UC San Diego

UC San Diego Electronic Theses and Dissertations

Title

Essays in Public Finance

Permalink

https://escholarship.org/uc/item/6p81w5bq

Author

Miyoshi, Yoshiyuki

Publication Date

2016

Peer reviewed|Thesis/dissertation

eScholarship.org Powered by the California Digital Library
University of California

UNIVERSITY OF CALIFORNIA, SAN DIEGO

Essays in Public Finance

A dissertation submitted in partial satisfaction of the
requirements for the degree
Doctor of Philosophy

in

Economics

by

Yoshiyuki Miyoshi

Committee in charge:

Professor Roger Gordon, Chair
Professor Julie Cullen
Professor Joseph Engelberg
Professor Joshua Graff Zivin
Professor Natalia Ramondo
Professor Alexis Toda

2016

Copyright
Yoshiyuki Miyoshi, 2016
All rights reserved.

The dissertation of Yoshiyuki Miyoshi is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

_____

_____

_____

_____

_____

_____

Chair

University of California, San Diego

2016

TABLE OF CONTENTS

Signature Page . . . iii

Table of Contents . . . iv

List of Figures . . . vi

List of Tables . . . vii

Acknowledgements . . . viii

Vita . . . ix

Abstract of the Dissertation . . . x

Chapter 1 Does Sales Factor Apportionment Benefit the Welfare of State? . . . 1
1.1 Introduction . . . 2
1.2 Model . . . 6
1.2.1 Overview and discussions . . . 6
1.2.2 Production technology and firms . . . 10
1.2.3 Workers . . . 13
1.2.4 Governments . . . 14
1.3 State tax policy . . . 16
1.3.1 Optimal tax policy in autarky . . . 16
1.3.2 Effects of taxing apportionment factors . . . 18
1.4 Calibration . . . 22
1.4.1 Parameters . . . 22
1.4.2 Results . . . 24
1.5 Discussions . . . 27
1.6 Conclusion . . . 31
1.7 Appendix . . . 32
1.7.1 Profit maximization problem for firms . . . 32
1.7.2 Proof for Proposition 1 . . . 34
1.7.3 Proof for Proposition 2 . . . 35

Chapter 2 Risk-Taking and Growth under Taxation with Moral Hazard . . . 37
2.1 Introduction . . . 38
2.2 Model . . . 44
2.3 Solving for equilibrium . . . 49
2.3.1 Individual problem . . . 50
2.3.2 Equilibrium conditions . . . 52
2.3.3 Welfare . . . 55
2.3.4 Stationary distribution . . . 57

2.4 Numerical example 60
2.4.1 Functional forms 61
2.4.2 Calibration 62
2.4.3 Results 63
2.5 Concluding remarks 69
2.6 Appendix 69
2.6.1 Solution algorithm 69

Chapter 3 Rental Housing Investment and Tax Reforms: An Empirical Study of Tax Clientele Model 72
3.1 Introduction 73
3.2 A tax clientele model for rental housing investment 76
3.3 Tax reforms in 1980s 80
3.4 Data and methodology 82
3.4.1 Survey of Consumer Finances 82
3.4.2 Summary information on owners of rental housing . 83
3.4.3 Estimating marginal tax rates 84
3.4.4 Empirical specification 85
3.5 Empirical findings 86
3.6 Conclusion 90

References 92

LIST OF FIGURES

Figure 1.1: The Effective Average Federal Personal Income Tax Rate 24
Figure 1.2: The Effects of Sales Factor Weights on Welfare and Aggregate Variables in the Two-State Model 25
Figure 1.3: The Effects of Sales Factor Weights on Welfare and Aggregate Variables in the 50-State Model 28

Figure 2.1: Comparative statics with respect to the consumption tax rate ($\eta = 0.05$) 64
Figure 2.2: Comparative statics with respect to the consumption tax rate ($\eta = 0.075$) 67

LIST OF TABLES

Table 1.1: Optimal State Tax Rates in Nash Equilibrium of the Two-State Model 25
Table 1.2: Optimal State Tax Rates in Nash Equilibrium of the 50-State Model . 27

Table 2.1: Parameter values 63

Table 3.1: Probability of ownership (percent) of residential rental properties . . 83
Table 3.2: Probit estimates for the ownership of residential rental properties . . 87
Table 3.3: Average marginal effects of marginal tax rate on the probability of ownership 89

ACKNOWLEDGEMENTS

I would like to acknowledge Roger Gordon for his support as the chair of my committee. Studying public finance under a master of the field has been an honor and a privilege. His mentoring and guidance have always been invaluable, and he taught me how to deepen my understanding of real world issues using the economic theory.

I would also like to thank Alexis Akira Toda for his mentoring and many hours of discussion. Working on an ambitious research project together with him was very helpful especially when I was struggling to make the transition from a student to a researcher.

Many additional thanks to Julie Cullen, Joseph Engelberg, Joshua Graff Zivin and Natalia Ramondo. Their time, thoughts, and feedback helped me immensely throughout this process.

Of course, the wonderful opportunity to spend four years following my intellectual passions would never have been possible without the love and support of my family. A special thanks to Natsuki Miyoshi, Mihoko Miyoshi, Shigeto Miyoshi and Kazuo Miyoshi.

Chapter 2, in full, is currently being prepared for submission for publication of the material. Miyoshi, Yoshiyuki; Toda, Alexis Akira. The dissertation author was the primary investigator and author of this material.

VITA

2002	Bachelor of Arts, University of Tokyo, Japan
2008	Master of Arts, Tufts University
2016	Doctor of Philosophy, University of California, San Diego

ABSTRACT OF THE DISSERTATION

Essays in Public Finance

by

Yoshiyuki Miyoshi

Doctor of Philosophy in Economics

University of California, San Diego, 2016

Professor Roger Gordon, Chair

Researches in public finance have clarified the effects of tax policies that are neither originally intended nor easily observable. This work contributes to the literature of public finance by rigorously examining the effects of three particular tax policies using quantitative models and econometric analysis.

Chapter 1 aims to answer the question as to what is the optimal apportionment formula for the US state corporate income tax. States have raised sales apportionment weight and lower payroll weight to stimulate local labor demand. However, the policy discussions often ignore the negative effect of sales apportionment tax; the tax distorts the sales allocation of firms across states and causes an increase in local price level in

the state. I construct a quantitative model that incorporates the effects of apportionment formula both on local labor demand and on the price level. The calibration suggests that the sales weight should be zero for the optimal tax policy of a single state because the negative effect on price level outweighs the positive effect on local labor under a range of plausible parameters.

Chapter 2 studies the effect of taxation on entrepreneurs' risk-taking, portfolio choice, and economic growth in the presence of a moral hazard problem in a dynamic general equilibrium model. The moral hazard problem occurs because the return of a type of entrepreneurial project depends on entrepreneurs' effort, which is private information. By collecting proportional consumption tax and redistributing the revenue, the government offers an opportunity for risk sharing to encourage risk-taking and spur economic growth. We show that when the moral hazard problem is absent, the full insurance is optimal in terms of welfare, and the economy grows faster. But if the moral hazard problem exists even in a slight degree, risk sharing through taxation cannot improve welfare.

Chapter 3 examines the relationship between household marginal tax rates and the probability of owning rental housing. I focus on the special provisions about passive losses introduced by 1986 TRA to test the theoretical prediction about this relationship. The empirical results based on the Surveys of Consumer Finances offer modest support for the prediction.

Chapter 1

Does Sales Factor Apportionment Benefit the Welfare of State?

Abstract

States have raised the sales apportionment weight and lowered the payroll weight to stimulate local labor demand. However, the policy discussions often ignore the negative effect of sales apportionment tax; the tax distorts the sales allocation of firms across states and causes an increase in the local price level in the state. This study examines the optimal tax policy from the perspective of a single state and predicts Nash equilibrium with a quantitative model that incorporates the effects of apportionment formula both on local labor demand and on the price level. The calibration suggests that the sales weight should be zero for the optimal state tax policy because the negative effect on the price level outweighs the positive effect on the local labor demand under a range of plausible parameters.

1.1 Introduction

Formula apportionment constitutes a major difference in the systems of corporate income tax between the federal and state governments in the U.S.[1] The apportionment formula determines what fraction of the taxable domestic profit of a multi-state firm is subject to state corporate income tax in each state. Adjusting this formula can significantly change the distribution of tax liability of a firm across states and then can affect the firm's decision about location, production and sales. The previous studies have found apportionment formula has impacts on the welfare and aggregate economic variables, including the prices and employment.

Most of the U.S. states use the three-factor apportionment formula, which defines the tax liability of multi-state firm j for state n as follows:

$$T^n(j) = t_C^n \left(\gamma_W^n \frac{W^n(j)}{W(j)} + \gamma_K^n \frac{K^n(j)}{K(j)} + \gamma_S^n \frac{S^n(j)}{S(j)} \right) \pi^T(j), \tag{1.1}$$

in which $W^n(j)$, $K^n(j)$, and $S^n(j)$ represent payroll, property, and sales of firm j in state n, respectively; $W(j)$, $K(j)$, and $S(j)$ represent total domestic payroll, property, and sales of firm j respectively; γ_W^n, γ_K^n, and γ_S^n are the weights for payroll, property, and sales factors respectively set by state n and $\sum_h \gamma_h^n = 1$; $\pi^T(j)$ is the taxable domestic profit for firm j. In the U.S., states can choose these factor weights independently as long as the weights sum up to one. From this formula, state corporate income tax can be viewed as a combination of three separate taxes on payroll, property, and sales of firms, as McLure (1981) points out.

Most states used to opt for the equally weighted formula, which puts the equal weight on three factors, because the Multistate Tax Compact recommended the formula.

[1] Formula apportionment is adopted in some of other developed countries: for example, Canada and Japan. But, unlike in the U.S., the formula is uniformly set by the national government in those countries.

After the Supreme Court upheld the right of states to use other formulas than the equal weight formula in 1978, many states started to lessen the weights on payroll and property factors and put more weight on the sales factor to stimulate the demand for local employment (Mazerov 2001). This trend in state tax policy still continues, if not accelerates, these days. Now 19 states have even adopted the single sales factor apportionment, which put the full weight on the sales factor. The conventional wisdom in policy discussions is that larger sales weight benefits the state, especially through an increase in the demand for local labor, but the policy is not desirable from the perspective of nation because the competition of tax policy among state governments eventurelly leads to a "prisoner's dilemma" type equilibrium.

It is often overlooked, however, that taxing the sales factor has its own cost. If a state sets a higher tax rate for sales factor ($t_C^n \gamma_S^n$ in equation 1.1) than the other states, the tax liability of a firm increases as the firm sells more in the state.[2] Thus the higher tax rate for the sales factor distorts the sales of firms in the way that firms sell less in the state. At the aggregate level, this leads to relatively less sales and higher prices in the state compared to other states with a lower tax rate for the sales factor. In sum, a greater sales weight will not only raise local real income, either through improved employment or through a higher wage rate, but it will also reduce local real income through a higher price level at the same time. As a whole, the total effect of change in the apportionment formula is ambiguous.

