth rate is high or the interest rate is low; it borrows less as the expected growth rate is low or the interest rate is high. Thus, the total capital return

[3]For example, Aiyagari and Gertler (1991), Alvarez and Jermann (2000), Attanasio, anks, and Tanner (2002), Bansal and Coleman (1996), Basak and Cuoco (1998), Brav, Constantinides and Geczy (2002), Cogley (2002), Constantinides and Duffie (1996), Danthine, Donaldson, and Mehra (1992), Detemple and Serrat (2003), He and Modest (1995), Heaton and Lucas (1996, 1997, 2000), Jacobs (1999), Krusell and Smith (1998), Lucas (1994), Luttmer (1996), Mankiw and Zeldes (1991), Mankiw (1986), Marcet and Singleton (1999), McGrattan and Prescott (2000,2001), Telmer (1993), Vissing-Jorgensen (2002), and Zhou (1999).

[4]They employ a 2 or 3-period overlapping generations model to incorporate life-cycle features. A common characteristic of these models is that aggregate risks concentrate on older agents so that the older generation demands a high equity premium.

rate and the stock return rate is larger than those acquired with a fixed leverage.

In the literature, the determination of capital structure is puzzling. Myers (1984) terms this as the capital structure puzzle[5]. Modigliani and Miller (1958) argue that the capital structure is irrelevant for a firm to maximize its value. Though this theoretical result is widely accepted, it is not helpful for our understanding of asset pricing. A good knowledge of the determination of capital structure is crucial for us to address the equity premium puzzle because the choice of debt level by the firm affects the dividend flow of its stocks and consequently stock returns.

Different from the literature, the capital structure in my model is determined by investors rather than the management of firms. Economists typically assume that the management of a firm determines the capital structure by optimizing a single objective function, say, the firm's value. However, from the perspective of corporate governance, the management is appointed or dismissed by the owners. The decision of the management is subject to the owners. A single objective, whatever it is, may never be agreed to by heterogenous owners whose interests are typically in conflict. For example, a young investor may weight more of future development of the firm and prefer the management to borrow more and distribute less dividends. However, an aged investor may do the opposite. As a matter of fact, financing decision is an optimization problem with multiple objectives. In my model, the owners of a firm are agents of overlapping generations who maximize their own life-time utility. I find that the free financial market is a natural way to reconciling the multiple and conflicting objectives. In the equilibrium, the total bond held by the consumers are equal to the debt of the firm. In sum, I study the determination of capital structure in the framework of asset pricing and the capital structure is essentially determined by the

[5]Three popular theories about the determination of capital structure are the trade-off theory, pecking order theory, and the theory of agency costs.

investors.

3.2. Motivation

The key problems of asset pricing models are the unanimous rejections of Euler equations. In this section, I offer a motivation for the usefulness of hump-shaped life-cycle consumption curve in correcting for the rejections.

Kocherlakota (1996) illustrates that the equity premium puzzle and risk-free rate puzzle are awfully robust by testing the Euler equations:

$$E_t\left[\beta\frac{c_{t+1}^{-\gamma}}{c_t^{-\gamma}}\left(\frac{P_{t+1}+D_{t+1}}{P_t}-R_t^f\right)\right]=0, \tag{3.1}$$

and

$$E_t\left[\beta\frac{c_{t+1}^{-\gamma}}{c_t^{-\gamma}}R_t^f\right]=1, \tag{3.2}$$

where γ is the degree of risk aversion in a CRRA utility function, β the constant discount factor, and P_t, D_t, R_t^f, and c_t respectively denoting the price of risky assets, dividend, gross risk-free rate and consumption at time t. Econometric tests indicate that no appropriate γ makes the above two Euler equations hold at the same time.[6] The following simple example may reveal that the aggregation of life-cycle consumption across different generations accounts for the rejections of Euler equation tests. Suppose agents in an economy live for three period. They are endowed with CRRA preferences. Their degree of risk aversion is γ. Let $M_1=\beta_1\frac{c_{2,t+1}^{-\gamma}}{c_{1,t}^{-\gamma}}$ and $M_2=\beta_2\frac{c_{3,t+1}^{-\gamma}}{c_{2,t}^{-\gamma}}$ be the stochastic discount factors for the first and second generations, where $\beta_1>1$, $\beta_2<1$, and $c_{i,t}$ denotes the consumption of the i-th generation at time t. In their first period, they value more of consumption in the next period than current con-

[6]A large γ may make the first Euler equation hold, but the second one will be rejected; a small γ may make the second Euler equation hold, but fails the first equation.

sumption; In their second period, they value more of current consumption than that in the next period. If β_1 and β_2 are far away from 1, I have that $E_t(c_{2,t+1}) > c_{1,t}$ and $E_t(c_{3,t+1}) < c_{2,t}$. That is, the life-cycle consumption curve is hump-shaped. At the equilibrium, two stochastic discount factors M_1 and M_2 are orthogonal to the excess return rate $\left(\frac{P_{t+1}+D_{t+1}}{P_t} - R_t^f\right)$. Let $\bar{c}_t = (c_{1,t} + c_{2,t} + c_{3,t})/3$ be the per capita consumption which is apparently much smoother than the life-cycle consumption. The term $\frac{\bar{c}_{t+1}^{-\gamma}}{\bar{c}_t^{-\gamma}}$ is generally not orthogonal to the excess return rate. It might be possible to construct a stochastic discount factor orthogonal to the excess return rate based on the per capital consumption. But, the degree of risk aversion may not be γ or the utility function for the aggregate agent will not be of CRRA class. Therefore, a representative agent model may never correctly identify the behaviors of micro-level agents with a realistic life-cycle consumption.

3.3. The Model

I construct an overlapping generations model with a pure exchange economy. The economy is similar to those studied by Lucas (1978) and Mehra and Prescott (1985). Two distinctions are (1) that consumers are equipped with micro-level life-cycle features, and (2) that I introduce an income distributor — the firm.

3.3.1. Consumers

Consider an economy with overlapping generations of agents who live a maximum of $I = 75$ periods, with ages denoted by $i \in \mathscr{I} \equiv \{1, \ldots, I\}$. Agents can die earlier. The probability of surviving between age i and $i+1$ is denoted by s_i, with $s_I = 0$. The unconditional probability of being alive at age i is $s^{(i)} = \prod_{j=1}^{i-1} s_j$, with $s^1 = 1$. Assume that the measure of newly-born agents is fixed. The size of the population and the sizes of different cohorts do not change over time. I normalize the population size to 1. Then the size of agents at age i is $\mu_i = \frac{s^{(i)}}{\sum_{i=1}^{I} s^{(i)}}$.

An agent born at time t wishes to maximize the lifetime expected utility,

$$E_t \Big\{ \sum_{i \in \mathscr{I}} \beta^{i-1} s^{(i)} U(c_{i,t+i-1}) \Big\}, \quad \beta > 1, \tag{3.3}$$

where c_t is a stochastic process of consumption of a perishable consumption good, β is a discount factor, $U(\cdot)$ is an instantaneous utility function, and $E\{\cdot\}$ is an expectation operator. The instantaneous utility function is restricted to be of the constant relative risk aversion class,

$$U(c) = \frac{c^{1-\gamma} - 1}{1 - \gamma}, \tag{3.4}$$

where γ measures the degree of relative risk aversion.

This lifetime utility function has been studied by José-Víctor Ríos-Rull (1996). Both β and s_i's govern the shape of the life-cycle consumption curve. I restrict β to be larger than 1, which implies that agents value future consumption more than today's consumption. When agents are young, s_i's are close to 1. The growth in consumption is determined primarily by β. This growing consumption when agents are young is consistent with the life-cycle literature. When agents are old, the survival rates s_i's decrease towards 0, and the values of the sequence $\beta^{i-1} s^{(i)}$ reduce to a number less than 1. Therefore, old agents value future consumption less, and consumption is decreasing with age.

The first generation at each period follows a social convention to chooses the consumption level. The convention is formulated as

$$c_{1,t} = \rho c_{20,t}, 0 < \rho < 1. \tag{3.5}$$

In other words, the first generation pegs to the consumption of the twentieth generation. This is similar to the idea of habit formation. But it is a social habit rather than an individual habit.

Agents are endowed with one unit of labor. Let ϵ_i denote an age-specific and exogenously given productivity parameter. The competitive market wage rate at time t is w_t. Agents make full use of labor. Thus, the labor income of an agent aged i at time t is $\epsilon_i w_t$.

Agents can invest in stocks and one-period riskless bonds[7]. Let P_t denote the ex-dividend price of the stock at time t, and D_t the dividend. The price of one unit of bond is 1. The bond claims a risk-free payoff R_t^f at time $t+1$. Let $a_{i,t+i-1}$ and $b_{i,t+i-1}$ denote the quantity of stocks and bonds held by an age-i agent born at time t. An agent may die leaving stocks and bonds behind. I assume that consumers of the newly-born generation equally inherit all the unclaimed assets. Let $a_{0,t-1}$ and $b_{0,t-1}$ denote the inherited securities by an agent, then

$$a_{0,t-1} = \frac{\sum_{i=1}^{I-1} \mu_i a_{i,t-1}(1-s_i)}{\mu_1}, \tag{3.6}$$

and

$$b_{0,t-1} = \frac{\sum_{i=1}^{I-1} \mu_i b_{i,t-1}(1-s_i)}{\mu_1}. \tag{3.7}$$

[7]Riskless investment technology must come from an institutional arrangement. Notice that a riskless apple tree does not exist In Lucas' world. The one-period return rate, from time t to $t+1$, of the i-th apple tree is defined as $r_{i,t+1} = \frac{p_{i,t+1}+d_{i,t+1}}{p_{it}}$. If the return rate of the tree is riskless, it is necessary that the numerator $p_{i,t+1} + d_{i,t+1}$ be without uncertainty conditional on all the information up to time t. However, this is impossible. Let us consider the pricing equation of the i-th apple tree at time $t+1$: $p_{i,t+1} = E_{t+1}\left[\sum_{j=1}^{\infty} \frac{\beta^j U'(c_{t+1+j})}{U'(c_{t+1})} d_{t+1+j}\right]$. Even if the future output process $\{d_{i,t+1}, d_{i,t+2}, \ldots\}$ is certain, the consumption process $\{c_{t+1}, c_{t+2}, \ldots\}$ may be still uncertain. The stochasticity of the consumption process $c_s = \sum_{i=1}^{n} d_{is}$ $(s \geq t+1)$ is attributed to the randomness of other apple trees' output. Therefore, $p_{i,t+1}$ is stochastic. So is the return rate $r_{i,t+1}$. Particularly, the trees producing a constant output at each time are not riskless. A riskless apple tree, if it exists, must live for no more than one period. At the beginning of next period, it produces a certain amount of apples and dies. Obviously, this tree does not exist at the steady-state equilibrium.

Therefore, the life-time budget constraint for an agent born at time t is:

$$\begin{aligned} c_{i,t+i-1} =&\epsilon_i w_{t+i-1} + a_{i-1,t+i-2}(P_{t+i-1} + D_{t+i-1}) + b_{i-1,t+i-2}R^f_{t+i-2} \\ &- a_{i,t+i-1}P_{t+i-1} - b_{i,t+i-1}, \;\; i \in \mathscr{I}. \end{aligned} \tag{3.8}$$

At each period, I have $2I-2$ Euler equations:

$$1 = E_t\left[\frac{\beta s_i c^{-\gamma}_{i+1,t+1}}{c^{-\gamma}_{i,t}}\right] R^f_t, \; (i = 1, \ldots, I-1), \tag{3.9}$$

$$P_t = E_t\left[\frac{\beta s_i c^{-\gamma}_{i+1,t+1}}{c^{-\gamma}_{i,t}} (P_{t+1} + D_{t+1})\right], \; (i = 1, \ldots, I-1). \tag{3.10}$$

I prohibit consumers from short-selling, i.e. $a_{i,t} \geq 0$ ($\forall i \in \mathscr{I}, \forall t$)[8]. Also, consumers are not allowed to borrow after retirement. Suppose I_1 is the retirement age. Thus, I restrict $b_{i,t} \geq 0$ ($I_1 + 1 \leq i \leq I, \forall t$). If the constraint is not binding, the Euler equation (3.9) or (3.10) holds. Otherwise, $a_{i,t}$ or $b_{i,t}$ is equal to zero. Allowing for constraints, the Euler equations (3.9) and (3.10) are upgraded to the following:

$$E_t\left[\beta s_i \left(\frac{c_{i+1,t+1}}{c_{i,t}}\right)^{-\gamma}\right] R^f_t - 1 = 0, \; (i = 1, \ldots, I_1), \tag{3.11}$$

$$b_{i,t}\left\{E_t\left[\beta s_i \left(\frac{c_{i+1,t+1}}{c_{i,t}}\right)^{-\gamma}\right] R^f_t - 1\right\} = 0, \; (i = I_1 + 1, \ldots, I-1), \tag{3.12}$$

$$a_{i,t}\left\{E_t\left[\beta s_i \left(\frac{c_{i+1,t+1}}{c_{i,t}}\right)^{-\gamma} \frac{(P_{t+1} + D_{t+1})}{P_t}\right] - 1\right\} = 0, \; (i = 1, \ldots, I-1). \tag{3.13}$$

3.3.2. The Firm and Income Distribution

A competitive firm is endowed with y_t consumption good at time t. I assume that the growth rate of the endowment $\lambda_{t+1} = \frac{y_{t+1}}{y_t}$ follows a two-state Markov chain process,

[8]The short-selling constraint is not binding in my calibration.

with transition probabilities given by

$$\phi_{ij} = \texttt{Prob}\{\lambda_{t+1} = \lambda_j | \lambda_t = \lambda_i\}, \tag{3.14}$$

where $i, j \in \{1, 2\}$, and $\sum_{j=1}^{2} \phi_{ij} = 1$, $\forall i$.

The firm has outstanding one unit of equity share which is perfectly divisible. Also, the firm issues one-period bonds which cost 1 at time t and promise a gross interest rate R_t^f at the beginning of time $t+1$. The competitive firm takes R_t^f as given.

The firm acts as an income distributor in the economy. It distributes a fixed proportion, α, of the endowment y_t to the owners of labor[9]. The firm pays the matured bonds $B_{t-1}R_{t-1}^f$ at time t. Now that the consumption good is perishable, the firm keeps no consumption good. All the remaining endowment, together with the revenue from issuing new bonds, B_t, is distributed to the owners of the equity shares. Thus, the dividend D_t is

$$D_t = (1-\alpha)y_t - B_{t-1}R_{t-1}^f + B_t. \tag{3.15}$$

I term equation (3.15) as the income distribution equation.

In this pure exchange economy, there is no physical capital. But this does not prevent us from studying the capital structure, since the observed ratio $\frac{B_t}{P_t}$ on the consumer side is equal to the debt-to-equity ratio on the firm side. Considering physical capital and production does not produce any new economic insight, but only complicates the computation. [10]

[9] The proportion of labor income is exogenously specified. Generally speaking, it is determined by the production function, say, Cobb-Douglas function.

[10] In an economy with production, it is required to assume that the management maximizes profits. This extra assumption can help to pin down the quantity of physical capital.

The income distribution equation is also the budget constraint of the firm. The firm's choice variable is bond, B_t or dividend D_t. The assumption that the consumption good is perishable enables us to focus on only the policy of bond issuing. The dividend policy is immediately implied by the policy of bond issuing.

The bond issuing changes the the probability distribution of the stock return. A rise (fall) of bond issued, B_t, increases (decreases) the current dividend, but decreases (increases) future payoffs of stocks. Thus, in my model, the probability distribution of the stock return is endogenously determined.

I assume that the management of the firm prefer to borrow a maximum amount of risk-less funds which are available in the market. The supply of the bond is perfectly elastic.

3.3.3. Equilibrium

In the equilibrium, four markets including the labor market, two asset markets, and the good market clear. We have,

$$\sum_{i=1}^{I} \mu_i \epsilon_i w_t = \alpha y_t, \tag{3.16}$$

$$\sum_{i=1}^{I-1} \mu_i a_{i,t} = 1, \tag{3.17}$$

$$\sum_{i=1}^{I-1} \mu_i b_{i,t} = B_t, \tag{3.18}$$

and

$$\sum_{i=1}^{I} \mu_i c_{i,t} = y_t.^{11} \tag{3.19}$$

At each time t, the dynamic system has $3I$ unknown variables, including P_t, R_t^f,

[11] It is easy to verify that equation (3.19) is redundant given the income distribution equation (3.15).

D_t, $a_{i,t}$,$(i = 1, 2, \ldots, I-2)$, $b_{i,t}$,$(i = 1, 2, \ldots, I-1)$, and $c_{i,t}$,$(i = 1, 2, \ldots, I)$. Also, it has $3I$ restrictions, including $2I - 2$ Euler equations (3.11) – (3.13), I number of budget constraints, the equation (3.5) and the income distribution equation (3.15). Thus, the solution of the dynamic system can be determined.

3.3.4. Transformation to a Stationary System

The solutions of P_t, R_t^f, $a_{i,t}(i = 1, \ldots, I-2)$ and $b_{i,t}(i = 1, \ldots, I-1)$ can be expressed as functions of the states including the endowment y_t, its growth rate λ_t, and the income distribution structure inherited from the previous period $(a_{1,t-1}, \ldots, a_{I-2,t-1}, b_{1,t-1}R_{t-1}^f, \ldots, b_{I-1,t-1}R_{t-1}^f)$.

Let

$$P_t = P(a_{1,t-1}, \ldots, a_{I-2,t-1}, b_{1,t-1}R_{t-1}^f, \ldots, b_{I-1,t-1}R_{t-1}^f, \lambda_t, y_t), \tag{3.20}$$

$$R_t^f = R^f(a_{1,t-1}, \ldots, a_{I-2,t-1}, b_{1,t-1}R_{t-1}^f, \ldots, b_{I-1,t-1}R_{t-1}^f, \lambda_t, y_t), \tag{3.21}$$

$$\begin{aligned} a_{i,t} =& a_i(a_{1,t-1}, \ldots, a_{I-2,t-1}, b_{1,t-1}R_{t-1}^f, \ldots, b_{I-1,t-1}R_{t-1}^f, \lambda_t, y_t), \\ & (i = 1, 2, \ldots, I-2), \end{aligned} \tag{3.22}$$

$$\begin{aligned} b_{i,t} =& b_i(a_{1,t-1}, \ldots, a_{I-2,t-1}, b_{1,t-1}R_{t-1}^f, \ldots, b_{I-1,t-1}R_{t-1}^f, \lambda_t, y_t), \\ & (i = 1, 2, \ldots, I-1). \end{aligned} \tag{3.23}$$

The income distribution structure should be stationary. As the endowment y_t is non-stationary, it is inferred that D_t, B_t, and $b_{i,t}$ $(i = 1, 2, \ldots, I-1)$ should be proportional to y_t up to a bounded functional. Considering that the risky return rate $\frac{P_{t+1}+D_{t+1}}{P_t}$ is a bounded functional, we infer that P_t is also proportional to y_t up to a bounded functional. Let $P_t = \tilde{P}_t y_t$, $b_{i,t} = \tilde{b}_{i,t} y_t$,$(i = 1, 2, \ldots, I-1)$ and $D_t = \tilde{D}_t y_t$. Thus,

$$D_t = \tilde{D}_t y_t = \left[(1-\alpha) - \frac{\tilde{B}_{t-1}\tilde{R}_{t-1}^f}{\lambda_t} + \tilde{B}_t \right] y_t. \tag{3.24}$$

The unknown functionals can be expressed as follows:

$$P_t = \tilde{P}(a_{1,t-1}, \dots, a_{I-2,t-1}, \tilde{b}_{1,t-1}R^f_{t-1}, \dots, \tilde{b}_{I-1,t-1}R^f_{t-1}, \lambda_t)y_t, \tag{3.25}$$

$$R^f_t = \tilde{R}^f(a_{1,t-1}, \dots, a_{I-2,t-1}, \tilde{b}_{1,t-1}R^f_{t-1}, \dots, \tilde{b}_{I-1,t-1}R^f_{t-1}, \lambda_t), \tag{3.26}$$

$$\begin{aligned} a_{i,t} =& \tilde{a}_i(a_{1,t-1}, \dots, a_{I-2,t-1}, \tilde{b}_{1,t-1}R^f_{t-1}, \dots, \tilde{b}_{I-1,t-1}R^f_{t-1}, \lambda_t), \\ & (i = 1, 2, \dots, I-2), \end{aligned} \tag{3.27}$$

$$\begin{aligned} b_{i,t} =& \tilde{b}_i(a_{1,t-1}, \dots, a_{I-2,t-1}, \tilde{b}_{1,t-1}R^f_{t-1}, \dots, \tilde{b}_{I-1,t-1}R^f_{t-1}, \lambda_t)y_t. \\ & (i = 1, 2, \dots, I-1). \end{aligned} \tag{3.28}$$

Define $\pi_i \equiv \frac{\alpha\epsilon_i}{\sum_{i=1}^{I}\mu_i\epsilon_i}$. Then labor income of age-i agent born at time t is $\epsilon_i w_{t+i-1} = \pi_i y_{t+i-1}$. The agent's life-time budget constraint becomes

$$\begin{aligned} c_{i,t+i-1} =& \pi_i y_{t+i-1} + a_{i-1,t+i-2}(P_{t+i-1} + D_{t+i-1}) + b_{i-1,t+i-2}R^f_{t+i-2} \\ & - a_{i,t+i-1}P_{t+i-1} - b_{i,t+i-1}, \quad i \in \mathscr{I}. \end{aligned} \tag{3.29}$$

Let $c_{i,t} = \tilde{c}_{i,t}y_t$. The budget constraint (3.29) now becomes

$$\begin{aligned} \tilde{c}_{i,t+i-1} =& \pi_i + a_{i-1,t+i-2}(\tilde{P}_{t+i-1} + \tilde{D}_{t+i-1}) + \frac{\tilde{b}_{i-1,t+i-2}R^f_{t+i-2}}{\lambda_{t+i-1}} \\ & - a_{i,t+i-1}\tilde{P}_{t+i-1} - \tilde{b}_{i,t+i-1}, \quad i \in \mathscr{I}. \end{aligned} \tag{3.30}$$

Substituting (3.15), (3.25) and (3.30) into (3.11) – (3.13), the Euler equations become

$$E_t\left[\beta s_i \lambda_{t+1}^{-\gamma}\left(\frac{\tilde{c}_{i+1,t+1}}{\tilde{c}_{i,t}}\right)^{-\gamma}\right]R^f_t - 1 = 0, \ (i = 1, \dots, I_1), \tag{3.31}$$

$$\tilde{b}_{i,t}\left\{E_t\left[\beta s_i \lambda_{t+1}^{-\gamma}\left(\frac{\tilde{c}_{i+1,t+1}}{\tilde{c}_{i,t}}\right)^{-\gamma}\right]R^f_t - 1\right\} = 0, \ (i = I_1 + 1, \dots, I-1), \tag{3.32}$$

$$a_{i,t}\left\{E_t\left[\beta s_i \lambda_{t+1}^{1-\gamma}\left(\frac{\tilde{c}_{i+1,t+1}}{\tilde{c}_{i,t}}\right)^{-\gamma}\frac{\left(\tilde{P}_{t+1} + \tilde{D}_{t+1}\right)}{\tilde{P}_t}\right] - 1\right\} = 0, \ (i = 1, \dots, I-1). \tag{3.33}$$

3.4. Calibration

3.4.1. Parameters

The technology parameters follow those derived by Mehra and Prescott (1985). Mehra and Prescott observe that a two-state Markov chain with $\lambda_1 = 1.054$, $\lambda_2 = 0.982$, $\phi_{11} = \phi_{22} = 0.43$ and $\phi_{12} = \phi_{21} = 0.57$, fairly well approximates the annual growth rate of per capita consumption over the period 1889-1978. The labor share of income α is set to be 0.7. It is consistent with the estimation by Alan B. Krueger (1999) and Douglas Gollin (2002).

The labor efficiency ϵ_i's are listed in Table (XI). The labor efficiency of males and females from age 21 to 60 is adopted from G. D. Hansen (1993). I assume that the agents enter my model at the age of 21 and retire at age 60 ($I_1 = 40$). After retirement, the efficiency is set to be 0. In the calibration, I apply the average efficiency of males and females. The survival rates s_i's are based on U.S. life table of 1901. The maximum age is set to be 95 ($I = 75$).

Table XI.: Life-cycle Labor Efficiency

Age	21-25	26-35	36-45	46-55	56-60	61-95
Average	0.735	1.015	1.135	1.130	1.085	0

3.4.2. Computation

The computation of a large-scale overlapping generations model with aggregate uncertainty is challenging. Since the portfolios of financial assets across generations have to be included as state variables in the argument list of endogenous choice functions, one faces a curse of dimensionality as the life span of agents increases. I provide an infeasible Taylor series approximation to show the difficulty in computation. Then, I present a "reduced-form" numerical solution. I express all the unknown variables as functions of current and past exogenous state variables which drive the whole dynamic system. A feasible solution is to approximate Euler equations using finite latest exogenous state variables.

The procedure of Taylor series approximation is as follows. First, I introduce separate functionals for different values of λ_t's. Hence, $\tilde{P}(\ldots,\lambda_t)$ is decomposed into two functionals: $\tilde{P}_1$ when $\lambda_t = \lambda_1$ and $\tilde{P}_2$ when $\lambda_t = \lambda_2$. Similarly, $\tilde{R}^f(\ldots,\lambda_t)$ is decomposed into $\tilde{R}^f_1$ and $\tilde{R}^f_2$, $\tilde{a}_i(\ldots,\lambda_t)$ into $\tilde{a}_{i,1}$ and $\tilde{a}_{i,2}$, and $\tilde{b}_i(\ldots,\lambda_t)$ into $\tilde{b}_{i,1}$ and $\tilde{b}_{i,2}$.

Second, let us approximate the unknown functionals using a first-order approximation. I recursively define

$$\hat{P}_{1,t} = x_{p1} \cdot (1, \hat{a}_{1,t-1}, \ldots, \hat{a}_{I-2,t-1}, \hat{b}_{1,t-1}\hat{R}^f_{t-1}, \ldots, \hat{b}_{I-1,t-1}\hat{R}^f_{t-1})', \tag{3.34}$$

$$\hat{P}_{2,t} = x_{p2} \cdot (1, \hat{a}_{1,t-1}, \ldots, \hat{a}_{I-2,t-1}, \hat{b}_{1,t-1}\hat{R}^f_{t-1}, \ldots, \hat{b}_{I-1,t-1}\hat{R}^f_{t-1})', \tag{3.35}$$

$$\hat{R}^f_{1,t} = x_{R^f1} \cdot (1, \hat{a}_{1,t-1}, \ldots, \hat{a}_{I-2,t-1}, \hat{b}_{1,t-1}\hat{R}^f_{t-1}, \ldots, \hat{b}_{I-1,t-1}\hat{R}^f_{t-1})', \tag{3.36}$$

$$\hat{R}^f_{2,t} = x_{R^f2} \cdot (1, \hat{a}_{1,t-1}, \ldots, \hat{a}_{I-2,t-1}, \hat{b}_{1,t-1}\hat{R}^f_{t-1}, \ldots, \hat{b}_{I-1,t-1}\hat{R}^f_{t-1})', \tag{3.37}$$

$$\hat{a}_{i,1,t} = x_{ai,1} \cdot (1, \hat{a}_{1,t-1}, \ldots, \hat{a}_{I-2,t-1}, \hat{b}_{1,t-1}\hat{R}^f_{t-1}, \ldots, \hat{b}_{I-1,t-1}\hat{R}^f_{t-1})', (i = \{1, \ldots, I-2\}), \tag{3.38}$$

$$\hat{a}_{i,2,t} = x_{ai,2} \cdot (1, \hat{a}_{1,t-1}, \ldots, \hat{a}_{I-2,t-1}, \hat{b}_{1,t-1}\hat{R}^f_{t-1}, \ldots, \hat{b}_{I-1,t-1}\hat{R}^f_{t-1})', (i = \{1, \ldots, I-2\}), \tag{3.39}$$

$$\hat{b}_{i,1,t} = x_{bi,1} \cdot (1, \hat{a}_{1,t-1}, \ldots, \hat{a}_{I-2,t-1}, \hat{b}_{1,t-1}\hat{R}^f_{t-1}, \ldots, \hat{b}_{I-1,t-1}\hat{R}^f_{t-1})', (i = \{1, \ldots, I-1\}) \tag{3.40}$$

and

$$\hat{b}_{i,2,t} = x'_{bi,2} \cdot (1, \hat{a}_{1,t-1}, \ldots, \hat{a}_{I-2,t-1}, \hat{b}_{1,t-1}\hat{R}^f_{t-1}, \ldots, \hat{b}_{I-1,t-1}\hat{R}^f_{t-1})'(i = \{1, \ldots, I-1\}), \tag{3.41}$$

to approximate the functionals $\tilde{P}_1$, $\tilde{P}_2$, $\tilde{R}^f_1$, $\tilde{R}^f_2$, $\tilde{a}_{i,1}(i = \{1, \ldots, I-2\})$, $\tilde{a}_{i,2}(i = \{1, \ldots, I-2\})$, $\tilde{b}_{i,1}(i = \{1, \ldots, I-1\})$, and $\tilde{b}_{i,2}(i = \{1, \ldots, I-1\})$. In these formulas, x's are time-invariant coefficient vectors.