The goal of this study is to derive the optimal corporate income tax policy for state governments, including the optimal approtionment weights, with a quantitative model that is calibrated plausibly. In particular, the model formalizes the trade-off between local labor demand and price distortion caused by the apportionment formula. The literature

[2] In this paper, I assume the sales in apportionment formula is defined as sales at destination rather than at origin. Gordon and Wilson (1986) follows the same definition because it matches the provisions of tax code of most U.S. states.

on formula apportionment has successfully illuminated the former effect, especially by Goolsbee and Maydew (2000). However, no theoretical studies have been done to examine the optimal apportionment formula with a model that incorporates this trade-off and is able to evaluate the combined effect.[3]

Another key feature of my model is that states use personal income tax as well as corporate income tax as a source of revenue. If states only considered the effect of corporate income tax on the local labor market and if personal income tax were available to them with no cost, then the best tax policy would be choosing zero (or even negative) corporate income tax and collect the necessary revenue by personal income tax. On the contrary, it is a conspicuous observation from the actual state tax policy in the U.S. that most states impose corporate income tax and, on average, the corporate income tax rate is close to the top marginal tax rate of personal income: these rates are 7.33% and 7.58% respectively in 2014, if weighted by state GDP.[4] Since the revenue from personal income tax acounts for much larger share of the total state tax revenue than corporate income tax, it is important to examine states' choice about personal income tax when a study wants to derive implications for the optimal state corporate income tax. However, the previous theoretical studies on the optimal apportionment formula do not allow states to use other sources for revenue than corporate income tax.[5]

States need both personal and corporate income taxes in my model because of income shifting behavior among workers. Workers can shift their income from labor income to capital income if their effective tax rate for labor income is higher than the

[3] As examples of studies regarding the effect of sales factor, Gordon and Wilson (1986) theoretically analyze the effect of sales weight on output prices separately from the effects of property or payroll weight, and derive the result of "cross-hauling" caused by uneven tax rates for sales factor. The empirical study of Edmiston and Arze del Granado (2006) reports that the increase in sales weight led to a drop in local sales of multi-state firms as well as a rise in payroll of them in State of Georgia. Fajgelbaum et al. (2015) focuses on the improvement of welfare that can be attained by harmonizing state tax policy, including apportionment formula, rather than the optimal tax policy from the perspective of a single state.

[4] This relation is also found in federal corporate and personal income taxes.

[5] For examples, see Anand and Sansing (2000), Pinto (2007), and Runkel and Schjelderup (2011).

rate for capital income. Thus if a state sets personal and corporate income tax rates being wildly different, the gap in the tax rates severely distorts such income shifting behavior of workers. In short, state govenrments use corporate income tax as the "backstop" for implementation of personal income tax system (Mirrlees 2011).

The model used in this study is a general equilibrium trade model in which multiple states set their tax policy strategically in the setting of static game. Firms respond to the state tax policy by choosing location in which to operate and the vector of price and output for each state. Basically, the tax rate for the payroll factor affects the location choice of firms while the tax rate for the sales factor does the prices and output that firms choose. In the calibration exercises, I look for Nash equilibrium of the model. Contrary to the popular policy recommendation and the implications from the existing literature, the model suggests that zero sales weight should be chosen in equilibrium under plausible paraemters because the negative effect of sales apportionment on the price level outweighs its positive effect on the local labor demand. In addition, even if the current equilbrium is distorted due to some exogenous restriction on apportionment weights, say that the sales weight should equal one third, the best response of a state is to reduce sales weight to zero once it starts to have the right to adjust the weights on its own.

Section 2 develops the model and defines the objective of state governments. Section 3 describes the specific features of the model: the optimal balance between corporate and personal income taxes and the effects of apportioment weights on the aggregate economic variables. Section 4 presents the results of calibration and simulation. Section 5 discusses the validity of the model, including comparing the model's prediction with the existing empirical study. Section 6 concludes.

1.2 Model

1.2.1 Overview and discussions

This study use a model to derive implications about the optimal tax policy from the perspective of a single state government and Nash equilibrium of state tax policy of multiple states. Therefore it is critical to define the objective of state government plausibly. State governments in this model aim to maximize the social welfare in the state, which is defined as the purely utilitarian welfare function, by using a restricted set of policy instruments. A state government collects its revenue through personal and corporate income taxes, and the schedules of those taxes are constrained to be linear. It produces services that the private sector does not produce, and provides an equal amount of the services (hereafter referred to as the state government services) to each of its residents. The federal government also imposes personal and corporate income taxes. Federal personal income tax has a progressive schedule while federal corporate income tax a linear one. Another important assumption in my model is that workers supply their endowment of labor inelastically regardless of the state and federal tax policies.

The key feature of my model is that state governments set personal and corporate income tax rates at the same time and keep a balance between these two tax rates. The mechanism to detemine the optimal balance for state govenrments is centered on income shifting behavior of high income workers. In this model, workers can choose to receive the payment for the labor they supply as capital income rather than as labor income at no cost if they want. Capital income is subject to corporate income tax as part of firm's taxable profit but not subject to any tax at the personal level. In addition, workers are heterogeneous in labor efficiency and then in income. For workers with high labor efficiency, it may be advantageous to shift their income from labor income to capital income since federal personal income tax has a progressive schedule while federal and

state corporate income tax has a linear one. Thus given federal and state tax rates, there is the cut-off value of income; workers whose income is above the cutoff value undertake income shifting.

The objective function of state governments in this model is defined as it is not sensitive to the allocation of income across workers. Although the utility function of workers is a Cobb-Douglas function that is *quasi*-concave in the private goods and state government services, it is defined to be linear in private consumption. Thus the social welfare function that adds up the utility of workers with an equal weight is linear in the aggregate private consumption, or equivalently in the sum of real after-tax income across workers in the state.

Given this objective function, state governments do not care about the distribution of effective tax rates across workers as the result of income shifting. State governments, however, do care about the aggregate income of workers after federal taxes are collected. A state can change the cutoff income for income shifting in the state by setting different tax rates for personal and corporate income taxes, and, as the cutoff income changes, federal tax collection also changes. As I argue in the next section in detail, if state tax policy did not affect the economic variables, including the wage rate and the price level, equal tax rates for state corporate and personal income taxes would minimize federal tax collection. In other words, it would be best for states not to use uneven tax rates for corporate and personal income taxes because such tax policy distorts income shifting behaviour of workers. The combination of inelastic labor supply and the social welfare function that is linear in private consumption enables the model to incorporate the balance between corporate and personal income taxes as an endogenous choice of state governments in a simple way.

Although my model makes states choose the optimal mix of corporate and personal income taxes endogenously, the limited policy instruments for the government

sector and inelastic labor supply of workers are restrictive and different from the standard optimal taxation literature, in which the government searches for the flexible tax schedule that maximizes the social welfare defined with a certain degree of redistributive taste under the constraint of private information.[6] However, the structure of my model makes the focus of this study—the trade-off between local labor demand and the price level—transparent. Moreover, it approximates the principal features of actual state government finances to a considerable extent, and can readily offer quantitative policy implications for state governments within the restriction of policy tools.

In practice, the progressivity of state personal income tax is moderate on average: among the 43 states that impose personal income tax (including the District of Columbia), 10 states have a linear tax schedule in 2015.[7] In addition, in the 18 states out of the 33 states that have progressive tax schedules, the top rate is reached at a household income level below the national average, and, in most cases, below $20,000. The average tax rate for the household with the national average income is 5.34%, if weighted by state GDP, while the weighted average tax rate for the highest bracket is 7.33% in 2014. As for the effective tax rates rather than the statutory rates, Gordon and Cullen (2012) report the effective marginal tax rates of federal and state personal income tax that are calculated based on the data including individual tax returns. The effective marginal tax rates of state income tax range between 3.5% and 6.2% across the top four income groups out of the five income groups, excluding the bottom group, while the federal rates range between 15.0% and 31.2% across the same groups.[8] As for corporate income tax, two

[6]Notable studies with policy implications in the standard optimal taxation literature are Saez (2001) and Diamond and Saez (2011).

[7]The states that do not impose personal income tax collect a large fraction of their revenue from general sales tax, except for Alaska and Wyoming. General sales tax with a single tax rate is considered to be equivalent to a linear personal income tax in my static model, although the model does not accomodate the option of sales tax as an endogeneous choice for state governments.

[8]Gordon and Cullen (2012) report a very high marginal tax rate of state personal income tax for the bottom income group. They attribute it to transfer income that the households in the group lose as their income goes up.

thirds of states have linear tax schedules. Therefore, it appears a plausible approximation of the actual U.S. state tax policy to restrict the choice of state personal and corporate income tax schedules to being linear.

On the spending side of state government finances, the largest share is spent on education (35.6%). State governments are also involved in provision of large-scale public infrastructure, most notably highway projects (6.7%). These facts motivate my model to have state governments produce services that the private sector does not and provide the services equally to each of their residents.[9]

The model is a general equilibrium trade model of multiple states. In the following two sections, the economy is assumed to consist of two states, A and B, for the purpose of presentation although it is relatively straightforward to extend the model to an economy with more than two states. State governments are players of a static game who aim to maximize their objective function taking account of general equilibrium outcome. There are continuum of firms of fixed measure in the economy. Each of them produces a differentiated good. Every firm is perfectly mobile and chooses one of the states to operate in to maximize its profit, considering the wage rates, local productivity, and state tax policies. On the contrary, workers are immobile in this model and consume a set of the differentiated goods and the state government services. Only state governments can produce the state government services, using the set of differentiated goods as inputs. The differentiated goods can be transported between states with no cost, but the state government services cannot be transported to the other state.

[9]The treatment of state expenditure here is almost equivalent to Fajgelbaum et al. (2015), in which the level of state government spending affects the level of local amenity that enters consumers' utility equally, although the level of state government spending also affects the productivity of firms in their model. On the contrary, state governments also engage in income redistribution not through a highly progressive income tax, but through the expenditure on public welfare programs, including Medicaid. That type of expenditure accounts for 30.8% of total expenditure in 2013. Gordon and Cullen (2012) focus on this role of state governments and provide a theoretical framework for the optimal taxation in which each of multi-level governments tries to maximize thier own welfare function by nonlinear labor income tax.

1.2.2 Production technology and firms

Firms of measure M produce the differentiated goods using labor as the only input. The production function for firm j is

$$x(j) = z_n(j)l(j),$$

in which $z_n(j)$ is the productivity of firm j if it operates in state n. A perfectly competitive sector in each state produces the final good from the differentiated goods according to the following technology:

$$X = \left(\int_M x(j)^{\frac{\sigma-1}{\sigma}} \mathrm{d}j \right)^{\frac{\sigma}{\sigma-1}}, \tag{1.2}$$

in which the substitution parameter $\sigma > 1$. The final good is either consumed by workers in the state or used to produce the state government services. The price of the final good in state n is expressed as

$$P^n = \left(\int_M p^n(j)^{1-\sigma} \mathrm{d}j \right)^{\frac{1}{1-\sigma}},$$

in which $p^n(j)$ is the price that firm j sets for its good sold in state n.

Firms maximize their profits as they choose their production levels, the prices for their goods, and the locations of production. A firm operates in a single state and is not allowed to merge with another firm in this model. All the firms are subject to federal and state corporate income tax and not allowed to choose any organizational forms to which corporate income tax is not applied.

The state corporate income tax system follows a two-factor formula apportionment.[10] In addition to their corporate income tax rates, state governments set their

[10]Production does not require capital as an input in this model. However, if production function is

apportionment weights on the payroll and sales factors. Let γ_S^n be the sales weight of state n, and then its payroll weight equals $1-\gamma_S^n$. If firm j chooses to locate in state n, the tax liability for the firm is

$$T_n(j) = \left(t_C^F + \bar{t}_n(j)\right)\pi_n^T(j),$$

in which t_C^F is federal corporate income tax rate; $\bar{t}_n(j)$ is the effective state corporate income tax rate for firm j, which is calculated based on the apportionment formula; $\pi_n^T(j)$ is the taxable profit for firm j.[11] The apportionment formula defines $\bar{t}_n(j)$ as

$$\bar{t}_n(j) = t_C^n(1-\gamma_S^n) + \frac{t_C^A\gamma_S^A p^A(j)x^A(j) + t_C^B\gamma_S^B p^B(j)x^B(j)}{p^A(j)x^A(j) + p^B(j)x^B(j)}, \tag{1.3}$$

in which t_C^n is the statutory rate for corporate income tax of state n; $x^m(j)$ is the amount of firm j's good that is sold in state m.