Third, I generate T observations of λ_t's. Next, I pick a set of initial values for x's, a, b and R^f for time $t = 1$. Starting from $t = 2$, I recursively approximate unknown functionals using (3.34) – (3.41), calculate $\hat{D}_t$, $\hat{a}_{0,t-1}$, $\hat{b}_{0,t-1}$ and $\hat{c}_{i,t}$'s based on equations (3.24), (3.6), (3.7) and (3.30), and then compute the approximation error of the new versions of Euler Equations (3.31) – (3.33). Let $SAE_{a,i,t}(i = 1, \ldots, I-1)$ and $SAE_{b,i,t}(i = 1, \ldots, I-1)$ denote the squared approximation errors of the corresponding restrictions at time t. I define the mean squared approximation error as

$$MSAE = \frac{1}{(T-1)(2I-1)} \sum_{t=2}^{T} \left\{ \sum_{i=1}^{I-1} [SAE_{a,i,t} + SAE_{b,i,t}] + (c_{1,t} - \rho c_{20,t})^2 \right\}. \tag{3.42}$$

The unknown coefficients x's can be solved by minimizing the MSAE.

The issue of the method of Taylor series approximation is the curse of dimensionality. Even the first order approximation involves an optimization problem with $4(I-1)(2I-1)$ (equals 47064 if $I = 75$) variables. It is obviously infeasible in practice. I turn to a discrete approximation method.

The motivation of the "reduced-form" (discrete) approximation method arises from the fact that the dynamic system is driven by the exogenous process $\{\lambda_t\}$ and each endogenous variable at time t should be a function of historical state values

$\{\lambda_s, s \leq t\}$. Let us define time-invariant discrete functions

$$\hat{P}_t = P(\lambda_{t-L+1}, \lambda_{t-L+2}, \ldots, \lambda_t) \tag{3.43}$$

$$\hat{R}^f_t = R^f(\lambda_{t-L+1}, \lambda_{t-L+2}, \ldots, \lambda_t) \tag{3.44}$$

$$\hat{a}_{i,t} = a_i(\lambda_{t-L+1}, \lambda_{t-L+2}, \ldots, \lambda_t),\ i = 1, \ldots, I-2 \tag{3.45}$$

$$\hat{b}_{i,t} = b_i(\lambda_{t-L+1}, \lambda_{t-L+2}, \ldots, \lambda_t),\ i = 1, \ldots, I-1 \tag{3.46}$$

to approximate the unknown functionals, where $P(\cdot)$, $Rf(\cdot)$, $a_i(\cdot)$, and $b_i(\cdot)$ are discrete functionals taking on 2^L values. I introduce a distinct variable for each value. Therefore, the optimization problem involves $(2I-1)2^L$ choice variables if I include L lags to approximate each functional. Obviously, the curse of dimensionality is also an issue for the approach of discrete approximation if L is too large. But, as L is small, a rough approximation is feasible. My strategy is start from a rough approximation and increase the precision of approximation by adding more and more lags. Each time I increase one more lag, the less precise solution can be the initial values of the choice variables in the extended optimization problem. For example, if $P(\lambda_{t-L+1}, \lambda_{t-L+2}, \ldots, \lambda_t) = \hat{P}$ as L lags are included, the initial values of $P(\lambda_{t-L} = \lambda_1, \lambda_{t-L+1}, \lambda_{t-L+2}, \ldots, \lambda_t)$ and $P(\lambda_{t-L} = \lambda_2, \lambda_{t-L+1}, \lambda_{t-L+2}, \ldots, \lambda_t)$ in the enlarged optimization problem will be assigned to be equal to $\hat{P}$.

The detailed procedure of L-order discrete approximation is as follows. First, pick a set of initial values for P_t, R^f_t, $a_{i,t}(i = 1, \ldots, I-2)$ and $b_{i,t}(i = 1, \ldots, I-1)$ for all of their discrete states. Since the discrete functions are time-invariant, I automatically assign the values for R^f_{t-1}, $a_{i,t-1}(i = 1, \ldots, I-2)$, $b_{i,t-1}(i = 1, \ldots, I-1)$, P_{t+1}, $a_{i,t+1}(i = 1, \ldots, I-2)$ and $b_{i,t+1}(i = 1, \ldots, I-1)$ for all of their discrete states. Second, calculate $\hat{D}_t$, $\hat{a}_{0,t-1}$, $\hat{b}_{0,t-1}$ and $\hat{c}_{i,t}$'s based on equations (3.24), (3.6), (3.7) and (3.30), and then compute the approximation errors of Euler Equa-

Table XII.: Calibration Results ($\gamma = 1.2, \beta = 1.03, \rho = 0.7$)

order	# of *Variables*	*MSAE*	Risk-Free Rate $E(R_t^f)$	Risky Rate $E(\frac{P_{t+1}+D_{t+1}}{P_t})$	Premium	Debt-to-Capital Ratio $E(\frac{B_t}{B_t+P_t})$
1	298	1.1152e-3	1.0047	1.0493	0.0446	0.3356
3	1192	6.5306e-4	1.0070	1.0477	0.0407	0.3342
5	4768	5.6824e-4	1.0072	1.0477	0.0405	0.3353
7	19072	5.3191e-4	1.0074	1.0478	0.0405	0.3359
9	76288	5.1759e-4	1.0074	1.0478	0.0404	0.3361

tions (3.31) – (3.33). Let $SAE_{a,i}(\lambda_{t-L}, \lambda_{t-L+1}, \lambda_{t-L+2}, \ldots, \lambda_t)(i = 1, \ldots, I-1)$ and $SAE_{b,i}(\lambda_{t-L}, \lambda_{t-L+1}, \lambda_{t-L+2}, \ldots, \lambda_t)(i = 1, \ldots, I-1)$ denote the squared approximation errors of the corresponding restrictions if the latest historical states are $(\lambda_{t-L}, \lambda_{t-L+1}, \lambda_{t-L+2}, \ldots, \lambda_t)$. I define the mean squared approximation error as

$$\begin{aligned} MSAE = & \frac{1}{2^{L+1}(2I-1)} \sum_{\lambda_k \in \{\lambda_1, \lambda_2\}, t-L \le k \le t} \\ & \left\{ \sum_{i=1}^{I-1} \left[SAE_{a,i}(\lambda_{t-L}, \ldots, \lambda_t) + SAE_{b,i}(\lambda_{t-L}, \ldots, \lambda_t) \right] + (c_{1,t} - \rho c_{20,t})^2 \right\}. \end{aligned} \tag{3.47}$$

The L-order discrete approximating functionals can be solved by minimizing the MSAE.

3.4.3. Calibration Results

The calibration results are reported in Tables (XII)-(XVII). The degree of relative risk aversion γ is set to be plausible: 1.2, 1.5, or 2.0. The discount factor β is set between 1.02 and 1.05. I include up to 9 lags to approximate the unknown functions, which involves $76,288$ variables in the optimization problem. The estimated equity premium and debt-to-capital ratio are relatively stable as I increase the lags from 1

Table XIII.: Calibration Results ($\gamma = 1.5, \beta = 1.02, \rho = 0.83$)

order	# of *Variables*	*MSAE*	Risk-Free Rate $E(R_t^f)$	Risky Rate $E(\frac{P_{t+1}+D_{t+1}}{P_t})$	Premium	Debt-to-Capital Ratio $E(\frac{B_t}{B_t+P_t})$
1	298	1.1106e-3	1.0131	1.0548	0.0418	0.3201
3	1192	6.0381e-4	1.0162	1.0526	0.0365	0.3184
5	4768	4.9135e-4	1.0177	1.0524	0.0347	0.3229
7	19072	4.4296e-4	1.0181	1.0525	0.0344	0.3251
9	76288	4.2698e-4	1.0182	1.0525	0.0343	0.3252

Table XIV.: Calibration Results ($\gamma = 1.5, \beta = 1.03, \rho = 0.8$)

order	# of *Variables*	*MSAE*	Risk-Free Rate $E(R_t^f)$	Risky Rate $E(\frac{P_{t+1}+D_{t+1}}{P_t})$	Premium	Debt-to-Capital Ratio $E(\frac{B_t}{B_t+P_t})$
1	298	1.2423e-3	1.0075	1.0512	0.0437	0.3261
3	1192	6.2926e-4	1.0101	1.0494	0.0392	0.3250
5	4768	5.3783e-4	1.0109	1.0493	0.0384	0.3276
7	19072	4.7867e-4	1.0111	1.0495	0.0384	0.3289
9	76288	4.5312e-4	1.0131	1.0500	0.0370	0.3311

to 9.

The estimated riskless rates are low, given a plausible degree of risk aversion γ and the discount factor β. As $\gamma = 1.2$ and $\beta = 1.03$, the calibrated riskless rate is around 0.74 percent. As $\gamma = 1.5$ and $\beta = 1.04$, the riskless rate is around 0.54 percent. In my model, the low riskless rate is primarily determined by the discount factor β. The more individuals care about future consumption, the more they would save. A high saving leads to a low interest rate. For example, an increase of β from 1.3 to 1.4 (γ=1.5), the riskless interest rate decreases from 1.31 percent to 0.54 percent. Another factor influencing the level of riskless rate is the degree of risk aversion. If individuals become more risk averse, the riskless rate may increases because they

Table XV.: Calibration Results ($\gamma = 1.5, \beta = 1.04, \rho = 0.72$)

$order$	# of $Variables$	$MSAE$	Risk-Free Rate $E(R_t^f)$	Risky Rate $E(\frac{P_{t+1}+D_{t+1}}{P_t})$	Premium	Debt-to-Capital Ratio $E(\frac{B_t}{B_t+P_t})$
1	298	1.1944e-3	1.0034	1.0481	0.0447	0.3305
3	1192	6.2503e-4	1.0055	1.0468	0.0413	0.3295
5	4768	5.3323e-4	1.0052	1.0470	0.0418	0.3303
7	19072	4.9246e-4	1.0053	1.0471	0.0418	0.3310
9	76288	4.7547e-4	1.0054	1.0471	0.0418	0.3312

Table XVI.: Calibration Results ($\gamma = 2.0, \beta = 1.04, \rho = 0.8$)

$order$	# of $Variables$	$MSAE$	Risk-Free Rate $E(R_t^f)$	Risky Rate $E(\frac{P_{t+1}+D_{t+1}}{P_t})$	Premium	Debt-to-Capital Ratio $E(\frac{B_t}{B_t+P_t})$
1	298	1.2814e-3	1.0094	1.0500	0.0406	0.3370
3	1192	5.7145e-4	1.0122	1.0482	0.0360	0.3356
5	4768	4.3522e-4	1.0129	1.0479	0.0350	0.3371
7	19072	3.7867e-4	1.0131	1.0480	0.0349	0.3380
9	76288	3.5358e-4	1.0131	1.0480	0.0349	0.3383

Table XVII.: Calibration Results ($\gamma = 2.0, \beta = 1.05, \rho = 0.74$)

$order$	# of $Variables$	$MSAE$	Risk-Free Rate $E(R_t^f)$	Risky Rate $E(\frac{P_{t+1}+D_{t+1}}{P_t})$	Premium	Debt-to-Capital Ratio $E(\frac{B_t}{B_t+P_t})$
1	298	1.3449e-3	1.0071	1.0475	0.0404	0.3443
3	1192	5.8591e-4	1.0095	1.0459	0.0365	0.3439
5	4768	4.5310e-4	1.0099	1.0459	0.0359	0.3451
7	19072	3.9572e-4	1.0100	1.0460	0.0360	0.3455
9	76288	3.7005e-4	1.0100	1.0460	0.0360	0.3458

would demand more bonds to hedge against risks. For examples, we can make a comparison of Tables (XII) and (XIV) and a comparison of Tables (XV) and (XVI).

The calibrated equity premium is around 4%, which is consistent with Jeremy J. Siegel (1998)'s estimate with a larger sample from 1802 to 1998. An equity premium as high as 6 percent can be generated only if I set β to be a large number. But the corresponding riskless rate becomes implausibly low. I conclude that the bulk of equity premium can be explained in my parsimonious model. The residual premium should be ascribed to other numerous factors such as rare disaster, market incompleteness and imperfections.

The estimated debt-to-capital ratio is around 1/3, which is consistent with empirical evidence in the U.S. presented by Raghuram G. Rajan and Luigi Zingales (1995) and Ronald W. Masulis (1988). Rajan and Zingales record that the median U.S. market debt-to-capital ratio in 1991 is 28%. Masulis reports that the debt-to-capital ratio of U.S. firms has varied between 13% and 44% from 1929 to 1986. The under-leverage puzzle in the literature can be explained by the fact that the equity return rate is much larger than the riskless rate. An average individual prefers to invest more in stocks than in bonds.