The payment to workers is the only cost of production for firms. In this model, a worker can negotiate with the firm to receive the payment either as labor income or as capital income. The payment of labor wage is deductible from the taxable profit of firms. But if a firm pays capital income to a worker rather than labor wage, the payment adds to the firm's taxable profit and is subject to federal and state corporate income tax. Thus the *wage rate* for capital income payment to the *income-shifters* at firm j that operates in state n is determined as

$$\tilde{w}_C^n(j) = \left(1 - t_C^F - \bar{t}_n(j)\right)w_N^n, \tag{1.4}$$

in which w_N^n is the prevailing market wage rate for labor income in state n. The tilde on a

homogeneous of degree one in labor and capital and if capital is also immobile (like land), the following results remain true.

[11] To make the model simple and transparent, the federal government in this model does not allow firms to deduct the payment of state corporate income tax from their federal taxable income, unlike the U.S. tax code.

variable in this paper means that the variable is the after-tax one.

The after-tax profit of firm j, when it chooses to locate in state n, is represented as

$$\begin{aligned}\tilde{\pi}_n(j) =&(1-t_C^F-\bar{t}_n(j))\left[p_n^A(j)x_n^A(j)+p_n^B(j)x_n^B(j)-w_N^n(1-\theta_n)\frac{x_n^A(j)+x_n^B(j)}{z_n(j)}\right]\\&-\tilde{w}_C^n\theta_n\frac{x_n^A(j)+x_n^B(j)}{z_n(j)},\end{aligned} \tag{1.5}$$

in which θ is the share of efficiency unit of labor that is provided by income-shifters; the subscripts n on some of the variables repersent being conditional on operating in state n. In this model, θ_n is not a choice variable of each firm, but every firm in a state is assumed to employ the same fraction of income shifters. On the other hand, the effective state tax rate $\bar{t}_n(j)$ is an endogenous variable that is determined by the firm's decision on the allocation of sales among the states. Nonetheless, all the firms in the same state choose the same allocation of sales shares as the solution to its profit maximization problem, even if they are heterogeneous in productivity. Refer to Section 1.7.1 for the proof.[12] Since it implies all the firms in the same state are subject to the same effective state tax rate $\bar{t}_n$, the wage rate for income-shifters $\tilde{w}_C^n$ is uniquely determined by equation 1.4. Note that this makes all the firms in state n indifferent about θ_n from equation 1.5. Finally, firm j chooses to locate in the state where the maximum of $\tilde{\pi}_n(j)$ is the larger.

Since firms engage in monopolistic competition, firms earn positive profits after paying out capital income to income-shifters. Those profits are pooled and distributed to workers by a perfectly competitive financial sector. ρ^A fraction of the total after-tax profit is distributed to workers in state A while the remaining, the fraction of $\rho^B = 1-\rho^A$, to

[12] Fajgelbaum et al. (2015) first prove this result of constant $\bar{t}_n(j)$ across all the firms in the same state for the standard Dixit-Stiglitz model of monopolistic competition. The proof in Section 1.7.1 shows their result can be extended to the model with income shifting with minor adjustments.

state B.

1.2.3 Workers

There are N^n workers in state n. Workers are heterogeneous in terms of their endowment of efficiency unit of labor, $l^n(i)$ for worker i, and they supply labor inelastically. L^n stands for the mean efficiency unit of labor of workers in state n, and then the aggregate labor supply in state n equals $N^n L^n$. In addition to labor income, and capital income for income-shifters, workers receive dividend from the financial sector. Workers in state n own shares for dividend in proportion to their efficiency unit of labor.[13] Thus the dividend payment for worker i in state n is

$$\tilde{d^n}(i) = \frac{\rho^n l(i) \tilde{\Pi}}{N^n L^n},$$

in which $\tilde{\Pi}$ is the aggregate profit of firms after the payments for income shifters and for the federal and state corporate tax liabilities.

Labor income is subject to federal and state personal income taxes while capital income for income-shifters and dividend income are not. Since federal personal income tax has a progressive schedule, the liability of that tax is larger for high income workers. Let $t_N^F(i,n)$ be a shorthand notation for the *average* tax rate for worker i in state n if the worker chooses not to be an income-shifter: $t_N^F(i,n) = T_N^F(w_N^n l^n(i))/(w_N^n l^n(i))$, in which $T_N^F(\mathbf{.})$ is the function for the schedule of federal personal income tax liability. Thus the

[13] I make this assumption regarding the distribution of financial income being motivated by the well-known fact that the distribution of wealth is significantly skewed. However, the following results are not affected by the assumption because, as I explain below, the state welfare function is linear in aggregate income.

total after-tax income of worker i in state n is

$$\tilde{y}^n(i) = \begin{cases} \tilde{w}_C^n l^n(i) + \tilde{d}^n(i) & \text{if } i \text{ is an income-shifter,} \\ (1 - t_N^F(i,n) - t_N^n) w_N^n l^n(i) + \tilde{d}^n(i) & \text{otherwise.} \end{cases} \tag{1.6}$$

in which t_N^n is state personal income tax rate in state n. Since the cost of income shifting for workers is zero in this model, they choose the type of income that makes $\tilde{y}^n(i)$ the larger. Thus equation 1.4 implies that worker i in state n becomes an income-shifter if the personal income tax liability is larger than the incidnce of federal and state corporate income tax, or $t_N^F(i,n) + t_N^n > t_C^F + \bar{t}_n$.

Workers consume the state government service as well as the private final good. Worker i in state n has the following utility function:

$$u^n(i) = \frac{\tilde{y}^n(i)}{P^n} (g^n)^{\alpha_G}, \tag{1.7}$$

in which g^n is the state government service provided to worker i by the state government in n; $\alpha_G > 0$ is the preference parameter for the state government services. Note that this utility function is a standard Cobb-Douglas function and exhibits the quasi-concave property although it is linear in (real) income.

1.2.4 Governments

State governments maximize the social welfare of residents that is defined as the purely utilitarian welfare function:

$$V^n = \int_{i \in n} u^n(i) f^n(i) N^n \mathrm{d}i, \tag{1.8}$$

in which $f^n(i)$ is the probability density function for the distribution of heterogeneous workers in state n. The budget constraint for state government is

$$N^n g^n \leq z_G^n (T^n / P^n), \tag{1.9}$$

in which z_G^n is the productivity for state government services in state n; T^n is the aggregate tax revenue for state n.[14] State governments take into account only how the federal tax collection affects the utility of their residents, but not how the changes in federal revenue caused by the state's fiscal policy affect the utility.[15] The federal government takes the final good from the economy as its tax collection.

Since the utility function defined as equation 1.7 is linear in real income for worker i and g^n is constant across all the workers in state n, the social welfare function for state n is linear in the aggregate income of the state. Therefore state governments in this model are not sensitive to the allocation of income among their residents. This feature isolates the aggregate effects of formula apportionment on the welfare of state from the issue of income redistribution, still keeping the heterogeneity across workers, which is needed to incorporate the choice of state governments about the optimal mix of personal and corporate income taxes.

[14] Since the state welfare function preserves the property of Cobb-Douglas utility function, z_G^n is irrelevant for the optimal tax policy.

[15] This assumption can be justified, for example, when the federal government covers a lot more states with which the two states have no economic interaction, and when the two states are too small to have a significant effect on the total federal revenue.

1.3 State tax policy

1.3.1 Optimal tax policy in autarky

I use a variant of my model, which turns the trading economy into the autarky economy, to illustrate how the environment of this model determines the optimal mix of state corporate and personal income taxes. In this case, state A has no economic interaction with state B: there is no trade between the two states, and the profits of firms are not pooled nationally, but are distriburted within the state. The federal government, however, still collects tax from both states. Since the apportionment weights are irrelevant in this case, the variables the state governments use are personal and corporate income tax rates.

From equation 1.2, the demand for good j is:

$$x(j) = p(j)^{-\sigma} \frac{Y^n}{(P^n)^{1-\sigma}}, \tag{1.10}$$

in which Y^n is the before-tax aggregate income in state n. The nominal price level is fixed in such a way that Y^n is normalized to one. The optimization problem for firm j is:

$$\max_{p(j),x(j)} (1 - t_C^F - t_C^n) \left(p(j)x(j) - \frac{w_N^n}{z(j)} x(j) \right),$$

subject to equation 1.10. This objective function is simplified from equation 1.5 by substituting equation 1.4. It can be shown that the equilibrium before-tax wage rate and price level do not depend on either federal or state tax policy:

$$w_N^n = \left(\frac{\sigma - 1}{\sigma} \right) \frac{1}{N^n L^n};$$

$$P^n = (\bar{z}^n)^{-1},$$

in which

$$\bar{z}^n = \left[\int_M z^n(j)^{\sigma-1} g^n(j) dj\right]^{\frac{1}{\sigma-1}},$$

in which $g^n(j)$ is the probability density function for firms in state n. The implication that state tax policy does not distort the production in the autarky economy makes the problem of optimal state tax policy simple as Proposition 1 shows:[16]

Proposition 1. [The optimal state tax rates in autarky] *In the case of autarky, the optimal rates for state personal and corporate income tax are equal:* $t_S^n = t_C^n$.

The proof for Proposition 1 is presented in Section 1.7.2. This result can be interpreted as follows. If there were no state taxes, the federal personal and corporate income taxes would completely determine the cutoff efficiency unit of labor $\bar{l}$ in state n such that $T_N^F(w_N^n \bar{l}) = t_C^F w_N^n \bar{l}$. Those who are more efficient than $\bar{l}$ engage in income shifting while those who are less efficient do not. With this being said for the no state tax case, if the state sets equal rates for both income taxes, the cutoff efficiency does not change. However, if the state tax rates do not match, it changes the cutoff efficiency and distort the behavior of workers that originally minimizes the *federal* tax burden. Therefore uneven state tax rates increase the federal tax collection, which flows out of the state economy. Since the welfare function for state government is linear in the aggregate after-tax income and is not sensitive to the allocation of after-tax income among the residents, the state government does not want to distort the cutoff efficiency.

The interpretation can be extended to a broader perspective. My model abstracts from elastic labor supply and the benefits and costs regarding firms' choice of organizational form. Although this feature makes the model tractable, its cost is that the optimal federal personal and corporate income tax schedules should be taken as exogenous.[17] If

[16] Although monopolistic competition is present in this model, the production activity is not distorted because workers supply labor inelastically and all the profits of firms are distributed to workers.

[17] Gordon and Mackie-Mason (1994) and Mackie-Mason and Gordon (1997) provide the empirical

the federal government implements its tax policy, especially the arrangement of personal and corporate income taxes, considering factors that my model cannot capture, Proposition 1 suggests state governments respect the choice of federal government except for a uniform linear tax that does not distort firms' choice of organizational form in the case of autarky. As the following sections argue, however, when the two states interact, state governments are faced with the trade-off between the opportunity for boosting local labor demand and the distortion to income shifting behaviour. In addition, it complicates the state governments' problem that part of the tax burden can be exported through taxing the profits of firms in the other state.

1.3.2 Effects of taxing apportionment factors

In the model of trading states, state corporate income tax affects the location and production choices of firms since firms are perfectly mobile across the states. It requires specifiying the distribution of productivity among firms to analyze the effect of state corporte tax on firms' location choice. I choose the assumption of homoegneous firms, in which $z^n(j) = z$ for all j and n in the rest of this section and the following calibration exercises. Section 1.5 discussses the implications of this assumption.

The corpoarte tax rate affects firms' choice, interacting with the apportionment weights. If the effective tax rate on the payroll factor in state A is higher than in state B, it reduces the demand for labor in state A and then lowers the before-tax wage rate there. If the effective tax rate on the sales factor in state A is higher, it raises output prices in the state since firms have an incentive to reduce the corporate income tax burden by selling less in state A and more in state B. The rest of this section examines the effects of state corporate income tax in the two polar cases regarding the apportionment weights

evidence that the difference between corporate and personal income tax distorts the choice of organizational form. Piketty et al. (2014) propose a model for the optimal labor income tax in the presence of income shifting.

to illustrate how the choice of the weights affects firms' behavior and the aggregate variables.