Table (XVIII) shows a cross-section snapshot of portfolio choices and consumption for different cohorts. The short-selling constraint is not binding, whereas the retirees-can-not-leverage constraint is binding. Relaxing this constraint will induce the retirees to borrow and buy more risky assets. This reduces the leverage effect and decreases the equity premium. Therefore, borrowing constraints matter.

3.4.4. Endogenous Leverage and Equity Premium

My calibration shows that an endogenous leverage produces an even larger premium than a fixed leverage does. From proposition II of MM theorem, the equity premium

Table XVIII.: Life-cycle Evolution of Portfolios and Consumption

Age	21	22	23	24	25	26	27	28	29	30
E(a)	0.854	0.760	0.634	0.703	0.735	0.750	0.574	0.556	0.560	0.707
$E(\tilde{b})$	5.941	7.375	9.162	8.268	7.888	7.993	10.605	11.006	11.100	9.267
$E(\tilde{c})$	0.688	0.704	0.720	0.736	0.752	0.768	0.784	0.799	0.814	0.830
Age	31	32	33	34	35	36	37	38	39	40
E(a)	0.737	0.565	0.516	0.675	0.786	0.914	0.673	0.649	0.631	0.724
$E(\tilde{b})$	9.057	11.522	12.245	10.174	8.808	7.360	10.869	11.347	11.723	10.594
$E(\tilde{c})$	0.844	0.859	0.874	0.888	0.902	0.916	0.930	0.943	0.956	0.969
Age	41	42	43	44	45	46	47	48	49	50
E(a)	0.754	0.795	0.875	0.999	1.037	0.989	1.039	1.055	1.116	1.195
$E(\tilde{b})$	10.338	9.940	9.036	7.573	7.313	8.193	7.725	7.726	7.136	6.315
$E(\tilde{c})$	0.982	0.994	1.007	1.019	1.030	1.042	1.053	1.063	1.073	1.084
Age	51	52	53	54	55	56	57	58	59	60
E(a)	1.231	1.213	1.269	1.357	1.347	1.427	1.439	1.520	1.643	1.675
$E(\tilde{b})$	6.099	6.602	6.105	5.203	5.639	4.824	4.942	4.146	2.832	2.775
$E(\tilde{c})$	1.093	1.103	1.112	1.121	1.129	1.136	1.142	1.147	1.152	1.156
Age	61	62	63	64	65	66	67	68	69	70
E(a)	1.685	1.663	1.697	1.672	1.639	1.615	1.579	1.538	1.497	1.448
$E(\tilde{b})$	2.098	1.840	0.842	0.650	0.553	0.313	0.217	0.153	0.079	0.058
$E(\tilde{c})$	1.160	1.157	1.155	1.159	1.164	1.170	1.181	1.194	1.209	1.226
Age	71	72	73	74	75	76	77	78	79	80
E(a)	1.397	1.342	1.284	1.221	1.155	1.086	1.013	0.938	0.860	0.783
$E(\tilde{b})$	0.039	0.028	0.020	0.016	0.013	0.013	0.014	0.016	0.028	0.029
$E(\tilde{c})$	1.245	1.267	1.291	1.316	1.339	1.360	1.373	1.378	1.373	1.354
Age	81	82	83	84	85	86	87	88	89	90
E(a)	0.705	0.626	0.545	0.471	0.406	0.325	0.227	0.221	0.183	0.140
$E(\tilde{b})$	0.032	0.058	0.151	0.181	0.128	0.352	0.867	0.206	0.066	0.053
$E(\tilde{c})$	1.324	1.280	1.225	1.158	1.083	1.002	0.913	0.816	0.728	0.641
Age	91	92	93	94	95					
E(a)	0.104	0.071	0.043	0.010	0.000					
$E(\tilde{b})$	0.029	0.026	0.025	0.136	0.000					
$E(\tilde{c})$	0.558	0.478	0.404	0.336	0.270					

$\gamma = 1.5$, $\beta = 1.04$, and $\alpha = 0.7$.

can be expressed as

$$R^r - R^f = (\bar{R} - R^f)(1 + \frac{B}{P})$$

where R^r, $\bar{R}$, and $\frac{B}{P}$ denote the risky return rate, average return rate, and debt-to-equity ratio respectively. Obviously, a fixed leverage is able to generate a large equity premium. In my model, the firm adjusts the debt level, thus the leverage ratio, according to the economy states. As follows is the correlation matrix of λ_{t+1}, $\bar{R}_t$, R^f_t, $\frac{B_t}{P_t}$, B_t and D_{t+1} calibrated from the model with $\gamma = 1.2$, $\beta = 1.03$.

$$corr(\lambda_{t+1}, \bar{R}_t, R^f_t, \frac{B_t}{P_t}, B_t, D_{t+1}) = \begin{pmatrix} 1.00 & 0.91 & -0.03 & 0.13 & 0.13 & -0.66 \\ 0.91 & 1.00 & 0.16 & \underline{0.10} & 0.02 & -0.56 \\ -0.03 & 0.16 & 1.00 & \underline{-0.35} & -0.58 & 0.21 \\ 0.13 & \underline{0.10} & \underline{-0.35} & 1.00 & 0.93 & -0.80 \\ 0.13 & 0.02 & -0.58 & 0.93 & 1.00 & -0.77 \\ -0.66 & -0.56 & 0.21 & -0.80 & -0.77 & 1.00 \end{pmatrix}$$

Intuitively, if the growth rate λ_{t+1} is high or the risk-free rate R^f_t is low, the firm chooses to raise the leverage ratio to increase the risky return rate; however, if the growth rate is low or the risk-free rate is high, the firm opts to lower the ratio to reduce the loss. Obviously, compared to a fixed leverage, the endogenous leverage results in an even larger equity premium. In the above model,

$$E(\bar{R}_t) = 1.0341 > E(\lambda_{t+1}) = 1.0180.$$

3.4.5. Robustness Checks

Another method of solving the dynamic system is to minimize the mean squared stochastic error. Define squared stochastic errors as follows

$$SSE_{b,i} = \left\{ \left[\beta s_i \lambda_{t+1}^{-\gamma} \left(\frac{\hat{c}_{i+1,t+1}}{\hat{c}_{i,t}} \right)^{-\gamma} \right] \hat{R}_t^f - 1 \right\}^2, \ (i = 1, \ldots, I_1), \tag{3.48}$$

$$SSE_{b,i} = 1(\hat{b}_{i,t} > 0) \left\{ \left[\beta s_i \lambda_{t+1}^{-\gamma} \left(\frac{\hat{c}_{i+1,t+1}}{\hat{c}_{i,t}} \right)^{-\gamma} \right] \hat{R}_t^f - 1 \right\}^2, \ (i = I_1 + 1, \ldots, I - 1), \tag{3.49}$$

$$SSE_{a,i} = 1(\hat{a}_{i,t} > 0) \left\{ \left[\beta s_i \lambda_{t+1}^{1-\gamma} \left(\frac{\hat{c}_{i+1,t+1}}{\hat{c}_{i,t}} \right)^{-\gamma} \frac{\left(\hat{P}_{t+1} + \hat{D}_{t+1} \right)}{\hat{P}_t} \right] - 1 \right\}^2, \ (i = 1, \ldots, I - 1). \tag{3.50}$$

The mean squared stochastic error is defined as

$$MSSE = \frac{1}{2^{L+1}(2I-2)} \sum_{\lambda_k \in \{\lambda_1, \lambda_2\}, t-L \le k \le t} \sum_{i=1}^{I-1} \left[SSE_{a,i}(\lambda_{t-L}, \ldots, \lambda_t) + SSE_{b,i}(\lambda_{t-L}, \ldots, \lambda_t) \right]. \tag{3.51}$$

The estimates are very close to the results reported in Tables (XII)-(XVII).

3.5. Summary

I resolve the equity premium puzzle, the risk free-rate puzzle, and the capital structure puzzle. In my large-scale overlapping generations model with a simple CRRA utility function and the rational expectation, the calibrated debt-to-capital ratio and risk-free rate are consistent with empirical evidence and the bulk of the equity premium is accounted for.

The capital structure is endogenously determined purely by consumers' hedging behavior. Modigliani and Miller (1958)'s capital structure irrelevance steps from the assumption that the risk-neutral management is the decision maker of corporate financing. However, as I argue, the usual one single objective applied by the management, whatever it is, is not valid to study the financing behavior of a firm. I show that investors collectively determine the capital structure through trading assets in the financial market. The endogenously determined capital structure is able to generate an even larger equity premium than a fixed leverage.

The endogenous formation of capital structure might help to understand why the CRRA utility function and rational expectation theory work well in the area of real business cycle but not in asset pricing. Typically, the asset or capital is not divided into the risk asset and the riskless asset in the real business cycle. In traditional asset pricing models, this division exists but only in a conceptual level. The riskless asset actually is not introduced in most models since the aggregate bond among consumers at the equilibrium is zero. Correspondingly, the returns of stocks are not volatile enough to generate a high equity premium.

Lastly, my resolution of asset pricing puzzles may help to understand why the asset pricing puzzles are robust. The unanimous rejections of Euler equations tests in the literature may arise from the application of per capita consumption. A test based on longitudinal data is promising to find empirical evidence to support the asset pricing models based on life-cycle consumption. This work is left to future research.

CHAPTER IV

WEAK COALITION EQUILIBRIUM

4.1. Introduction

A drawback of Nash's (1951) fundamental solution concept, non-cooperative equilibrium, is Pareto inefficiency. The notion of Nash equilibrium assumes that each participant is rational and acts independently. In some cases, however, players limited to individual rationality may be trapped in an inefficient outcome which can be avoided by collaboration.

Several theories have been developed to incorporate the consideration of coalitional collaboration into the non-cooperative game theory. Aumann (1959) pioneers the study of multilateral deviations by defining the concept Strong Nash equilibrium. Bernheim, Peleg, and Whinston (1987) propose the concept of Coalition-Proof Nash equilibrium which is widely applied in economic contexts. Assuming that the game players are farsighted, Chwe (1994) and Mariotti (1997) consider the dynamics of coalition formation and offer notions of Largest Consistent Set and Coalition Equilibrium, respectively. Ambrus (2006) has recently defined a non-equilibrium theory, Coalitional Rationalizability. These concepts and many others greatly enhance our understanding of group behavior.

However, most existing coalition solution concepts violate the fundamental principle of individual rationality. For example, for the well-known game of the Prisoner's Dilemma, the notions of Largest Consistent Set and Coalition Equilibrium predict the dominated strategy (Don't Confess, Don't Confess), whereas only the strategy (confess, confess) is individually rationalizable. The violation of the principle of individual rationality may exclude reasonable strategies. Let us take the game in Table

(XIX) as an example. In this game, strategies (A_1, B_1, C_1) and (A_2, B_2, C_2) are Nash equilibria. The strategy (A_1, B_1, C_1) is obviously a reasonable coalitional outcome. Although players A and B can jointly deviate from (A_1, B_1, C_1) to (A_2, B_2, C_1), this deviation is perhaps individually irrational if each individual is farsighted. The maximum potential gain from this deviation is 0.1, while the possible loss caused by player C's further deviation is 19. A convincing notion may have to retain the strategy (A_1, B_1, C_1). Unfortunately, the notion of Coalition-Proof Nash equilibrium and the theory of Coalitional Rationalizability predict the strategy (A_2, B_2, C_2) but completely exclude the strategy (A_1, B_1, C_1). Further, the violation of the principle of individual rationality may lead to the non-existence of a solution. For the game in Table (XIX), the set of Strong Nash equilibrium is empty. In section 4.3, we will see two examples of the non-existence of Coalition-Proof Nash equilibrium.

Table XIX.: A Modified Version of Ambrus' Example

C1			C2		
	B1	B2		B1	B2
A1	20, 20, 20	0, 0, 0	A1	0, 0, 0	0, 0, 0
A2	0, 0, 0	20.1, 20.1, 0	A2	0, 0, 0	1, 1, 1

The objective of this essay is to develop a coalition concept which is in line with the fundamental principle of individual rationality and guarantees the existence of the solution. This is done by defining a weakening criterion of coalition-proofness that requires players to deviate only when the deviation is strongly self-enforceable, i.e., credible not only within but also across coalitions. Based on the notion of within-coalition self-enforceability derived from the concept of Coalition-Proof Nash equilibrium, I define three new notions: cross-coalition self-enforceability, strong self-enforceability, and Weak Coalition equilibrium. A deviation is strongly self-enforceable if it is self-enforceable both within the coalition and across the coalitions. I term the solution concept as Weak Coalition equilibrium since fully farsighted players in our concept are very cautious when considering a coalitional deviation. Weak Coalition equilibrium is a generalization of Coalition-Proof Nash equilibrium. All the Coalition-Proof Nash equilibria are contained in the set of Weak Coalition equilibria which is further a subset of Nash equilibria. Weak Coalition equilibrium exists under a very weak condition. It exists if a Nash equilibrium exists.

This essay is organized as follows. In section 4.2, I will define the solution concept Weak Coalition equilibrium and discuss its properties. In section 4.3, I present some examples.

4.2. Definitions and Characterization

I assume that players in a normal form game are individually rational and fully farsighted. An arbiter prescribes a strategy for all the players before they simultaneously play the game. Players can freely discuss the strategy and can coordinate to deviate. However, players cannot make binding commitments. Introducing an arbiter helps us understand the forthcoming concepts. But it is inessential.

I denote the normal form game as $G = [I, (S_i)_{i \in I}, (U_i)_{i \in I}]$, where $I = \{1, 2, ..., n\}$ denotes a finite set of players, S_i the strategy set of player i, and $U_i : \prod_{i \in I} S_i \mapsto \mathbb{R}$ the payoff function of player i. A coalition J is a nonempty subset of I. Let $\mathscr{C} = \{J : J \neq \emptyset, J \subseteq I\}$ denote all the possible coalitions. I call J a singleton coalition if its size is 1, a complete coalition if its size is n. Let $-J$ denote the complement of J in I. Let $\mathscr{C}_J = \{q \in \mathscr{C} : q \neq J, q \subseteq J\}$ be the set of proper sub-coalitions of coalition J. If J is a singleton coalition, $\mathscr{C}_J = \{\}$.[1] Let $S_J \equiv \prod_{i \in J} S_i$, particularly

[1] To avoid possible confusion, I designate $\emptyset$ and $\{\}$ to be the empty subsets of I and $\mathscr{C}$.

$S \equiv S_I$. Finally, for each $s_{-J} \in S_{-J}$, let G/s_{-J} denote the game induced on the subgroup J by the actions s_{-J} for coalition $-J$, i.e.,

$$G/s_{-J} \equiv \left[J, (S_i)_{i \in J}, (\tilde{U}_i)_{i \in J}\right],$$

where $\tilde{U}_i : S_J \mapsto \mathbb{R}$ is given by $\tilde{U}_i(s_J) = U_i(s_J, s_{-J})$ for all $i \in J$ and $s_J \in S_J$.

Let us recall the definition of Coalition-Proof Nash equilibrium. The solution concept is recursively defined as follows.

Definition 1. (i) In a single player game G, $s^* \in S$ is a *Coalition-Proof Nash equilibrium* if and only if s^* maximizes $U_1(s)$.[2]
(ii) Let $n > 1$ and assume that Coalition-Proof Nash equilibrium has been defined for games with fewer than n Players. Then,
(ii.a) For any game G with n players, $s^* \in S$ is *self-enforceable* if, for all $J \in \mathscr{C}_I$, s^*_J is a Coalition-Proof Nash equilibrium in the game G/s^*_{-J}.
(ii.b) For any game G with n players, $s^* \in S$ is a *Coalition-Proof Nash equilibrium* if it is self-enforceable and if there is not another self-enforceable strategy vector $s \in S$ such that $U_i(s) > U_i(s^*)$ for all $i \in I$.

Compared to the concept of Strong Nash equilibrium[3] which implicitly assumes that one-step coalitional deviation would always be made if each member is better off than without deviating, Coalition-Proof Nash equilibrium requires that the deviation be self-enforceable within the coalition. Alternatively, players in the concept of Coalition-Proof Nash equilibrium are farsighted, but the farsightedness is restricted within the coalition. I illustrate this point by offering an equivalent definition of the

[2]The subscript 1 is the index number of the unique player.

[3]A strategy $s^* \in S$ is a Strong Nash equilibrium if for all $J \in \mathscr{C}$, there is no strategy $s_J \in S_J$ such that for every agent $i \in J$, $U_i(s_J, s^*_{-J}) > U_i(s^*)$.

concept.

Definition 2. In a game G, $s^* \in S$ is a *Coalition-Proof Nash equilibrium* if there is no coalition $J \in \mathscr{C}$ such that there exists one strategy vector $s_J \in S_J$ in the game G/s^*_{-J} such that (i) $\tilde{U}_j(s_J) > \tilde{U}_j(s^*_J)$ for all $j \in J$; and (ii) For an arbitrary proper sub-coalition $Q \in \mathscr{C}_J$, s_Q is *Coalition-Proof Nash equilibrium* in game G/s_{-Q}, where $s = (s_J, s^*_{-J})$.

Definition 1 is equivalent to definition 2. If $n = 1$, then $J = \{1\}$ and $-J = \emptyset$. Definition 2 says that there is no $s_1 \in S_1$ such that $\tilde{U}_1(s_1) > \tilde{U}_1(s^*_1)$, which is equivalent to saying that s^*_1 maximizes $\tilde{U}_1(s)$. Note that the proper sub-coalition of a singleton coalition is empty and the condition (ii) of definition 2 holds. If $n > 1$, I have two subcases to consider. First, when $J = I \in \mathscr{C}$, definition 2 is equivalent to saying that there is not another self-enforceable strategy making every player better off. Second, when $J \in \mathscr{C}_I$, definition 2 means that $s^* \in S$ is self-enforceable.

The limited farsightedness of the notion of Coalition-Proof Nash equilibrium can lead to violation of individual rationality. Short-sighted players may agree to a coalitional deviation without realizing that ensuing cross-coalition deviations may hurt them so much that they are worse off than without supporting the first deviation.

I will extend the notion of farsightedness beyond a coalition. First, I define a sequence of within-coalition self-enforceable deviations.

Definition 3. A coalition J's deviation from s^* to $s = (s_J, s^*_{-J})$ is *within-coalition self-enforceable* if (i) $\forall j \in J, U_j(s) > U_j(s^*)$; (ii) for an arbitrary proper sub-coalition $Q \in \mathscr{C}_J$, s_Q is a Coalition-Proof Nash equilibrium in the game G/s_{-Q}.

Let $s^* \underset{J}{\rightarrow} s^0$ denote a within-coalition self-enforceable deviation by coalition J from s^* to s^0. Within-coalition self-enforceable deviations may form a sequence.

Denote $\pi = \{s^{(k-1)} \underset{J_k}{\rightarrow} s^k\}$ as a sequence of within-coalition self-enforceable deviations. Let $l(\pi)$ denote the length of the sequence. The sequence may be infinite. For example, within-coalition self-enforceable deviations may constitute a circle.

Since our objective is to develop a weakening of the Coalition-Proof Nash equilibrium, I only consider subsequent deviations which are within-coalition self-enforceable. A deviation by coalition J may trigger a multiple of sequences of within-coalition self-enforceable deviations. The final outcome should be one of the strategies in the sequences. Since I assume that every player is farsighted, the final outcome should not be the strategy from which some members forming a coalition always benefit by a coordinate deviation. Therefore, the final equilibrium strategies are either the terminal nodes in the sequences or the intermediate nodes from which all the within-coalition self-enforceable deviations are not credible across coalitions. I assume that players are very cautious and that the deviation by J is taken only if all members get a higher payoff at all the possible final strategies than without moving. The cross-coalition self-enforceability is formally defined as follows.

Definition 4. The deviation from s^* to $s^0 = (s_J, s^*_{-J})$ by coalition $J \in \mathscr{C}$ is *cross-coalition self-enforceable*, if it satisfies one of the following two conditions: (i) J is a singleton coalition; or (ii) any subsequent sequence of within-coalition self-enforceable deviations $\pi = \{s^0 \underset{J_1}{\rightarrow} s^1, s^1 \underset{J_2}{\rightarrow} s^2, s^2 \underset{J_3}{\rightarrow} s^3, \ldots\}$ satisfies (a) $l(\pi) < \infty$; and (b) $\forall m \in \{0, 1, \ldots, l(\pi)\}$, if s^m is the terminal node or if all the within-coalition self-enforceable deviations starting from s^m are NOT *cross-coalition self-enforceable*, and if the consecutive deviations $s^{k-1} \underset{J_k}{\rightarrow} s^k (k = 1, \ldots, m)$ are all *cross-coalition self-enforceable*, it is the case that $\forall j \in J$, $U_j(s^m) > U_j(s^*)$.

I emphasize that the definition relies on a recursive argument rather than a cyclical one. The definition of cross-coalition self-enforceability of a deviation recursively depends on the definition for all the subsequent within-coalition deviations. But the terminal deviations in all the sequences are automatically defined. They are always cross-coalition self-enforceable.

Condition (i) in definition (4) stems from the principle of individual rationality. This restriction rules out the non-Nash-equilibrium strategies from the solution concept which I will define based on this definition. Condition (ii.a) is necessary due to the recursive nature of the definition. The definition of cross-coalition self-enforceability of $s^* \underset{J}{\rightarrow} s^0$ depends on the definition of $s^0 \underset{J_1}{\rightarrow} s^1$, which further depends on the definition of $s^1 \underset{J_2}{\rightarrow} s^2$, and so on. If the deviation sequence is infinite, the definition will not be well-defined. [4]

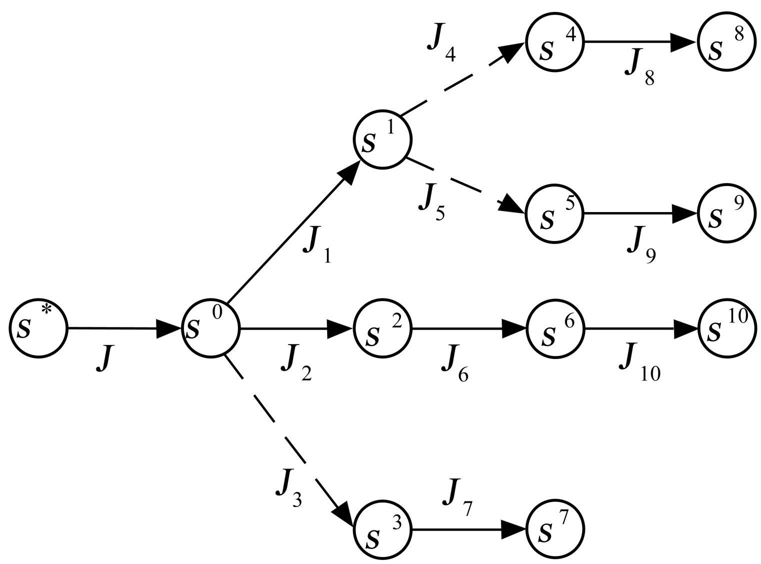

Fig. 4.: Cross-coalition Self-enforceability

The key points of this definition arise from condition (ii.b). I design an illustrative example as in Figure (4) to interpret the ideas. If coalition J moves from s^* to s^0, the deviation can trigger a sequence of within-coalition self-enforceable deviations which

[4]The restriction (ii.a) can be found in the literature, say, Mariotti (1997).

is one of the four sequences: $\pi_1 = \{s^0 \underset{J_1}{\rightarrow} s^1, s^1 \underset{J_4}{\rightarrow} s^4, s^4 \underset{J_8}{\rightarrow} s^8\}$, $\pi_2 = \{s^0 \underset{J_1}{\rightarrow} s^1, s^1 \underset{J_5}{\rightarrow} s^5, s^5 \underset{J_9}{\rightarrow} s^9\}$, $\pi_3 = \{s^0 \underset{J_2}{\rightarrow} s^2, s^2 \underset{J_6}{\rightarrow} s^6, s^6 \underset{J_{10}}{\rightarrow} s^{10}\}$, or $\pi_4 = \{s^0 \underset{J_3}{\rightarrow} s^3, s^3 \underset{J_7}{\rightarrow} s^7\}$. Since the definition of cross-coalition self-enforceability of the deviation $s^* \underset{J}{\rightarrow} s^0$ recursively depends on the definition of all the subsequent deviations, I apply the procedure of backward induction which is common in the subgame perfection. First, the terminal deviations $s^4 \underset{J_8}{\rightarrow} s^8$, $s^5 \underset{J_9}{\rightarrow} s^9$, $s^6 \underset{J_{10}}{\rightarrow} s^{10}$, and $s^3 \underset{J_7}{\rightarrow} s^7$ are cross-coalition self-enforceable. This is because condition (ii) holds as no sequence of within-coalition self-enforceable deviations can be derived from Coalition-Proof Nash equilibria s^8, s^9, s^{10} and s^7. In Figure (4), I use solid arrows to denote cross-coalition self-enforceability, while dashed arrows indicate that the deviations are not cross-coalition self-enforceable.

Along all the sequences, we move one step backward. The deviation $s^1 \underset{J_4}{\rightarrow} s^4$ (similarly, $s^1 \underset{J_5}{\rightarrow} s^5$ and $s^0 \underset{J_3}{\rightarrow} s^3$) is indicated not cross-coalition self-enforceable, which means that $\exists j \in J_4$ such that $U_j(s^8) \leq U_j(s^1)$. In contrast, that the deviation $s^2 \underset{J_6}{\rightarrow} s^6$ is indicated cross-coalition self-enforceable implies that $\forall j \in J_6, U_j(s^{10}) > U_j(s^2)$. The subsequence of deviations $\{s^2 \underset{J_6}{\rightarrow} s^6, s^6 \underset{J_{10}}{\rightarrow} s^{10}\}$ is interesting. If the status quo is s^2, all the members in coalitions J_6 and J_{10} together might have difficulty in reaching an agreement moving from s^2 to s^{10} because some members in coalition J_{10} may prefer the strategy s^2 to s^{10} and oppose the deviation. But a careful analysis reveals that the opposition is futile. All the members in J_6 are better off than keeping status quo no matter whether the outcome is s^6 or s^{10}. The opposers in J_{10} know that the members of coalition J_6 always deviate. Therefore, as rational players, they should move to s^{10} rather than stick to s^2. However, if the deviation $s^2 \underset{J_6}{\rightarrow} s^6$ is not within-coalition self-enforceable, the opposers in J_{10} may be able to deter the deviation from s^2 to s^{10}. This is why I confine our interest of the triggered coalitional moves in within-coalition self-enforceable deviations rather than other types of deviations when defining the cross-coalition self-enforceability.

I call a deviation *strongly self-enforceable* if it is self-enforceable both within the coalition and across the coalitions. Implicitly, the definition (4) assumes that players deviate if and only if the deviation is strongly self-enforceable. This restriction helps avoiding the possible conflict between the coalitional rationality and individual rationality, because nobody will be worse off after this kind of deviations.

Along the sequences π_1, π_2, and π_3, we continue to move one step backward. The deviation $s^0 \underset{J_1}{\rightarrow} s^1$ is cross-coalition self-enforceable. Though some members of J_1 may be worse off if the outcome is s^4, s^5, s^8, or s^9, these outcomes are not reachable as J_4 and J_5 do not deviate from s^1. Thus, a member j in J_1 can safely get the payoff $U_j(s^1)$. That the deviation $s^0 \underset{J_2}{\rightarrow} s^2$ is indicated as cross-coalition self-enforceable implies (i) the deviations $s^2 \underset{J_6}{\rightarrow} s^6$ and $s^6 \underset{J_{10}}{\rightarrow} s^{10}$ are strongly self-enforceable; and (ii) $\forall j \in J_2$, $U_j(s^{10} > U_j(s^0))$. It is possible that some members in J_2 are worse off than keeping the status quo if the outcome is s^6. But, s^6 is not stable. All the members in J_2 will agree to the deviation $s^0 \underset{J_2}{\rightarrow} s^2$ knowing that $J_6 \cup J_{10}$ will accommodate their actions and move from s^2 to s^{10}.