Payroll factor

First assume the states put the full weight on payroll factor and zero on sales factor. Since firms are homogeneous and mobile, they are indifferent about the location in equilibrium:

$$\tilde{\pi}_A(j) = \tilde{\pi}_B(j) \quad \text{for all } j, \tag{1.11}$$

in which

$$\tilde{\pi}_n(j) = \max_{p_n(j),x_n(j)} (1 - t_C^F - t_C^n) \left[p_n^A(j) x_n^A(j) + p_n^B(j) x_n^B(j) - w_N^n \frac{x_n^A(j) + x_n^B(j)}{z} \right].$$

The first order conditions imply that frims sell their goods at the same price in both states:

$$p_n^A(j) = p_n^B(j) = \left(\frac{\sigma}{\sigma - 1} \right) \frac{w_N^n}{z}.$$

It is possible to derive the relation between the corporate income tax rates and the equilibrium wage rates:

$$\frac{w_N^A}{w_N^B} = \left(\frac{1 - t_C^F - t_C^A}{1 - t_C^F - t_C^B} \right)^{\frac{1}{\sigma - 1}}. \tag{1.12}$$

This implies the elasticity of wage rate with respect to net-of-tax rate, $1 - t_C^F - t_C^n$, is $1/(\sigma - 1)$ if the tax rate in the other state is fixed.

Equation 1.12 means that state governments can increase the local wage rate by reducing its corporate income tax rate, or its effective tax rate for the payroll factor more precisely.[18] This result fits the previous empirical studies: for example, Goolsbee and

[18] In the limiting case, state corporate tax does not affect the local wage as σ goes to infinity. For the market is perfectly competitive and firms earn no economic proft in this case. Corporate income tax revenue only comes from income shifters in the state and the full incidence falls on those workers without

Maydew (2000) and Edmiston and Arze del Granado (2006).

However, there are two considerations that offset this positive effect of lower corporate income tax rate. First, if firms earn economic profits and state tax policy does not affect the distribution of ownerships of firms, as this model presumes, the incidence of state corporate income tax partly falls on the owners outside the state. This makes taxing the profits of firms a less costly way to raise revenue. Second, if a state government taxes firms lightly, the gap between personal and corporate tax rates changes the income shifting behavior from one that federal tax policy intends and, as a result, it increases the federal tax collection, which is taken away from the state economy. Therefore, the quantitative implications for optimal state tax policy in this case depends not only on the substitution parameter σ, but also on the distribution of firm ownerships and federal tax policy.

Sales factor

Next assume the states put the full weight on the sales factor and zero on the payroll factor. In this case, the effective corporate tax rate for firms is not affected by the location of firms, but it is completely determined by the ratio of sales between the two states.[19] State corporprate income tax does not affect the production cost, and homogeneous firms locate in whichever state that offers the lower production cost. Thus, $w_N^A = w_N^B = w_N$ in equilibrium.

distortion.

[19] In practice, the effective tax rate also depends on the location of firms because of the nexus rules in the U.S. tax code. Section 1.5 discusses this issue further.

The optimization problem for firm j is:

$$\max_{p_n(j),x_n(j)} (1-t_C^F-\bar{t}_n)\left[p_n^A(j)x_n^A(j)+p_n^B(j)x_n^B(j)-(1-\theta_n)w_N^n\frac{x_n^A(j)+x_n^B(j)}{z}\right] - \theta_n w_C^n\frac{x_n^A(j)+x_n^B(j)}{z},$$

in which

$$\bar{t}_n(j)=\frac{t_C^A p_n^A(j)x_n^A(j)+t_C^B p_n^B(j)x_n^B(j)}{p_n^A(j)x_n^A(j)+p_n^B(j)x_n^B(j)}.$$

The closed-form solution for equilium is not derived in this case, but the qualitative effect of taxing sales factor on the price level can be shown as follows:

Proposition 2. [Distortion by unequal sales factor taxation] *When states use the single sales factor, namely $\gamma_S^A=\gamma_S^B=1$ in equation 1.3, the following statements concerning the prices of final good, P^A and P^B, hold.*

1. *If $t_C^A=t_C^B$, then $P^A=P^B=\bar{P}$, in which*

$$\bar{P}=\left(\frac{\sigma}{\sigma-1}\right)\frac{w_N}{z}.$$

2. *If $t_C^A>t_C^B$, then $P^A>\bar{P}>P^B$.*

Refer to Section 1.7.3 for the proof. Proposition 2 shows a greater weight on the sales factor carries some cost to the state economy as the local price level rises. Note that the structure for the optimal tax policy is similar to the one in the case of the full payroll weight; a low corporate tax rate may improve the state's welfare through a lower price level, which leads to higher real income because w_N is equal among states. But there are two offsetting effects: the distortion caused by the gap between corporate and personal income tax rates and the possible benefit of exporting tax burden through taxing the corporate profits of firms in the other state.

1.4 Calibration

The arguments in the previous section illustrate the effects of taxing the payroll and sales factors separately by factors. It shows state governments are faced with the trade-off when they consider apportionment weights: if a state raises its weight on sales, then the local wage will go up, but so will the price of final good in the state at the same time, which harms the state welfare by reducing the real income. Calibration exercises are necessary to evaluate the total effect and find the optimal tax policy for states. I first explain the parameters for the calibration, and then report the results.

1.4.1 Parameters

The income distribution among workers is important because it determines how many workers undertake income shifting. Considering labor income is linear in efficiency unit of labor and only the highest part of the distribution matters for income shifting behavior in the model, I assume the distribution of efficiency unit of labor follows a Pareto distribution.[20] The density function of efficiency in state n is defined as

$$f^n(l(i)) = \frac{\eta(l_m^n)^\eta}{(l(i))^{\eta+1}},$$

in which l_m^n is the efficiency unit of the least efficient workers in state n and $\eta > 1$. I set $\eta = 5/3$ following Jones (2015). Since the mean efficiency in state n is L^n, $l_m^n = L^n(\eta-1)/\eta$.

Federal personal income tax rate is the other determinant for income shifting behaviour. I replicate its progressive schedule following Gouveia and Strauss (1994).[21]

[20] In particular, the distribution of middle and low income workers does not affect the state welfare because the welfare function is linear in aggregate income in this model. For the highest part of the distribution, Piketty and Saez (2013) show the U.S. before-tax income closely follows a Pareto distribution especially above the income of $400,000.

[21] Conesa et al. (2009) and Conesa and Krueger (2006) also use this function to parametrize the optimal

In their model, the effective average tax rate for federal personal income tax is expressed as:

$$t_N^F(y_N) = b_2 - b_2(b_1(y_N)^{b_0} + 1)^{-1/b_0}, \tag{1.13}$$

in which y_N is labor income. They estimate the parameters in equation 1.13 and report values of $b_0 = 0.768$, $b_1 = 0.031$, and $b_2 = 0.258$ for year 1989. Note that the limit of t_N^F equals b_2 as labor income goes to infinity. Since the highest marginal tax rate is lower in 1989 than today, I use value of $b_2 = 0.396$, which is the highest marginal tax rate in the current tax schedule, in the following calibration exercises. Since equation 1.13 is not linear in labor income, I have to adjust the nominal price level to replicate the tax schedule properly. The nominal price level is set such that the mean before-tax labor income in the model equals \$76,000 when the two states are symmetric.[22] The effective tax rate for federal personal income tax by income level is presented in Figure 1.1. Combined with the Pareto distribution of labor efficiency, the share of shifted income out of total labor income (θ) equals 0.137 when two identical states use a symmetric tax policy.[23]

The elasticity of substitution σ in the final good production function is the key parameter in this model since it affects the elasticity of wage rate with respect to tax rates and the taxable profits of firms. I use the value of $\sigma = 4$ following Fajgelbaum et al. (2015) as the benchmark. The preference for state government services α_G is set at 0.116 so as to make the optimal state corporate income tax rate in the autarky case equal the

federal personal income tax schedule.

[22] The mean household income in the U.S. is \$75,738 in 2014.

[23] As a related statistic, the share of state corporate income tax revenue out of the sum of state personal and corprate income tax revenue is 12.7%. Thus the value of θ here does not seem extreme. Moreover, the following result does not change much if I use the parametric function for the effective federal personal income tax schedule that is adopted by Benabou (2002) and Heathcote et al. (2014):

$$t_N^F(y_N) = \frac{y_N - \lambda(y_N)^{1-\tau}}{y_N}.$$

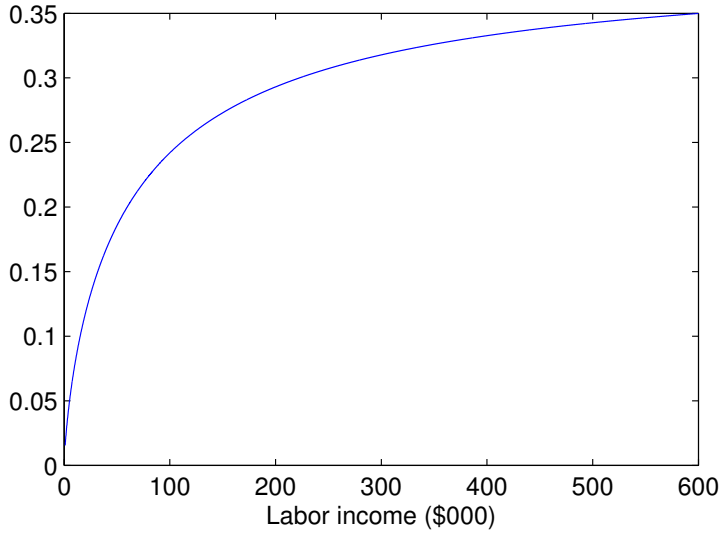


Figure 1.1: The Effective Average Federal Personal Income Tax Rate

average tax rate of U.S. states weighted by state GDP, which is 7.58% in 2014.

1.4.2 Results

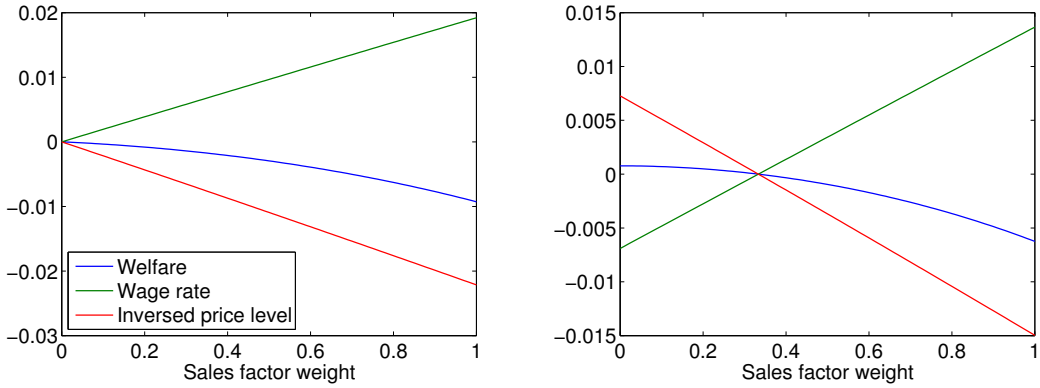
In the calibration exercises, I derive Nash equilibrium of state tax policy by iteration. I focus on the case of symmetric states and homogeneous firms although my model allows heterogeneity in the key variables, including the distribution of productivity of firms across the states. I start with the model of two states, and then extend the model to the economy of 50 states to calibrate the optimal tax policy for an average state in the U.S.

The first column of Table 1.1 reports the optimal state tax policy in Nash equilibrium of the model of two symmetric states. The model suggests the optimal sales apportionment weight is zero: the state governments should use the full payroll weight.

To examine this result, Panel (a) of Figure 1.2 reports how state welfare, the local wage rate, and the local price level change as the sales factor weight in one of the states increases from zero through 100% while the tax policy in the other state is fixed.