Finally, let us discuss the cross-coalition self-enforceability of the deviation $s^* \underset{J}{\rightarrow} s^0$. Starting from s^0, we move forward step by step along all the solid arrows. We will stop at either s^1 or s^{10}, which are reachable strategies. If, for all $j \in J$, $U_j(s^1) > U_j(s^*)$ and $U_j(s^{10}) > U_j(s^*)$, the deviation $s^* \underset{J}{\rightarrow} s^0$ is cross-coalition self-enforceable. The payoffs derived from the strategies s^3, s^4, s^5, s^7, s^8, and s^9 are irrelevant because the deterrence by coalition J_3, J_4 and J_5 makes these strategies unreachable.

Given the definitions of within-coalition self-enforceability and cross-coalition self-enforceability, I can define a new solution concept, Weak Coalition equilibrium, which is a weakening of the notion of Coalition-Proof Nash equilibrium.

Definition 5. In a game G, $s^* \in S$ is a *Weak Coalition equilibrium* if there are no strongly self-enforceable deviations.

The conciseness of the definition 5 enables us to associate it with other three concepts: Nash equilibrium, Coalition-Proof Nash equilibrium, and Strong Nash equilibrium. In a Nash equilibrium, there is not an individually conceivable deviation; in a Weak Coalition equilibrium, there is not a strongly self-enforceable deviation; in a Coalition-Proof Nash equilibrium, there is not a within-coalition self-enforceable deviation; in a Strong Nash equilibrium, there is not any conceivable coalition deviation. The equilibrium requirements of the solution concepts, Nash equilibrium, Weak Coalition equilibrium, Coalition-Proof Nash equilibrium, and Strong Nash equilibrium, become increasingly stronger because {d | d is an individually conceivable deviation}⊆{d | d is a strongly self-enforceable deviation}⊆{d | d is a within-coalition self-enforceable deviation}⊆{d | d is a conceivable coalition deviation}. Therefore, in a game, we have {e | e is a Nash equilibrium}⊇ {e | e is a Weak Coalition equilibrium}⊇{e | e is a Coalition-Proof Nash equilibrium}⊇{e | e is a Strong Nash equilibrium}.

A Weak Coalition equilibrium always exists, if a Nash equilibrium exists in the game. The arguments are as follows. Starting from an arbitrary Nash equilibrium s^*, I construct sequences of with-coalition self-enforceable deviations. Two cases need to be considered. First, if no strongly self-enforceable deviation is derived from s^*, then s^* is a Weak Coalition equilibrium. Second, if there is one strongly self-enforceable deviation from s^* to s, then the terminal node of every subsequent sequence of within-coalition self-enforceable deviations is a Coalition-Proof Nash equilibrium, hence is a Weak Coalition equilibrium. In sum, at least a Weak Coalition equilibrium exists in the game.

Definition 6. In a game G, $s^* \in S$ is an *Efficient Weak Coalition equilibrium* if s^* is a Weak Coalition equilibrium and there is not another Weak Coalition equilibrium s^{**} such that, $\exists J \in \mathscr{C}$, (i) $\forall j \in J, U_j(s^{**}) > U_j(s^*)$; and (ii) $\forall j \in -J, U_j(s^{**}) \geq U_j(s^*)$.

The notion of Efficient Weak Coalition equilibrium selects Pareto-efficient strategies from the set of Weak Coalition Equilibria.

4.3. Examples

In this section, I discuss a series of examples in order for the reader to gain more insights into the notion of Weak Coalition equilibrium. I associate the examples with the concept Coalition-Proof Nash equilibrium to justify the definition of Weak Coalition Equilibrium.

Example 1. The Prisoner's Dilemma

Table XX.: The Prisoner's Dilemma

	Don't Confess	Confess
Don't Confess	-2, -2	-10, -1
Confess	-1, -10	-5, -5

Table (XX) presents the well-known game of the Prisoner's Dilemma. The strategy (Confess, Confess) is the unique Weak Coalition equilibrium in this game. Though both row and column players are better off by the deviation (Confess, Confess) $\underset{\textit{All Players}}{\longrightarrow}$ (Don't Confess, Don't Confess), it is not within-coalition self-enforceable. Additionally, from the status quo strategy (Don't Confess, Don't Confess), we can derive two sequences of strongly self-enforceable deviations: {(Don't Confess, Don't Confess) $\underset{\textit{Row Player}}{\longrightarrow}$ (Confess, Don't Confess), (Confess, Don't Confess) $\underset{\textit{Column Player}}{\longrightarrow}$

(Confess, Confess)} and {(Don't Confess, Don't Confess) $\underset{Column\ Player}{\longrightarrow}$ (Don't Confess, Confess), (Don't Confess, Confess) $\underset{Row\ Player}{\longrightarrow}$ (Confess, Confess)}.

Example 2. Ambrus' Example

Table (XIX) in the introduction is modified from Ambrus's example. The strategies (A_1, B_1, C_1) and (A_2, B_2, C_2) are two Weak Coalition equilibria. The deviation $(A_1, B_1, C_1) \underset{A,B}{\rightarrow} (A_2, B_2, C_1)$ is not cross-coalition self-enforceable because it may trigger the deviation $(A_2, B_2, C_1) \underset{C}{\rightarrow} (A_2, B_2, C_2)$. The notion of Weak Coalition equilibrium can not rule out the strategy (A_2, B_2, C_2). Given the strategy (A_2, B_2, C_2) as the status quo, the deviation $(A_2, B_2, C_2) \underset{A,B,C}{\rightarrow} (A_1, B_1, C_1)$ is in the interest of players A and B, whereas player C may deter this deviation worrying about a further deviation by players A and B. With an efficiency restriction, the notion of Efficient Weak Coalition equilibrium predicts (A_1, B_1, C_1).

Example 3. Coalitional Betrayal and Coalitional Antagonization

Table (XXI) shows a three-player game in which players are assumed to play pure strategies. The strategies (A1, B1, C1), (A1, B2, C2) and (A2, B2, C1) are three Nash equilibria. Given the strategy (A1, B1, C1), players B and C have a within-coalition self-enforceable deviation from (B1, C1) to (B2, C2); given (A1, B2, C2), players A and C have a within-coalition self-enforceable deviation from (A1, C2) to (A2, C1); given (A2, B2, C1), players A and B have a within-coalition self-enforceable deviation from (A2, B2) to (A1, B1). Within-coalition self-enforceable deviations constitute a circle of cross-coalition betrayals. In this game, Coalition-Proof Nash equilibrium does not exist because there is a within-coalition self-enforceable deviation for all Nash equilibria. In fact, an individual player is not necessarily agree to the first betrayal deviation if he/she anticipates there might be further deviations which punish him/her. Therefore, the concept of Coalition-Proof Nash equilibrium violates the basic principle of individual rationality. Any coalitional solution concept should treat (A1, B1, C1), (A1, B2, C2) and (A2, B2, C1) equally, in view of their symmetric position. All of them are Weak Coalition equilibria.

Table XXI.: An Example of Coalitional Betrayal

C1			C2		
	B1	B2		B1	B2
A1	3, 2, 1	0, 0, 0	A1	0, 0, 0	1, 3, 2
A2	.5, .5, .5	2, 1, 3	A2	0, 0, 0	0, 0, 0

Table (XXII) shows a four-player game, in which players are assumed to play pure strategies. The strategies (A1, B1, C1, D1), (A2, B2, C1, D1), (A1, B1, C2, D2) and (A2, B2, C2, D2) are four Nash equilibria. Given strategy (A1, B1, C1, D1), players C and D have a within-coalition self-enforceable deviation from (C1, D1) to (C2, D2); given strategy (A1, B1, C2, D2), players A and B have a within-coalition self-enforceable deviation from (A1, B1) to (A2, B2); given strategy (A2, B2, C2, D2), players C and D have a within-coalition self-enforceable deviation from (C2, D2) to (C1, D1); given strategy (A2, B2, C1, D1), players A and B have a within-coalition self-enforceable deviation from (A2, B2) to (A1, B1). In this game, Coalition-Proof Nash equilibrium does not exist because there is a within-coalition self-enforceable deviation for all Nash equilibria. Any coalitional solution concept should treat (A1, B1, C1, D1), (A2, B2, C1, D1), (A1, B1, C2, D2) and (A2, B2, C2, D2) equally, in view of their symmetric position. All of them are Weak Coalition equilibria.

The games of coalitional betrayal and coalitional antagonization are two representative examples of a circle of deviations. If the strategy space is infinite, it is possible that distinct deviations form an infinite sequence. It is uncertain which strategy in the circle or the infinite sequence will happen; we are not able to determine wether the first deviation will be carried out or not. Now that we aim at defining a weak criterion, we impose restriction (ii.a) in the definition (4) to rule out these cases.

Table XXII.: An Example of Coalitional Antagonization

		D1			D2	
		B1	B2		B1	B2
C1	A1	2, 2, 1, 1	0, 0, 0, 0	A1	.5, 0, 0, 0	0, 0, 0, 0
	A2	0, 0, 0, 0	1, 1, 2, 2	A2	0, 0, 0, 0	.5, 0, 0, 0
		B1	B2		B1	B2
C2	A1	.5, 0, 0, 0	0, 0, 0, 0	A1	1, 1, 2, 2	0, 0, 0, 0
	A2	0, 0, 0, 0	.5, 0, 0, 0	A2	0, 0, 0, 0	2, 2, 1, 1

Example 4. Solution Refinement

Table XXIII.: An Example of Solution Refinement

		D1			D2	
		B1	B2		B1	B2
C1	A1	2, 2, 3, 3	0, 0, 0, 0	A1	1.5, 1.5, 4, 4	0, 0, 0, 0
	A2	0, 0, 0, 0	1, 1, 7, 7	A2	0, 0, 0, 0	0, 0, 0, 0
		B1	B2		B1	B2
C2	A1	0, 0, 0, 0	0, 0, 0, 0	A1	.5, .5, .7, .7	0, 0, 0, 0
	A2	0, 0, 0, 0	0, 0, 0, 0	A2	0, 0, 0, 0	2, 2, 3, 3

The notion of Weak Coalition equilibrium is a good refinement of Nash set. For example, in Table (XXIII), (A2, B2, C1, D1) is a Nash equilibrium. If the arbiter prescribes this strategy to the players, then players A and B jointly deviate from (A2, B2) to (A1, B1). This deviation is strongly self-enforceable. After this deviation, A and B always get a payoff value larger than the status quo value 1. If D would not

Table XXIV.: An Example without Solution Refinement

		D1			D2	
		B1	B2		B1	B2
C1	A1	2, 2, 3, 3	0, 0, 0, 0	A1	.6, 1.5, 4, 4	0, 0, 0, 0
	A2	0, 0, 0, 0	1, 1, 7, 7	A2	0, 0, 0, 0	0, 0, 0, 0
		B1	B2		B1	B2
C2	A1	0, 0, 0, 0	0, 0, 0, 0	A1	.5, .5, .7, .7	0, 0, 0, 0
	A2	0, 0, 0, 0	0, 0, 0, 0	A2	0, 0, 0, 0	2, 2, 3, 3

deviate, A and B get a payoff value 2. Even if D deviates from (D1) to (D2), the payoff value is 1.5, which is also larger than 1. Therefore, (A2, B2, C1, D1) is not a Weak Coalition equilibrium. However, in Table (XXIV), (A2, B2, C1, D1) is a Weak Coalition equilibrium. If A and B move first, D deviates to D2. Then the payoff of A will decrease to 0.6 in the end. Knowing this, player A would be irrational to agree to deviate from (A2, B2) to (A1, B1), even though this deviation is within-coalition self-enforceable.

4.4. Summary

I recursively define a concept Weak Coalition equilibrium in which coalitional rationality is reconciled with the fundamental principle of individual rationality. Starting from an assignment prescribed by an arbiter, the game players who can freely discuss their strategies but cannot make binding commitments agree with a coalition to coordinate and deviate only when the deviation is strongly self-enforceable, i.e., credible not only within but also across coalitions. Weak Coalition equilibrium is a generalization of Coalition-Proof Nash equilibrium. All the Coalition-Proof Nash equilibria are contained in the set of Weak Coalition equilibria which is further a subset of Nash equilibria. A Weak Coalition equilibrium exists if a Nash equilibrium exists.

CHAPTER V

SUMMARY

Three essays show that microfoundations are important not only in the macroeconomic analysis, but also in the analysis of welfare, finance and coalition etc. Economists have a tendency to oversimplify the economic structures. On the one hand, the simplification helps us to capture the key logic underlying complex economic phenomena. On the other hand, economists may neglect some pivotal microeconomic factors. Many puzzles in the literature may step from the weak microfoundations in the economic models.

With stronger microfoundations of economic models, three essays contribute to the literature in different respects. In the first essay, I propose a brand-new methodology to nonparametrically estimate the structural labor supply and exact welfare change and deadweight loss due to tax reforms. Different from Hausman, I specify an ordinary indirect utility function. Compared to the specification of labor supply function in the Hausman's framework, the indirect utility function facilitates the calculation under non-convex piecewise-linear budget sets. My method is able to address many problems such as individual heterogeneity, nonconvex budget sets, labor nonparticipation, and measurement errors. This method is a good generalization and substitute of Vartia's method. In the second essay, I contribute to the literature by resolving the equity premium puzzle, risk-free rate puzzle, and capital structure puzzle. I point out that the microfoundations in a representative-agent, consumption-based asset pricing model is too weak to correctly price the assets. To generate the historical data of equity premia and risk-free rates, an asset pricing model should incorporate the hump-shaped life-cycle consumption curve. In addition, this essay presents a new solution to the determination of capital structure. Different from Modigliani and

Miller, the capital structure in my model is determined by investors. The third essay proposes a new coalition solution concept, Weak Coalition Equilibrium, in which the coalitional deviations are in line with the fundamental principles of individual rationality.

REFERENCES

ABEL, A. B. (1990): "Asset Prices under Habit Formation and Catching up with the Jones," *American Economic Review*, 80, 38-42.