Table 1.1: Optimal State Tax Rates in Nash Equilibrium of the Two-State Model

σ	4	3	2	1.5
t_C^n	6.79%	6.50%	6.65%	7.01%
γ_S^n	0	0	0.603	1
t_N^n	7.75%	7.77%	7.04%	5.79%



(a) From the Equilibrium of Unrestricted γ (b) From the Equilibrium of $\gamma = 1/3$

Figure 1.2: The Effects of Sales Factor Weights on Welfare and Aggregate Variables in the Two-State Model

The y-axis represents the rate of change in the variables. As expected, the local wage rate increases as the state puts more weight on the sales factor. If the state adopts the single sales factor, the local wage rate will go up by 1.9%. However, the local price level increases at a faster pace than wage rate at the same time. For example, when the single sales factor is used, the price level goes up by 2.2%. As a whole, the state welfare keeps decreasing as the state raises its sales factor weight. The degree of decrease in welfare is equivalent to a 0.93% decrease in private consumption for the change in sales weight from zero to 100%.

The optimal state corporate income tax rate is lower than personal inecome tax rate due to the negative effect of corporate income tax on local wage rate. However, t_C^n does not deviate far from t_N^n since the distortion in income shifting behavior and the resulting increase in federal tax payment discourage states from reducing t_C much.

Equation 1.12 implies that the local wage rate is more sensitive to t_C^n, or the tax

rate on the payroll factor $t_C^n(1-\gamma_S^n)$ in this context, when σ is lower. This suggests the positive effect of sales apportionment on the local wage may outweigh its negative effect on the price level in case of a lower σ. The second through fourth columns of Table 1.1 report the optimal state tax policy under a few different values of σ that are lower than 4. While the optimal sales weight is still zero even if $\sigma = 3$, the optimal weight increases rapidly to 0.6 and then to 1 as σ goes down to 2 and then to 1.5 respectively. Nonetheless, to obtain a positive γ_S^n as the optimum requires a considerably smaller value of σ than the standard values in the trade literature.

Another observation for the various values of σ is that the relationship between the tax rates for corporate and personal income taxes gets reversed as σ goes down. The optimal corporate income tax rate is higher than the personal income tax rate when $\sigma = 1.5$. When σ is close to one, the markup rate $\sigma/(\sigma-1)$ is very high, which leads to large corporate profits of firms. In this case, a state can export a significant portion of corporate tax liability to the owners of firms who live in the other state. Therefore when σ is close to one, corporate income tax becomes an attractive tool for revenue for states.

Most of the states in the U.S. used to follow the Multistate Tax Compact that recommended states should adopt the equally weighted formula. Thus it is worth examining how Nash equilibirum looks if γ_S^n is fixed at one third rather than considered as one of the choice variables of states. In this case, $t_C^n = 7.16\%$ and $t_N^n = 7.34\%$ in equilibrium. If states start to be allowed to freely choose the value of γ_S^n suddenly, for example due to the ruling by the Court, adopting zero sales weight is the best response in this case too as Panel (b) of Figure 1.2 shows. The panel presents the result of same simulation as Panel (a) but starting from the Nash equilibrium of fixed γ_S^n being equal to one third. Even though the local wage rate will decrease if state lowers γ_S^n from one third, the positive effect of decrease in the price level outweighs the effect on wage rate.

Although the current model has described the economy of two states, it is readily

extended to the economy of many states if they are symmetric. To simulate results for an average state in the U.S., I also calibrate the Nash equilibrium of the model of 50 symmetric states. Table 1.2 reports the optimal tax policy in Nash equilibrium of the 50-state economy for various values of σ. When σ equals 4 or 3, the optimal sales weight is zero again while t_C^n is lower and t_N^n is slightly higher than the two-state model. In this economy, the optimal sales weight is zero even when σ equals as low as 2. If the sales apportionment weight is exogenously fixed at one third, the two tax rates come closer as in the two-state economy: $t_C^n = 6.70\%$ and $t_N^n = 7.22\%$.

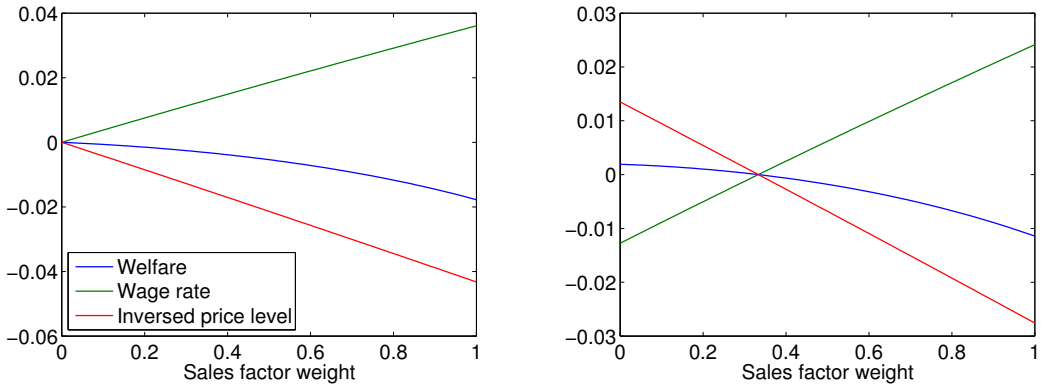
Table 1.2: Optimal State Tax Rates in Nash Equilibrium of the 50-State Model

σ	4	3	2	1.5
t_C^n	6.18%	5.82%	5.16%	7.82%
γ_S^n	0	0	0	1
t_N^n	7.88%	8.02%	8.59%	3.22%

Figure 1.3 examines how state welfare, local wage rate, and local price level change as sales weight deviates from the equilibrium value. The shape of curves in both panels look very similar to those in Figure 1.2. However, the magnitude of changes is almost doubled compared to the two-state case. For example, when the sales apportionment is not restricted, raising the sales weight from zero to one increases wage rate by 3.6% and price level by 4.3%.

1.5 Discussions

Goolsbee and Maydew (2000) is one of the first rigorous empirical studies that estimate the impact of payroll apportionment tax on the local labor demand. They use extensive panel data to find the statistically significant impact of apportionment weights on the local labor demand: if the sales weight is raised from one third to one half and the payroll and property factors are lowered accordingly, manufacturing employment in the



(a) From the Equilibrium of Unrestricted γ (b) From the Equilibrium of $\gamma = 1/3$

Figure 1.3: The Effects of Sales Factor Weights on Welfare and Aggregate Variables in the 50-State Model

state goes up by 1.1%.[24] Their result provides some test of how relevant the model of this study is to the U.S. economy. The dataset of Goolsbee and Maydew (2000) covers the period from 1978 through 1994. Most states used the equally weighted formula at the beginning of the period, and then more and more states started putting a greater weight on the sales factor towards the end of the period. Thus I choose Nash equilibrium of the model of 50 states in which all the states are required to set $\gamma_S^n = 1/3$ and look at the growth rate of local wage rate when one of the states deviates from the equilibrium by raising γ_S^n from one third to one half. My model predicts the growth rate equals 0.62%. This value is not equal to 1.1%, the point estimate of Goolsbee and Maydew (2000), but it is in the same order and within one standard error of their point estimate.[25] Basically, the prediction of my model is not far from their estimate even though the model is meant to be parsimonious for transparency and tractability.

[24]Goolsbee and Maydew (2000) calculate this result based on their estimate for the elasticity of manufacturing employment with respect to state payroll tax burden (-1.92) and the mean corprate income tax rate in their dataset (7.3%). Although the mean tax rate predicted in the equilibrium of my model is slightly different (6.7%), the result of calculation replacing their mean tax rate with the predicted rate is not significantly different from the value of 1.1%.

[25]Moreover, Goolsbee and Maydew (2000) find that the estimate for elasticity of non-manufacturing employment is smaller than manufacturing. If the average of those estimates are taken, their estimate will be closer to the prediction of my model.

The use of some continuous distribution to represent heterogeneity in locations and firms is standard in the recent trade literature.[26] Although the calibration in the previous section assumes homogeneous productivity across states and firms, the model can accommodate such distributions in productivity. Firms may have some idiosyncratic preferences about their location as in Fajgelbaum et al. (2015) and Suárez Serrato and Zidar (2014); for example, an oil company may be more profitable in Alaska, and some kind of agricultural production more suitable in Iowa. If such heterogeneity is brought into my model, the elasticity of the number of firms in a state with respect to its corporate income tax rate will be lessened compared to the case of perfectly homogeneous firms. This leads to a smaller positive effect of reducing the payroll factor tax rate on local labor demand, and the positive effect is less likely to outweigh the negative effect of higher price level. Therefore the assumption of homogeneous firms used in the calibration can be considered as the most favorable for a positive sales weight. Since the calibration suggests there is no point of using a positive sales weight under the assumption of homogeneous firms, a positive sales weight will not be justified in the model of heterogeneous firms either, at least under the baseline value of $\sigma = 4$.

Another restrictive assumption of the model is that workers cannot move across state borders. But this assumption is also the most favorable for a positive sales weight. The mobility of workers can be introduced in the model, for example by heterogeneous preferences of workers regarding locations. The mobility, however, will reduce the gap in local wage rates between states with different tax rates. Thus a decrease in payroll apportionment tax will raise the wage rate less than the model of immobile workers. Moreover, the assumption of immobile workers has another advantage; the model can avoid the issue of how to define the state welfare with changing population.

The complexity of U.S. state tax rules that is beyond the scope of the current

[26]The classic examples from this literature include Eaton and Kortum (2002) and Melitz (2003).

study includes tax nexus and throwback rules. The U.S. state tax rules do not allow a state to impose corporate income tax on a firm unless the firm establishes nexus in the state; for exapmle, a firm does not establish nexus if it has no contact with a state except for soliciting sales of tangible products in the state. In addition, the majority of U.S. states have the throwback rule; under the rule, the sales of tangible goods are counted as sales in the origin state if the seller does not establish nexus in the destination state.

If a perfectly competitive retail industry exists in the economy and if producing firms do not have to perform any business activity in the destination states, then producing firms sell all the goods to retail companies in a state that has zero tax rate for the sales factor. The retail companies distribute the goods to the final consumers across states. If this story is true, firms never report positive sales to states with a positive tax rate for the sales factor as long as there exists a state that adopts zero rate for the sales factor. However, Edmiston and Arze del Granado (2006) examine the data set of corporate income tax returns filed to the State of Georgia by multistate firms and report the share of sales in Georgia is 4.4% in 1992, when the corporate income tax rate and sales apportionment weight in the state were 6% and 1/3 respectively. Moreover, they estimate the sales share in Georgia decreased by 6.3% when the state raised the sales apportionment weight to 1/2. Thus firms do not seem to simply minimize the tax burden on the sales factor by concentrating their nexus in a zero-rate state, but they have nexus in various states for some reasons and respond to a change in apportionment formula of a state in a nontrivial way.

In this context, my model can be interpreted in such a way that firms establish nexus in all the states that they sell goods to, for example by setting up an office at an infinitesimal cost. In reality, probably it is possible for firms in some industries to sell in states without establishing nexus, for example online retail companies, but it is hard for some industries to do so, for example, car makers. Therefore it is necessary to extend

the model to incorporate nexus choice of firms and heterogeneity across industries in terms of production locations if one wants to include these aspects of U.S. tax rules in the analysis.

1.6 Conclusion

Recently, increasing the sales apportionment weight, including adopting the single sales factor, is a popular policy choice among U.S. states, and its effect of stimulating the local labor demand is well recognized in the literature. However, its possible negative effect on the local price level is often overlooked in policy discussions. This paper searches for the optimal state tax policy including apportionment weights and predicts Nash equilibrium of multiple states with a model that incorporates these effects of apportionment formula. Contrary to the popular argument, the model suggests the optimal sales weight should be zero even from the perspective of a single state because a positve sales weight increases the local price level more than the local wage rate. This result is consistent under a wide range of plausible parameters.