ABEL, A. B. (1999): "Risk Premia and Term Premia in General Equilibrium," *Journal of Monetary Economics*, 43, 3-33.

AIYAGARI, S. R., AND M. GERTLER (1991): "Asset Returns with Transactions Costs and Uninsured Individual Risk," *Journal of Monetary Economics*, 27, 311-331.

ALVAREZ, F., AND U. JERMANN (2002): "Efficiency, Equilibrium, and Asset Pricing with Risk of Default," *Econometrica*, 68, 775-797.

AMBRUS, A. (2006): "Coalitional Rationalizability," *Quarter Joural of Economics*, 121, 903-929.

ATTANASIO, O. P., J. BANKS, AND S. TANNER (2002): "Asset Holding and Consumption Volatility," *Journal of Political Economy*, 110, 771-792.

AUMANN, R. J. (1959): *Acceptable Points in General Cooperative n-Person Games*. Princeton, NJ: Princeton University Press.

BANSAL, R., AND J. W. COLEMAN (1996): "A Monetary Explanation of the Equity Premium, Term Premium and Risk-Free Rate Puzzles," *Journal of Political Economy*, 104, 1135-1171.

BARRO, R. J. (1993): *Macroeconomics*. Cambridge, MA: MIT Press.

BARRO, R. J. (2006): "Rare Disasters and Asset Markets in the Twentieth Century," *Quarterly Journal of Economics*, 121, 823-866.

BARRO, R. J. (2009): "Rare Disasters, Asset Prices, and Welfare Costs," *American Economic Review*, 99, 243-264.

BASAK, S., AND D. CUOCO (1998): "An Equilibrium Model with Restricted Stock Market Participation," *Review of Financial Studies*, 11, 309-341.

BENARTZI, S., AND R. H. THALER (1995): "Myopic Loss Aversion and the Equity Premium Puzzle," *Quarterly Journal of Economics*, 110, 73-92.

BENNINGA, S., AND A. PROTOPAPADAKIS (1990): "Leverage, Time Preference, and the 'Equity Premium Puzzle'," *Journal of Monetary Economics*, 25, 49-58.

BERNHEIM, B. D., B. PELEG, and M. D. WHINSTON (1987): "Coalition-proof Nash Equilibrium I. Concepts," *Journal of Economic Theory*, 42, 1-12.

BLOMQUIST, S. (1988): "Nonlinear Taxes and Labor Supply," *European Economic Review,* 32, 1213-1226.

BLOMQUIST, S. (1996): "Estimation Methods for Male Labor Supply Functions–How to Take Account of Nonlinear Taxes," *Journal of Econometrics*, 70, 383-405.

BLOMQUIST, S., AND W. K. NEWEY (2002), "Nonparametric Estimation with Nonlinear Budget Sets," *Econometrica*, 70, 2455-2480.

BLUNDELL, R., C. Meghir, E. Symons, and I. Walker (1988): "Labor Supply Specification and the Evaluation of Tax Reforms," *Journal of Public Economics*, 36, 23-52.

BOLDRIN, M., L. J. CHRISTIANO, AND J. D. M. FISHER (2001): "Habit Persistence, Asset Returns, and the Business Cycle," *American Economic Review*, 91, 149-166.

BRAV, A., G. M. CONSTANTINIDES, AND C. C. GECZY (2002): “Asseting Pricing with Heterogeneous Consumers and Limited Participation: Empirical Evidence,” *Journal of Political Economy*, 110, 793-824.

BREEDEN, D. T. (1979): “An Intertemporal Asset Pricing Model with Stochastic Consumption and Investment Opportunities,” *Journal of Financial Economics*, 7, 265-296.

BROWN B. W., AND M. B. WALKER (1989): “The Random Utility Hypothesis and Inference in Demand Systems,” *Econometrica*, 57, 815-829.

BURTLESS G., AND J. A. HAUSMAN (1978): “The Effect of Taxation on Labor Supply: Evaluating the Gary Negative Income Tax Experiment,” *Journal of Political Economy*, 86, 1103-1130.

CAMPBELL, J. Y., AND H. H. COCHRANE (1999) “By Force of Habit: A consumption-Based Explanation of Aggregate Stock Market Behavior,” *Journal of Political Economy*, 107, 205-251.

CHAN, Y. L., AND L. KOGAN (2002): “Catching Up with the Joneses: Heterogeneous Preferences and the Dynamics of Asset Prices,” *Journal of Political Economy*, 110, 1255-1285.

CHWE, M. S. (1994): “Farsighted Coalitional Stability,” *Journal of Economic Theory*, 63, 299-325.

COGAN, J. F. (1981): “Fixed Costs and Labor Supply,” *Econometrica*, 49, 945-963.

COGLEY, T. (2002): “Idiosyncratic Risk and the Equity Premium: Evidence from the Consumer Expenditure Survey,” *Journal of Political Economy*, 49, 309-334.

CONSTANTINIDES, G. M. (1990): "Habit Formation: A Resolution of the Equity Premium Puzzle," *Journal of Political Economy*, 98, 519-543.

CONSTANTINIDES, G. M., J. B. DONALDSON, AND R. MEHRA (2002): "Junior Can't Borrow: A New Perspective on the Equity Premium Puzzle," *Quarterly Journal of Economics*, 117, 269-296.

CONSTANTINIDES, G. M., AND D. DUFFIE (1996): "Asset Pricing with Heterogeneous Consumers," *Journal of Political Economy*, 104, 219-240.

COX, J. C., J. E. ENGERSOLL JR., AND S. A. ROSS (1985): "An Intertemporal General Equilibrium Model of Asset Prices," *Econometrica*, 53, 363-384.

DANTHINE, J. P., J. B. DONALDSON, AND R. MEHRA (1992): "The Equity Premium and The Allocation of Income Risk," *Journal of Economic Dynamics and Control*, 16, 509-532.

DETEMPLE, J. B., AND A. SERRAT (2003): "Dynamic Equilibrium with Liquidity Constraints," *Review of Financial Studies*, 16, 597-629.

EBRAHIM, M. S., AND I. MATHUR (2001): "Investor Heterogeneity, Market Segmentation, Leverage and the Equity Premium Puzzle," *Journal of Banking and Finance*, 25, 1897-1919.

EPSTEIN, L. G., and S. E. ZIN (1989): "Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework," *Econometrica*, 57, 937-969.

EPSTEIN, L. G., AND S. E. ZIN (1991): "Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: An Empirical Analysis," *Journal of Political Economy*, 99 , 263-286.

FERSON, W. E., AND G. M. CONSTANTINIDES (1991): "Habit Persistence and Durability in Aggregate Consumption: Empirical Tests," *Journal of Financial Economics*, 29, 199-240.

FULLERTON D., AND L. GAN (2004): "A Simulation-Based Welfare Loss Calculation for Labor Taxes with Piecewise-linear Budgets," *Journal of Public Economics*, 88, 2339-2359.

GOLLIN, D. (2002): "Getting Income Shares Right," *Journal of Political Economy*, 110, 458-474.

GROSSMAN S. J., A. MELINO, AND R. J. SHILLER (1987): "Estimating the Continuous-Time Consumption-Based Asset-Pricing Model," *Journal of Business & Economic Statistics*, 5, 315-327.

HANSEN, G.D. (1993): "The Cyclical and Secular Behavior of the Labour Input: Comparing Efficiency Units and Hours Worked," *Journal of Applied Econometrics*, 8, 71-80.

HANSEN, L. P., AND R. JAGANNATHAN (1991) "Implications of Security Market Data for Models of Dynamic Economies," *Journal of Political Economy*, 99, 225-262.

HANSEN, L. P., AND J. S. KENNETH (1982): "Generalized Instrumental Variables Estimation of Nonlinear Rational Expectations Models," *Econometrica*, 50, 1269-1286.

HAUSMAN J. A. (1981a): "Exact Consumer's Surplus and Deadweight Loss," *American Economic Review*, 71, 662-676.

HAUSMAN, J. A. (1981b): "Labor Supply,"in *How Taxes Affect Economic Behavior,* ed. by H. J. Aaron and J. A. Pechman. Washington: The Brookings Institution.

HAUSMAN, J. A. (1985): "Taxes and Labor Supply,"in *Handbook of Public Economics*, Vol. 1., ed. by A. Auerbach and M. Feldstein. Amsterdam Netherland: North-Holland.

HAUSMAN, J. A., AND W. K. NEWEY (1995): "Nonparametric Estimation of Exact Consumers Surplus and Deadweight Loss," *Econometrica*, 63, 1445-1476.

HE, H., AND D. M. MODEST (1995): "Market Frictions and Consumption-Based Asset Pricing,"*Journal of Political Economy*, 103, 94-117.

HEATON, J., AND D. J. LUCAS (1996): "Evaluating the Effects of Incomplete Markets on Risk Sharing and Asset Pricing,"*Journal of Political Economy*, 104, 443-487.

HEATON, J., AND D. J. LUCAS (1997): "Market Frictions, Savings Behavior and Portfolio Choice,"*Journal of Macroeconomic Dynamics*, 1, 76-101.

HEATON, J., AND D. J. LUCAS (2000): "Portfolio Choice and Asset Prices: The Importance of Entrepreneurial Risk,"*Journal of Finance*, 55, 1163-1198.

HECKMAN J. J. (1979): "Sample Selection Bias as a Specification Error," *Econometrica*, 47, 153-161.

JACOBS, K. (1999): "Incomeplete Markets and Security Prices: Do Asset-Pricing Puzzles Result from Aggregation Problems?,"*Journal of Finance*, 54, 123-163.

JERMANN, U. J. (1998): "Asset Pricing in Production Economies,"*Journal of Monetary Economics*, 41, 257-275.

KOCHERLAKOTA, N. R. (1996): "The Equity Premium: It's Still a Puzzle," *Journal of Economic Literature*, 34, 42-71.

KRUEGER, A. B. (1999): "Measuring Labor's Share," *American Economic Review*, 89, 45-51.

KRUSELL, P., AND A. A. SMITH (1998): "Income and Wealth Heterogeneity in the Macroeconomy," *Journal of Political Economy*, 106, 867-896.

KUMAR A. (2008): "Labor Supply, Deadweight Loss and Tax Reform Act of 1986: A Nonparametric Evaluation Using Panel Data," *Journal of Public Economics*, 92, 236-253.

LIANG C. (2009): "Nonparametric Structural Estimation of Labor Supply in the Presence of Censoring," Unpublished Manuscript, Department of Economics, Uppsala University.

LUCAS, D. J. (1994): "Asset Prices with Undiversifiable Risk and Short Sales Constraints: Deepening the Equity Premium Puzzle," *Journal of Monetary Economics*, 34, 325-341.

LUCAS JR., R. E. (1978): "Asset Prices in An Exchange Economy," *Econometrica*, 46, 1429-1445.

LUTTMER, E. G. J. (1996): "Asset Prices in Economies with Frictions," *Econometrica*, 64, 1439-1467.

MANKIW, N. G. (1986): "The Equity Premium and the Concentration of Aggregate Shocks," *Journal of Financial Economics*, 17, 211-219.

MANKIW, N. G., AND S. P. ZELDES (1991): "The Consumption of Stockholders and Nonstockholders," *Journal of Financial Economics*, 29, 97-112.

MARCET, A., AND K. J. SINGLETON (1999): "Equilibrium Asset Prices and Savings of Heterogeneous Agents in the Presence of Portfolio Constraints," *Macroeconomic Dynamics*, 3, 243-277.

MARIOTTI, M. (1997): "A Model of Agreements in a Strategic Form Games," *Journal of Economic Theory* 74, 196-217.

MASULIS, R. W. (1988): *The Debt-Equity Choice*. Cambridge, MA: Ballinger Press.

MCGRATTAN, E. R., AND E. C. PRESCOTT (2000): "Is the Stock Market Overvalued?" *Federal Reserve Bank of Minneapolis Quarterly Review*, 24, 20-40.

MCGRATTAN, E. R., AND E. C. PRESCOTT (2001): "Taxes, Regulations, and Asset Prices," Working Paper, Federal Reserve Bank of Minneapolis.

MEHRA, R., AND E. C. PRESCOTT (1985): "The Equity Premium: A Puzzle," *Journal of Monetary Economics*, 15, 145-161.

MERTON, R. C. (1971): "Optimum Consumption and Portfolio Rules in a Continuous-Time Model," *Journal of Economic Theory*, 3, 373-413.

MERTON, R. C. (1973): "An Intertemporal Capital Asset Pricing Model," *Econometrica*, 41, 867-887.

MODIGLIANI, F., AND M. H. MILLER (1958): "The Cost of Capital, Corporation Finance and the Theory of Investment," *American Economic Review*, 48, 261-297.

MYERS, S. C. (1984): "The Capital Structure Puzzle," *Journal of Finance*, 39, 575-592.

NASH, J. (1951): "Non-Cooperative Games," *Annals of Mathematics*, 54, 286-295.

RAJAN, R. G., AND L. ZINGALES (1995): "What Do We Know About Capital Structure? Some Evidence from International Data," *Journal of Finance*, 50, 1421-1460.

RIETZ, T. A. (1988): "The Equity Premium: A Solution," *Journal of Monetary Economics*, 22, 117-131.

RÍOS-RULL, J. (1996): "Life-Cycle Economies and Aggregate Fluctuations," *Review of Economic Studies*, 63, 465-489.

RUBINSTEIN, M. (1976): "The Valuation of Uncertain Income Streams and the Pricing of Options," *Bell Journal of Economics*, 7, 407-425.

SIEGEL, J. J. (1998): *Stocks for the Long Run: The Definitive Guide to Financial Market Returns and Long-Term Investment Strategies,* 2nd ed. New York: McGraw-Hill.

SOEST A. V., M. DAS AND X. GONG (2002): "A Structural Labour Supply Model with Flexibile Preferences," *Journal of Econometrics*, 107, 354-374.

STORESLETTEN, K., C. I. TELMER, AND A. YARON (2007): "Asset Pricing with Idiosyncratic Risk and Overlapping Generations," *Review of Economic Dynamics*, 10, 519-548.