Two explanations are possible to reconcile the model's suggestion and the dominant trend among U.S. states. First, the negative effect of sales apportionment tax may actually be overlooked by state policy makers and constituents. The positive effect on local labor demand is often visible; for example, some firms may open new plants or cancel layoffs because of a change in tax policy. However, changes in the price level are harder to detect. It requires a rigorous empirical study to find out the causality between state tax policy and the local price level.

The second possibility is that this model may not capture how tax rules affect firms' decision about production and sales perfectly and overestimate the negative effect. Since the tax rules of U.S. states are complex as discussed in the previous section, firms

may be actively avoiding the tax burden on sales apportionment by adjusting nexus, though not so perfectly as the simple story predicts. To verify these hypotheses, empirical studies are needed to find out how firms allocate and report sales shares across states in practice.

1.7 Appendix

1.7.1 Profit maximization problem for firms

This subsection proves that all the firms in the same state choose the same allocation of sales shares among the states. The proof is a variant of the one by Fajgelbaum et al. (2015). If firm j chooses to operate in state n, the profit maximization problem for firm j is:

$$\begin{aligned} \max_{\{p_n(j),x_n(j)\}} \quad & \tilde{\pi}_n(j) \\ \text{subject to} \quad & x_n^m(j) = p_n^m(j)^{-\sigma} \frac{Y^m}{(P^m)^{1-\sigma}} \quad \text{for } m = \{A,B\}, \end{aligned}$$

in which $\tilde{\pi}_n(j)$ is defined in equation 1.5. Dividing the first order condition of $\tilde{\pi}_n(j)$ with respect to $p_n^A(j)$ by $p_n^A(j)^{-\sigma-1} Y^A (P^A)^{\sigma-1}$ and using the constraint give:

$$\begin{aligned} (1 - t_C^F - \bar{t}_n(j)) \left[(1-\sigma) p_n^A(j) + \sigma(1-\theta_n) \frac{w_N^n}{z_n(j)} \right] + \sigma\theta_n \frac{\tilde{w}_C^n}{z_n(j)} \\ - (1-\sigma) \left[t_S^A - (t_S^A s_n^A(j) + t_S^B s_n^B(j)) \right] \frac{p_n^A(j)}{S_n(j)} \pi_n^T(j) = 0, \quad (1.14) \end{aligned}$$

in which $t_S^m = t_C^m \gamma_S^m$, the effective tax rate for the sales factor in state m; $S_n(j)$ is the total sales of firm j; $s_n^m(j)$ is the share of sales in state m out of $S_n(j)$. Solving equation 1.14

for $p_n^A(j)$ by using $\tilde{w}_C^n = (1 - t_C^F - \bar{t}_n(j))w_N^n$ gives:

$$p_n^A(j) = \frac{1}{1 - \tilde{t}_n^A(\pi_n^T(j)/S_n(j))} \frac{\sigma}{\sigma - 1} \frac{w_N^n}{z_n(j)}, \tag{1.15}$$

in which

$$\tilde{t}_n^A = \frac{t_S^A - (t_S^A s_n^A(j) + t_S^B s_n^B(j))}{1 - t_C^F - \bar{t}_n(j)}. \tag{1.16}$$

The taxable profit can be expressed as:

$$\pi_n^T = S_n(j) \sum_{m=\{A,B\}} s_n^m(j) \left[1 - (1 - \theta_n) \frac{w_N^n}{z_n(j) p_n^m(j)} \right]. \tag{1.17}$$

Substituting equation 1.15 into equation 1.17 gives $\pi_n^T = S_n(j)[1 + \theta_n(\sigma - 1)]/\sigma$. This implies

$$p_n^A(j) = \frac{\sigma}{\sigma - \tilde{t}_n^A[1 + \theta_n(\sigma - 1)]} \frac{\sigma}{\sigma - 1} \frac{w_N^n}{z_n(j)}. \tag{1.18}$$

Finally, note that the sales shares are independent of productivity, $z_n(j)$:

$$\begin{aligned} s_n^A(j) &= \frac{p_n^A(j)^{1-\sigma} \bar{Y}^A}{\sum_{m=\{A,B\}} p_n^m(j)^{1-\sigma} \bar{Y}^m} \\ &= \frac{\{\sigma - \tilde{t}_n^A[1 + \theta_n(\sigma - 1)]\}^{\sigma-1} \bar{Y}^A}{\sum_{m=\{A,B\}} \{\sigma - \tilde{t}_n^m[1 + \theta_n(\sigma - 1)]\}^{\sigma-1} \bar{Y}^m}, \end{aligned} \tag{1.19}$$

in which $\bar{Y}^m = Y^m (P^m)^{\sigma-1}$. The symmetric equation holds for $s_n^B(j)$. Equations 1.16 and 1.19 and the corresponding equations for state B define a system for $\{\tilde{t}_n^m\}$ and $\{s_n^m(j)\}$ whose solution is independent from $z_n(j)$. Threfore $s_n^m(j) = s_n^m$ and $\bar{t}_n(j) = \bar{t}_n$ for all the firms in state n. This result can easily be extended to the case of more than two states.

1.7.2 Proof for Proposition 1

From equations 1.7, 1.8 and 1.9, the problem for the government of state n is:

$$\max_{t_C^n, t_N^n} \quad \int_{i\in n} \frac{\tilde{y}^n(i)}{P^n} (g^n)^{\alpha_G} f^n(i) N^n \mathrm{d}i$$
$$\text{subject to} \quad N^n P^n g^n = z_G^n T^n.$$

If the before-tax aggregate income is normalized to one, as in Section 1.3.1, it can be shown $\Pi = 1/\sigma$ and

$$\tilde{d^n}(i) = (1 - t_C^F - t_C^n) \frac{l(i)}{\sigma N^n L^n}, \tag{1.20}$$

in which Π is the before-tax aggregate markup. By substituting equations 1.4, 1.6 and 1.20, dropping the variables that are exogeneous for the state government, and changing the variable for integration from i to $l(i)$, the state government's problem can be rewritten:

$$\max_{t_C^n, t_N^n} \quad (T^n)^{\alpha_G} \left[\int_{\bar{l}}^{\infty} (1 - t_C^F - t_C^n) w_N^n l f_L^n(l) \mathrm{d}l \right.$$
$$\left. + \int_0^{\bar{l}} (1 - t_N^F(l) - t_N^n) w_N^n l f_L^n(l) \mathrm{d}l + \frac{1 - t_C^F - t_C^n}{\sigma N^n} \right]$$

$$\text{subject to} \quad T^n = t_C^n \int_{\bar{l}}^{\infty} w_N^n N^n l f_L^n(l) \mathrm{d}l + t_N^n \int_0^{\bar{l}} w_N^n N^n l f_L^n(l) \mathrm{d}l + t_C^n \frac{1}{\sigma},$$

in which $f_L^n(l)$ is the probability density function of labor efficiency in state n; $\bar{l}$ is the cutoff labor efficiency for income shifting, which is defined implicitly as $t_N^F(\bar{l}) + t_N^n = t_C^F + t_C^n$. I assume the federal personal income tax schedule $t_N^F(l)$ is differentiable and $\partial t_N^F / \partial l$ is strictly positive everewhere. Dividing the first order condition with respect to

t_N^n by that with respect to t_C^n gives:

$$\frac{\partial T^n/\partial t_N^n}{\partial T^n/\partial t_C^n} = \frac{\int_0^{\bar{l}} w_N^n N^n l f_L^n(l)\mathrm{d}l}{\int_{\bar{l}}^{\infty} w_N^n N^n l f_L^n(l)\mathrm{d}l + 1/\sigma}. \tag{1.21}$$

The partial derivatives of T^n with respect to t_N^n and t_C^n are:

$$\frac{\partial T^n}{\partial t_N^n} = \int_0^{\bar{l}} w_N^n N^n l f_L^n(l)\mathrm{d}l + (t_N^n - t_C^n)\frac{\partial \bar{l}}{\partial t_N^n} w^n N^n \bar{l} f_L^n(\bar{l}); \tag{1.22}$$

$$\frac{\partial T^n}{\partial t_C^n} = \int_{\bar{l}}^{\infty} w_N^n N^n l f_L^n(l)\mathrm{d}l + \frac{1}{\sigma} + (t_N^n - t_C^n)\frac{\partial \bar{l}}{\partial t_C^n} w^n N^n \bar{l} f_L^n(\bar{l}). \tag{1.23}$$

From equations 1.22 and 1.23, note that equation 1.21 holds when $t_N^n = t_C^n$. Therefore the optimal rates for state personal and corporate income taxes are equal in the case of autarky.

1.7.3 Proof for Proposition 2

Since the assumptions made in Proposition 2 constitute a special case of Section 1.7.1, equation 1.18 still holds:

$$p_n^A(j) = \frac{\sigma}{\sigma - \tilde{t}_n^A[1+\theta_n(\sigma-1)]}\frac{\sigma}{\sigma-1}\frac{w_N}{z}. \tag{1.24}$$

From equation 1.16, note $\tilde{t}_n^A = 0$ if $t_C^A = t_C^B$. Then equation 1.24 implies:

$$p_n^A(j) = \left(\frac{\sigma}{\sigma-1}\right)\frac{w_N}{z}. \tag{1.25}$$

Since equation 1.25 holds for all j and does not depend on the state, the first statement of Proposition 2 holds.

Equation 1.16 implies $\tilde{t}_n^B < 0 < \tilde{t}_n^A$ if $t_C^A > t_C^B$. It is straightforward from equation

1.24 to show

$$p_n^B < \left(\frac{\sigma}{\sigma - 1}\right) \frac{w_N}{z} < p_n^A.$$

Therefore the second statement of Proposition 2 holds.

Chapter 2

Risk-Taking and Growth under Taxation with Moral Hazard

Abstract

We study the effect of taxation on entrepreneurs' risk-taking, portfolio choice, and economic growth in the presence of a moral hazard problem in a dynamic general equilibrium model. The model economy is populated by a continuum of agents who invest capital in projects that are subject to aggregate and idiosyncratic investment risks. The moral hazard problem occurs because the return of a type of project depends on entrepreneurs' effort, which is private information. By collecting proportional consumption tax and redistributing the revenue, the government offers an opportunity for risk sharing to encourage risk-taking and spur economic growth. When the moral hazard problem is absent, the full insurance is optimal in terms of welfare, and the economy grows faster. But if the moral hazard problem exists even in a slight degree, risk sharing through taxation cannot improve welfare.

2.1 Introduction

It has long been recognized that taxation can encourage risk-taking since the seminal work of Domar and Musgrave (1944). This somewhat surprising view holds because the government can share the risk of investment as a "partner" for private entrepreneurs. Since entrepreneurial risk-taking is commonly considered as an important driver of macroeconomic growth, it is a natural direction of research to embed this intuition regarding taxation and risk-taking in the context of economic growth. Some of the subsequent studies extended the literature by examining this intuition in stochastic endogenous growth models (Eaton, 1981; Asea and Turnovsky, 1998; Kenc, 2004).

In the literature of corporate finance, the study of entrepreneurial activity stresses the importance of incentive problems, including the moral hazard problem of entrepreneurs. For example, if the performance (*e.g.*, expected return) of entrepreneurial investment depends on the degree of entrepreneur's effort and if the effort is not publicly observable, it will lead to distortion for socially optimal risk-taking. Although it is easy to expect that the moral hazard problem will alter the established conclusions about taxation and risk-taking, however, the literature has not formally incorporated this problem into the analytical framework until recently, especially in the setting of dynamic general equilibrium.[1]

Our study presents quantitative evaluations about how introducing a moral hazard problem changes the policy implications regarding taxation and risk-taking using an overlapping generations, incomplete-market dynamic general equilibrium model. In the model, there is a continuum of ex ante identical agents (entrepreneurs) who are subject

[1]There are some exceptions. The literature has argued for no loss offset rule, which is commonly adopted in many developed countries, based on an incentive problem; the proposed rationale is that entrepreneurs try to mingle personal consumption with investment losses for tax purposes if the full offset is allowed in the tax code, while it is difficult for the government to detect this behaviour (Atkinson and Stiglitz, 2015, Poterba, 2002). Atkinson and Stiglitz (2015), which is originally published in 1980, also present a simple model in which there is information friction between "capitalists" and "managers" and capital income taxation tends to reduce risk-taking (p. 97-99).

to uninsurable idiosyncratic investment risks. Agents can invest their capital in two technologies, one with only aggregate risk (interpreted as the stock market), and another with both aggregate and idiosyncratic risks (interpreted as private equity) in the absence of government intervention. The expected return on private equity investment depends on entrepreneur's effort, which is private information since the effort is not observable. When the government intervenes in the allocation of risk, the choice of the effort level becomes a source of moral hazard problem.

We consider a rather non-traditional policy tool instead of capital income tax, which has drawn attention in the related literature. The government in our model imposes a flat-rate consumption tax and supplies a financial instrument—which we simply refer to as a "government bond"—which is the right for receiving an equal share of government tax revenue in every period in the future. Agents are allowed to trade this government bond at any point in time and allocate their wealth among the two real investment technologies, the government bond, and current consumption. The combination of consumption tax and government bond provides agents with an opportunity to insure themselves against private idiosyncratic risk, for which there is no private insurance market.

If the outcome of private investment did not depend on effort, in other words, if there were no moral hazard problem, the welfare would be maximized by the full insurance through the government fiscal policy. The optimal consumption tax rate would be infinite (equivalent to 100% income tax), and the economy would grow at a higher rate. That is because the risk sharing through government intervention perfectly eliminates the idiosyncratic risk of private equity (but not its aggregate risk) to make the investment less risky, and agents invest more on private equity, which is assumed to have a higher expected return than the stock market. The government intervention improves the welfare of agents by raising the expected utility in the future due to a faster economic growth as

well as by the direct effect of risk sharing.

Once the moral hazard problem is considered, however, the conclusion dramatically changes. With an arguably low elasticity of effort, which is translated as the elasticity of the expected return of private equity with respect to net-of-tax rate to be 0.03 in equilibrium,[2] our quantitative model implies that the optimal tax rate for consumption tax should be 63%, which is far from the full insurance. Since no other tax is imposed in our model, this tax rate can be regarded as being equivalent to 39% proportional income tax. This result is brought about because risk sharing through taxation exacerbates the moral hazard problem and makes agents exert less effort on private equity, even though taxation still successfully encourages agents to allocate a larger share of their wealth to private equity (the riskier asset) rather than the market asset. Moreover, with a slightly higher elasticity (0.06), we find risk-sharing through taxation cannot lead to a Pareto improvement among all the generations. While taxation at a low rate can encourage risk-taking, spur economic growth, and improve the welfare of future generations, the intratemporal wedge caused by the moral hazard problem hinders the welfare improvement for the initial generation. This result sheds light on the important issue that previous studies have not addressed: more risk-taking induced by taxation does not necessarily coincide with welfare improvement. Therefore it is essential to take into accout of the moral hazard problem explicitly when the effects of taxation on risk-taking are studied.

Another feature of our model is that it endogenously generates the cross-sectional wealth distribution that is consistent with the stylized observation, as the stationary distribution of dynamic general equilibrium model that accommodates heterogeneity. Thus we can see how the combination of consumption tax and government bond affects the stationary distribution of wealth. The cross-sectional wealth distribution generated by our model is double Pareto—a distribution with two Pareto tails—and therefore the

[2]For the difinition of net-of-tax rate in this study, see footnote 14 in subsection 2.4.3.

inequality is captured by the Pareto exponent. Because the government intervention enables more risk sharing, it reduces the idiosyncratic volatility in individual portfolios, and therefore reduces inequality (increases the Pareto exponent).

Our paper adds new quantitative insights to the literature of taxation and risk-taking. After Domar and Musgrave (1944) pioneered the research in this area, Mossin (1968) and Stiglitz (1969) formalized the analysis in the expected utility framework. These studies conclude capital income taxation with full loss offset encourages risk-taking in the economy under certain types of utility functions including the Constant Relative Risk Aversion utility function. In addition, Atkinson and Stiglitz (2015) show the conclusion remains true in a broader environment when the tax revenue is redistributed to households, eliminating idiosyncratic risk due to the law of large numbers, rather than is used in a way that it affects households' utility separately (or does not affect at all). While these studies take a partial equilibrium approach,[3] the next generation of studies in this literature examine how taxation affects the allocation of risk between private and public sectors in general equilibrium models, particularly considering that the government budget constraint also becomes stochastic when the government collects revenue from a tax on stochastic returns from investment. Gordon (1985) shows corporate income tax leaves risk-taking in the economy unchanged in a general equilibrium model. He argues that the government can be interpreted as collecting a fraction of expected return from private investment and charging the market price for risk sharing that it offers to investors in his model, in which the government is no better than private markets at spreading risk. The *AK* type endogenous growth models of Eaton, 1981, Asea and Turnovsky, 1998, and Kenc, 2004 assume the government are required to spend a constant fraction of macroeconomic output and imply the government risk sharing does not necessarily increase aggregate risk-taking.[4]

[3] Sandmo (1985) includes detailed review of this strand of literature.

[4] While empirical examination about whether taxation encourages risk-taking is difficult because of

Our study deviates from this line of the literature towards a new direction. In our general equilibrium model, the government can definitely improve the welfare through taxation if there is no moral hazard problem. Since we assume there is no private market for insurance against idiosyncratic risk of private equity investment, the government has a power to offer an insurance against the risk by means of taxation. Instead of allocation of risk between government and private sectors, we focus on how the presence of moral hazard problem undermines the power of government. Our results suggest the moral hazard problem reduces the power of government significantly.

Our paper also provides support for the relevance of the recent literature called New Dynamic Public Finance. The studies in the literature look for the conditions for optimal taxation in the presence of information friction in dynamic settings. Although most of the studies examine cases in which labor skills are private information (Golosov et al., 2003; Kocherlakota, 2005), Albanesi (2011) addresses the optimal taxation problem for entrepreneurs. In her model, entrepreneurs' effort is private information while entrepreneurs are ex ante identical, and it affects the expected return of investment, as in our model. Admittedly our approach is different from that of the New Dynamic Public Finance literature, including Albanesi (2011), notably because we restrict our attention to a specific tax structure while the studies of New Dynamic Public Finance search broader class of tax instruments only considering the constraint of private information.[5] Nonetheless, we view our study has valuable implications in support for this literature. The quantitative results from our model show the optimal tax policy varies substantially according to the degree of moral hazard problem and underscore the importance of taking into account the private information issues when we study the tax policy regarding risky investment.

complicated incentives due to the provisions in the tax code, Cullen and Gordon (2007) provide convincing evidence that tax provisions have large effects on the amount of risk-taking.

[5]This point is well articulated by Golosov et al. (2007).

This paper is also related to the small but growing literature on dynamic general equilibrium models with heterogeneous agents that are analytically solvable and generate heavy-tailed cross-sectional distributions, such as Benhabib et al. (2011) and Benhabib et al. (2016), Toda (2014), and Aoki and Nirei (2014). In all of these papers, agents are subject to uninsurable investment risks, and the upper tail of the cross-sectional wealth distribution obeys the power law (consistent with empirical findings). In Benhabib et al. (2011), agents live for fixed finite periods and are subject to both idiosyncratic capital and labor income risks, and have a bequest motive. They study the effect of estate tax and redistribution on inequality. In Benhabib et al. (2016), agents die at a constant probability of death and receive means-tested subsidies financed by capital taxation. The cross-sectional distribution is *double Pareto*, which has two Pareto tails. The double Pareto distribution has been shown to robustly arise in a large class of dynamic general equilibrium models (Toda, 2014) as well as fit the U.S. income and consumption data well (Toda, 2012; Toda and Walsh, 2015). Aoki and Nirei (2014) jointly study the firm size distribution and top income distribution and obtain the transitory dynamics of the top income share when the marginal tax rate changes.

Our paper is closest to Benhabib et al. (2016) in that the model features a redistributive fiscal policy and a stationary double Pareto distribution. There are important differences, however. First, in our paper the fiscal tools are proportional comsumption tax and tradable government bond as opposed to means-tested subsidies. Second, our model features the moral hazard problem, aggregate shocks, and endogenous asset prices. This point, although it may seem only technical at first glance, is important because without aggregate shocks the government's fiscal policy does not affect the asset span. With aggregate shocks, the government bond will typically increase the asset span and improve welfare. Finally, we study the interaction of moral hazard problem and risk sharing through taxation affects growth, welfare, and inequality and derive the optimal

tax rate.

2.2 Model

The model is a continuous-time stochastic growth model (AK model) with heterogeneous agents and a government.[6] Time is continuous and is denoted by $0 \leq t < \infty$. At $t = 0$, there is a continuum of ex ante identical agents with mass 1. Each agent dies at a constant Poisson rate $\delta > 0$. As soon as an agent dies, a new agent is born (Yaari, 1965; Blanchard, 1985). The agents are not altruistically connected, and there is no bequest motive.

Preference Preferences are represented by the continuous-time analog of the Epstein-Zin constant relative risk aversion (CRRA), constant elasticity of intertemporal substitution (CEIS) utility function with discount rate $\beta > 0$, relative risk aversion $\gamma > 0$, and elasticity of intertemporal substitution $\varepsilon > 0$ with disutility from effort. More precisely, the continuation value J satisfies

$$J(t, w(t)) = \mathrm{E}_t \int_t^\infty f(c(s)h(e(s)), J(s, w(s)))\mathrm{d}s, \tag{2.1}$$

where

$$f(x, y) = \frac{\beta + \delta}{1 - 1/\varepsilon} \frac{x^{1-1/\varepsilon} - ((1-\gamma)y)^{\frac{1-1/\varepsilon}{1-\gamma}}}{((1-\gamma)y)^{\frac{\gamma-1/\varepsilon}{1-\gamma}}} \tag{2.2}$$

is the aggregator,[7] $c(t)$ is consumption, $w(t)$ is wealth, and $h(e(t))$ is the disutility from effort $e(t)$.

[6] Saito (1998) studies the asset pricing implication of idiosyncratic investment risks in a similar settings. Toda (2014) studies a general AK model in discrete time without a government. Toda (2015) considers the optimal taxation of physical and human capital in a Markov setting but without intergenerational considerations or moral hazard issues.

[7] See Duffie and Epstein (1992) for details.

Technologies There are two types of technologies in which agents can invest their capital. The first is the aggregate stock market, which is interpreted as a stochastic constant-returns-to-scale (AK) technology. The capital invested in the stock market evolves according to

$$\mathrm{d}K/K = \mu_m \mathrm{d}t + \sigma_m \mathrm{d}B_m, \tag{2.3}$$

where μ_m is the expected return, $\sigma_m > 0$ is volatility, and B_m is a standard Brownian motion.

The second is the private equity, which is also an AK technology but requires effort to operate. The capital invested in agent i's private equity evolves according to

$$\mathrm{d}K/K = \mu_p(e)\mathrm{d}t + \sigma_p \mathrm{d}B_p + v\mathrm{d}B_i, \tag{2.4}$$

where $\mu_p(e)$ is the expected return when the effort level is e, B_p and B_i are independent standard Brownian motions that represent the aggregate and idiosyncratic shock to capital, and $\sigma_p, v > 0$ are their volatilities. Each agent can only invest in his own private equity. B_i is assumed to be i.i.d. across agents and is uninsurable. Note that only the expected return $\mu_p(e)$ depends on the effort level. This is because if the idiosyncratic volatility v also depends on effort, since v is observable by calculating the quadratic variation of the value of invested capital, it is possible to back-out the effort level. Thus it is necessary that v is independent of e so that e remains unobservable.[8]

The stock market and the private equity may well be correlated: let

$$\mathrm{d}B_m \mathrm{d}B_p = \rho \mathrm{d}t,$$

[8]There is an additional reason that we leave v independent of effort. Even if less effort were associated with larger v for the purpose of representing the moral hazard problem in another dimension, it would not harm the attractivenss of government bond as a financial asset; the government can perfectly eliminate the fraction of variation in its tax revenue that is attributed to the idiosyncratic risk of private equity by the law of large numbers because of the i.i.d. assumption.

where $-1 < \rho < 1$ is the correlation coefficient between the stock market and the aggregate component of private equity.