TELMER, C. I (1993): "Asset Pricing Puzzles and Incomplete Markets," *Journal of Finance*, 48, 1803-1832.

TRIEST, R. K. (1990): "The Effect of Income Taxation on Labor Supply in the Unites States," *Journal of Human Resources*, 25, 491-516.

VARTIA Y. (1983): "Efficient Methods of Measuring Welfare Change and Compensated Income in Terms of Ordinary Demand Functions," *Econometrica*, 51, 79-98.

VISSING-JORGENSEN, A. (2002): "Limited Stock Market Participation and the Elasticity of Intertemporal Substitution," *Journal of Political Economy*, 110, 825-853.

WEIL, P. (1989): "The Equity Premium Puzzle and the Risk-Free Rate Puzzle," *Journal of Monetary Economics*, 24, 401-421.

ZHOU, C. (1999): "Informational Asymmetry and Market Imperfections: Another Solution to the Equity Premium Puzzle," *Journal of Financial and Quantitative Analysis*, 34, 445-464.

APPENDIX A

VARTIA'S METHOD AND ITS APPLICATION TO PIECEWISE-LINEAR BUDGET SETS

A.1. A Framework for Calculation of Welfare Change under Piecewise-Linear Budgets Sets

Let T denote a tax system. For an individual with a before-tax wage rate w^*, non-labor income Y^n and a "virtual" lump-sum subsidy S, we express his/her budget set as $B(w^*, T, Y^n, S)$. The "virtual" subsidy S is only conceptually meaningful. The individual's true budget is $B(w^*, T; Y^n, S = 0)$. The subsidy shifts the position of the budget set, but does not change the after-tax wage rate at each budget segment.

Define the government expenditure function as:

$$e(w^*, T, Y^n; u) = inf\{S \mid \exists\,(h, g) \in B(w^*, T; Y^n, S) \; and \; u(h, g) \geq u\}. \tag{A.1}$$

The expenditure $e(w^*, T, Y^n; u)$ is the minimum subsidy from the government to help the individual to attain a given utility level u. It is possible to for e to take a negative value. For notational convenience, I suppress w^* and Y^n in the expenditure function and the budget set hereafter.

Consider a case where a change of tax system from T^0 to T^1 leads to a change of the individual's utility level from u^0 to u^1 ($u^i = max\{u(h, g) \mid (h, g) \in B(T^i, S = 0)\}$). Obviously, we have $e(T^0, u^0) = 0$ and $e(T^1, u^1) = 0$. The compensating variation (CV) and equivalent variation (EV) can be formally defined as:

$$CV = e(T^1; u^0) - e(T^1; u^1) = e(T^1; u^0), \tag{A.2}$$

$$EV = e(T^0; u^0) - e(T^0; u^1) = -e(T^0; u^1). \tag{A.3}$$

We will only discuss CV. The EV can be calculated similarly.

The key issue in calculating the CV is to obtain the touching point of the budget set $B[T^1; e(T^1; u^0)]$ and the indifference curve with utility u^0. Suppose the touching point is $(\bar{h}, \bar{g})$, then

$$CV = \bar{g} - y^a(\bar{h}, T^1), \tag{A.4}$$

where the second term denotes the after-tax income under T^1 tax system if the actual working hour is $\bar{h}$.

Vartia's numerical algorithm depends on four types of information: (1) the market labor supply function $\pi(w, y)$; (2) the choice point (h^0, g^0) on the budget set $B[T^0, 0]$; (3) the net wage rate w^0 and the virtual income y^0 at the choice point (h^0, g^0) before the tax reform; (4) the net wage rate $\bar{w}$ at the touching point between the compensated budget set $B[T^1, e(T^1; u(h^0, g^0))]$ and the indifference curve with utility $u(h^0, g^0)$. Equipped with these four types of information, we can work out the compensated virtual income $\bar{y}$ at the touching point by numerically solving the differential equation (2.29). Thus, the compensated choice of labor supply will be $\bar{h} = \pi(\bar{w}, \bar{y})$ and the corresponding consumption of the numeraire good will be $\bar{g} = \bar{w}\bar{h} + \bar{y}$.

A.2. Application of Vartia's Method to Convex Budget Sets

We assume that the budget set is convex and the preference is strictly convex. In this case, the information of a labor supply function $\pi(w, y)$ will be shown to be adequate to measure the welfare change. That is, the direct or indirect utility function is not needed to calculate the welfare change if we apply Vartia's method.

It is relatively easy to obtain the initial point (h^0, g^0) and the corresponding w^0 and y^0. The labor supply choice h^0 on this convex budget set is unique, which can be determined by employing the necessary and sufficient condition (2.6). If h^0 falls on a segment of the budget set $B[T^0; 0]$, the net wage rate w^0 and the virtual income y^0 correspond to the slope and the intercept of the segment (extended to the zero-hour axis). If h^0 falls on a kink point, w^0 and y^0 can be calculated according to the methodology in section (2.2.3).

We need to obtain the wage rate $\bar{w}$ at the touching point. At this touching point, the tangent line to the indifference curve (with utility $u(h^0, g^0)$) separates the curve from the budget set $B[T^1, e(T^1; u^0)]$. If the touching point falls on a budget segment, say, the j-th segment, then $\bar{w} = w_j^*$. If it falls on a kink, say, the j-th kink point, then $\bar{w}$ will be in the interval $[w_{j+1}^*, w_j^*]$. As the budget set is convex, its slope is decreasing if we move to the left (increase the labor supply) along the budget set and increasing if we move to the right (decrease the labor supply). Similarly, the slope of the budget set is increasing if we move to the left along the indifference curve and deceasing if we move to the right.

To characterize the slope of the budget set $B(T^1, 0)$, we define a correspondence $\omega(\cdot, B(T^1, 0)) : R^+ \mapsto I$ as follows:

$$\omega(h, B(T^1, 0)) = \begin{cases} [w_j^*, w_j^*], & \text{if } H_{j-1} < h < H_j, \forall j; \\ [w_{j+1}^*, w_j^*], & \text{if } h = H_j, \forall j, \end{cases} \tag{A.5}$$

where I is a set of intervals. The correspondence is monotonically decreasing, i.e.

$$inf\left[\omega(h_1, B(T^1, 0))\right] \geq sup\left[\omega(h_2, B(T^1, 0))\right],$$

whenever $h_1 \leq h_2$. At the unique touching point, we have $\bar{w} \in \omega(\pi(\bar{w}, \bar{y}), B(T^1, 0))$. But for any slope of the indifference curve w different from $\bar{w}$ and its corresponding compensated virtual income y, it is the case that $w \notin \omega(h(w, y), B(T^1, 0))$.

The correspondence is useful in determining the direction of the movement along the indifference curve. If $w^0 > sup\,[\omega(h^0, B(T^1, 0))]$, the touching point should lie to the right of the initial point (h^0, g^0) and the desired wage rate $\bar{w}$ should be larger than $sup\,[\omega(h^0, B(T^1, 0))]$. Similarly, the touching point should lie to the left of the initial point if $w^0 < inf\,[\omega(h^0, B(T^1, 0))]$ and the desired wage rate $\bar{w}$ should be less than $inf\,[\omega(h^0, B(T^1, 0))]$.

We numerically solve the differential equation (2.29). If the touching point is to the right of the initial point, we appropriately choose a differential and monotonically decreasing function $w(t)$ such that $w(0) = w^0$ and $w(1) = sup\,[\omega(h^0, B(T^1, 0))]$. Suppose that the numerical method for the differential equation takes the following recursive formulae[1]:

$$y(t_k) = \varphi\left[y(t_{k-1}), w(t_k), w(t_{k-1})\right], \tag{A.6}$$

where $t_k = k/K$, $k = 0, 1, \ldots, K$. The iteration will be carried out till

$$t_k = min\{t_k \mid w(t_k) \leq sup\left[\omega(\pi(w(t_k), y(t_k)), B(T^1, 0))\right], k = 0, 1, \ldots, K\}.$$

The wage rate $w(t_k)$ will be a good approximation of $\bar{w}$ if K is large enough. If the touching point is to the left of the initial point, we choose a differential and monotonically increasing function $w(t)$ such that $w(0) = w^0$ and $w(1) = inf\,[\omega(h^0, B(T^1, 0))]$. The stopping criteria will be

$$t_k = min\{t_k \mid w(t_k) \geq inf\left[\omega(\pi(w(t_k), y(t_k)), B(T^1, 0))\right], k = 0, 1, \ldots, K\}.$$

A.3. Application of Vartia's Method to nonconvex Budget Sets

If the budget set is nonconvex, we are still able to estimate the exact welfare change use only the information of the market labor supply $\pi(w, y)$, without invoking the

[1] See Vartia (1983) and Hausman and Newey (1995) for other numerical methods for solving the differential equation.

direct or indirect utility function. However, the estimation procedure is much more complicated and computationally demanding.

The choice point (h^0, g^0) on the budget set $B[T^0; 0]$ can be solved in two steps. First, we apply the necessary condition in (2.6) to search for the local optimum bundles. Second, if there are more than one bundle satisfying (2.6), we single out the best by applying Vartia's method. We pairwise compare bundles to eliminate the less preferred. Staring from one bundle, we move along the indifference curve. If the second choice is above the indifference curve, then the first bundle will be discarded, and vice versa.[2]

We divide the nonconvex budget set into a multiple of convex sets. Thus, the methodology applied to the convex budget sets can be utilized here. To illustrate the idea, we consider a case in which there is only one concave point in the budget set. It is straightforward to extend the method to the cases with a multiple of concave points.

Assume that the k-th kink point is concave. We divide the budget into two halves. For the left part, we define the slope correspondence to be

$$\omega(h, B(T^1, 0)) = \begin{cases} [w_j^*, w_j^*], & \text{if } H_{j-1} < h < H_j, \forall j > k; \\ [w_{j+1}^*, w_j^*], & \text{if } h = H_j, \forall j > k; \\ [w_{k+1}^*, M_1], & \text{if } h = H_k; \\ [M_1, M_1], & \text{if } h < H_k, \end{cases} \tag{A.7}$$

where $M_1 = max\{w_j^* \mid 1 \leq j \leq k+1\}$. Similarly, for the right part, we define the

[2]Implicitly, we assume that the preference is monotonically increasing with respect to the numeraire.

slope correspondence to be

$$\omega(h, B(T^1, 0)) = \begin{cases} [w_j^*, w_j^*], & \text{if } H_{j-1} < h < H_j, \forall j \leq k; \\ [w_{j+1}^*, w_j^*], & \text{if } h = H_j, \forall j < k; \\ [M_2, w_k^*], & \text{if } h = H_k; \\ [M_2, M_2], & \text{if } h > H_k, \end{cases} \tag{A.8}$$

where $M_2 = min\{w_j^* \mid j \geq k-1\}$. If the compensated variation calculated based on the left and right convex budget sets are CV_1 and CV_2, then the desired CV will be $min\{CV_1, CV_2\}$.

APPENDIX B

IMPUTATION OF WIVES'S WAGE RATES

This appendix provides the estimated parameters used to impute the latent wage rates of nonparticipating wives. We follow Triest's procedure of applying Heckman's technique for correcting for sample selection bias. Table (XXV) and (XXVI) show the estimated parameters of wives' decisions of labor participation for 1983 and 2000 respectively. Slightly different from Triest's, we include the log of yearly mortgage payment as a regressor to predict the probability of working.

Variable	Coefficient
Constant	1.403
	(0.260)
Number of kids (Age$\leq$ 6)	-0.332
	(0.072)
Family Size	-0.118
	(0.044)
Age (35-45)	0.004
	(0.025)
Age-45	-0.014
	(0.018)
Wife's nonlabor income (in dollars)	-1.860e-05
	(2.910e-06)
Bad health	-0.269
	(0.183)
Log of yearly mortgage payment	0.021
	(0.013)
College education	0.436
	(0.101)
Log likelihood= -537.645	
Number of observations: 1004	

Table XXV.: Probit Model: Wives' Labor Participation (1983)

Table XXVI.: Probit Model: Wives' Labor Participation (2000)

Variable	Coefficient
Constant	0.167
	(0.346)
Age of the youngest child	0.053
	(0.012)
Number of kids (Age $<$ 18)	-0.034
	(0.057)
Age (35-45)	0.006
	(0.033)
Age-45	-0.014
	(0.018)
College	0.256
	(0.098)
Wife's nonlabor income (in dollars)	-2.220e-06
	(7.510e-07)
Bad health	-0.652
	(0.164)
Log of yearly mortgage payment	0.061
	(0.015)
Log likelihood= -463.729	
Number of observations: 1166	

Table (XXVII) and (XXVIII) present the results of wage regression for 1983 and 2000 respectively. Different from Triest (1990), we control for the variable indicating whether the women joins a union in the regression.

Table XXVII.: Wives' Wage Imputation Regression (1983)

Variable	Coefficient
$(\text{Age}/10)^2$	-0.380
	(0.137)
Education	-1.127
	(0.672)
Education^2	0.324
	(0.229)
Education*Age/10	0.298
	(0.081)
Nonwhite	-0.140
	(0.574)
Union	1.772
	(0.431)
County unemployment rate	-0.025
	(0.046)
Inverse Mills' ratio	0.918
	(0.829)
Constant	7.518
	(4.430)
R^2= 0.18	
Number of observations: 725	

Table XXVIII.: Wives' Wage Imputation Regression (2000)

Variable	Coefficient
Age	0.267
	(0.239)
Education	-4.266
	(2.068)
Education^2	0.324
	(0.229)
Education*Age/10	2.407
	(0.719)
Nonwhite	0.561
	(1.030)
Union	2.203
	(0.759)
Inverse Mills' ratio	0.227
	(1.933)
Constant	22.814
	(16.346)
R^2= 0.21	
Number of observations: 984	

VITA

Name: Gaosheng Ju

Address: Department of Economics, Texas A&M University
College Station, TX 77843-4228

Email Address: gju@econmail.tamu.edu

Education: B.S., Economics, Shandong Institute of Economics, 1999
M.S., Management, Peking University, 2003
Ph.D., Economics, Texas A&M University, 2011

The typist for this dissertation was Gaosheng Ju.