This specification of the technologies is not the most general one for which we can obtain closed-form solutions. The model is equally tractable even if aggregate capital and private equity jointly enters the production function nonlinearly, provided that it is homogeneous of degree 1.[9] The reason why we assume that the stock market and private equity are each AK technologies is because with nonlinearities, the equilibrium becomes constrained inefficient (Toda, 2014), and hence the government can improve welfare even without introducing new assets. With AK technologies and no effort choice, the equilibrium is constrained efficient (Toda, 2014), so any welfare improvement from the government intervention will be entirely due to the expansion in the asset span.

Government bond To provide agents with an opportunity to insure against the idiosyncratic investment risk, for which there is no private market, the government issues bond and redistributes tax revenue to the owners of bond. Although the government bond is handed out equally to each agent in the economy at the initial time point, agents can trade government bond afterwards. While the simplest way of redistribution would be to give an equal share of tax revenue to each agent in every period over lifetime (demogrant or "basic income"), the tradable government bond allows us to focus on the pure effect of extending the asset span by government fiscal policy on entrepreneurial risk-taking without wealth effect.

In our model, the choice of tax strucute is also different from the standard approach in the literature on taxation and risk-taking. The government imposes proportional consumption tax on every agent rather than capital income tax. The reason for this choice is to make information friction as transparent as possible. The rules about loss offset has

[9] See Angeletos (2007) and Toda (2014) for such models.

been one of the most discussed topic in the literature since Domar and Musgrave (1944). Although the studies mostly agree that allowing loss offset encourages risk-taking, at least in a partial equilibrium setting, many developed countries adpot no offset rule. The common rational for no offset rule proposed in the literature is that it is difficult for government to distinguish between losses from risky investment and private consumption (Atkinson and Stiglitz, 2015; Poterba, 2002). Thus capital income tax would cause another type of moral hazard problem in which agents try to take advantage of loss offset. On the contrary, consumption tax is immune to this type of problem and allows us to focus on the moral hazard problem due to entrepreneurial effort. A consumption tax (value-added tax) is also attractive from an administrative point of view (especially in developing countries) since the invoice system discourages tax evasion at a low administrative cost, at least as long as the tax rate is not so high that people start to engage in barter trades. It is a formidable task for the government to capture captial income of individuals from all sources, espcially in the era of global integration. In fact, the value-added tax is a major source of revenue in most developing and developed countries (except in U.S., which does not have a nation-wide value added tax). Another advantage of taxing consumption is that it is a better proxy of permanent income than just income.

Throughout this paper, we assume that the government taxes consumption at a flat rate τ and distribute the entire tax revenue to the owners of government bond, whose unit is normalized such that the aggregate supply equals unity at all time. Therefore if aggregate consumption is C, the dividend of government bond is τC per unit.

Financial assets and transfers In addition to the two types of physical capital and government bond, agents can trade a risk-free asset in zero net supply, whose risk-free rate r is determined in equilibrium. If there is no aggregate risk (so $\sigma_m = \sigma_p = 0$), then the government bond will be risk-free and therefore will have the same rate as the

risk-free asset.

As mentioned earlier, each agent dies at constant Poisson rate $\delta > 0$ and gets reborn. Assets are transfered between deceased and newborn agents as follows. First, there is an annuity market for stock and risk-free asset: an agent receives an annuity δ per asset holdings when alive and transfers the asset to the annuity company upon death. This arrangement is optimal from agents' point of view and the annuity company breaks even. Second, private equity cannot be pledged for annuity, possibly because in reality the value of private equity may be difficult to observe or evaluate, or the value of a private business may be highly dependent on the ability of the entrepreneur. When an agent dies, all of her capital in private equity is transfered to the government, which in turn is split equally among newborn agents. This transfer might be broadly interpreted as public goods such as mandatory elementary education. Third, the government bond is not pledgeable for annuity either. Although the government bonds are tradable among agents, all shares must be returned to the government upon death.[10] The government then gives 1 share of the bond to each newborn agent.

Equilibrium The definition of the competitive equilibrium is standard: given the consumption tax rate τ, an equilibrium consists of individual choices (consumption, effort, and portfolio) and asset prices (risk-free asset and government bond) such that (i) individual choices are optimal, (ii) asset markets clear, and (iii) the government budget is balanced.

[10]This rule might be difficult to enforce in practice, since if agents have some private information that predicts death (*e.g.*, sickness), those agents will have the incentive to sell off the bond before death. A practical solution would be to make the bonds nontradable (which is actually what is happening in reality since social security benefits cannot be used as collateral), although it will make the theoretical analysis difficult since we lose homogeneity. At least theoretically, making the bonds tradable and requiring to return upon death cause no difficulty since the death rate δ is constant.

2.3 Solving for equilibrium

This section explains how to solve for the equilibrium. Let $C(t)$ denote the aggregate consumption. From the structure of the model, C is some geometric Brownian motion

$$\mathrm{d}C/C = \mu_c \mathrm{d}t + \sigma_c \mathrm{d}B_c, \tag{2.5}$$

where the drift μ_c and volatility σ_c are determined in equilibrium and B_c is some combination of the fundamental shocks B_m and B_p. Let $P(t)$ be the price of the government bond and $D(t)$ be the dividend. Since consumption is taxed at a flat rate τ and the tax revenues are paid out as dividend to the government bond, we have

$$D(t) = \tau C(t).$$

From the structure of the model, the government bond price must be proportional to aggregate consumption. Therefore $P(t)$ follows the same geometric Brownian motion as aggregate consumption (except for the difference in levels), (2.5). Since the government bond pays out dividend, its instantaneous total return is

$$\frac{\mathrm{d}P}{P} + \frac{D}{P}\mathrm{d}t = (\mu_c + d)\mathrm{d}t + \sigma_c \mathrm{d}B_c, \tag{2.6}$$

where $d = D/P$ is the dividend yield to be determined in equilibrium. Let $\mu_g = \mu_c + d$ be the expected total return on the government bond.

2.3.1 Individual problem

By (2.3), (2.4), and (2.6), the budget constraint of agent i is

$$\begin{aligned} \mathrm{d}w_i/w_i = ((\mu_m+\delta)\theta_1+\mu_p(e)\theta_2+\mu_g\theta_3)\mathrm{d}t+(1-\theta_1-\theta_2-\theta_3)(r+\delta)\mathrm{d}t \\ +\theta_1\sigma_m\mathrm{d}B_m+\theta_2(\sigma_p\mathrm{d}B_p+v\mathrm{d}B_i)+\theta_3\sigma_c\mathrm{d}B_c-(1+\tau)m\mathrm{d}t, \quad (2.7) \end{aligned}$$

where m is the propensity to consume out of wealth ($m = c_i/w_i$, where c_i is the consumption rate), w_i is wealth, $(\theta_1,\theta_2,\theta_3)$ are the fraction of wealth invested in the stock, private equity, and government bond, r is the risk-free rate, δ is the insurance premium on mortality risk, and τ is the consumption tax rate. Note that δ appears only in the terms corresponding to the stock market and the risk-free asset because those are the only assets insurable against the mortality risk. For notational simplicity, let

$$\theta = \begin{bmatrix} \theta_1 \\ \theta_2 \\ \theta_3 \end{bmatrix}, \ \mu(e) = \begin{bmatrix} \mu_m+\delta \\ \mu_p(e) \\ \mu_g \end{bmatrix}, \ \mathrm{d}X = \begin{bmatrix} \sigma_m\mathrm{d}B_m \\ \sigma_p\mathrm{d}B_p+v\mathrm{d}B_i \\ \sigma_c\mathrm{d}B_c \end{bmatrix},$$

and define the instantaneous variance matrix Σ by $(\mathrm{d}X)(\mathrm{d}X)' = \Sigma\mathrm{d}t$. Then the budget constraint (2.7) becomes

$$\mathrm{d}w_i/w_i = (r+\delta+\theta'(\mu(e)-r-\delta)-(1+\tau)m)\mathrm{d}t+\theta'\mathrm{d}X. \quad (2.8)$$

Maximizing the recursive utility defined by (2.1) and (2.2) subject to the budget constraint (2.8) is a standard Merton (1971)-type optimal consumption-portfolio problem except for the presence of an effort choice, a consumption tax, and recursive preferences.[11] Before solving the individual problem, we impose a few equilibrium conditions in order

[11] Svensson (1989) solves the optimal portfolio problem with Epstein-Zin preferences.

to simplify the computation of equilibrium. First, since the utility function as well as the budget constraint (2.8) are homothetic in wealth, the optimal consumption rate, effort, and portfolio choice are common across all agents. Since the risk-free asset is in zero net supply, there will be no trade in the risk-free asset in equilibrium. Therefore it must be $\sum_{j=1}^{3} \theta_j = 1$. Furthermore, since technologies cannot be operated in reverse and the government bond is in positive net supply, we have $\theta_j \geq 0$ for all j. Let $\Delta = \left\{ \theta \in \mathbb{R}^3_+ \,\middle|\, \sum_{j=1}^{3} \theta_j = 1 \right\}$ be the set of admissible portfolios.

The Hamilton-Jacobi-Bellman (HJB) equation of the optimal consumption-effort-portfolio problem is

$$0 = \max_{\theta, m, e} \left[\mathcal{D}J + f(mh(e)w, J) \right], \tag{2.9}$$

where

$$\mathcal{D}J = J_t + wJ_w(\theta' \mu(e) - (1+\tau)m) + \frac{1}{2} w^2 J_{ww} \theta' \Sigma \theta.$$

(Here the risk-free rate does not appear because $\theta \in \Delta$.) By the homotheticity of the utility function, it should be clear that the value function takes the form

$$J(w) = \frac{a^{1-\gamma}}{1-\gamma} w^{1-\gamma}$$

for some $a > 0$. Substituting this functional form into (2.9) and retaining terms that contain θ, the optimal portfolio is the maximizer of

$$\mu(e)' \theta - \frac{\gamma}{2} \theta' \Sigma \theta, \tag{2.10}$$

which is quadratic in $\theta \in \Delta$.

The first-order condition with respect to the consumption rate m is

$$
\begin{aligned}
&-(1+\tau)wJ_w + \frac{\beta+\delta}{((1-\gamma)J)^{\frac{\gamma-1/\varepsilon}{1-\gamma}}}(mh(e)w)^{-1/\varepsilon}hw = 0 \\
\iff m &= \left(\frac{\beta+\delta}{1+\tau}\right)^{\varepsilon} \frac{h(e)^{\varepsilon-1}}{w\left[J_w((1-\gamma)J)^{\frac{\gamma-1/\varepsilon}{1-\gamma}}\right]^{\varepsilon}} = \left(\frac{\beta+\delta}{1+\tau}\right)^{\varepsilon}\left(\frac{a}{h(e)}\right)^{1-\varepsilon}.
\end{aligned}
$$

Substituting this optimal m into the HJB equation (2.9), after some algebra we obtain

$$
a = (\beta+\delta)^{\frac{1}{1-1/\varepsilon}}\frac{h(e)}{1+\tau}\left[\varepsilon(\beta+\delta) + (1-\varepsilon)\left(\mu(e)'\theta - \frac{\gamma}{2}\theta'\Sigma\theta\right)\right]^{\frac{1}{1-\varepsilon}}. \tag{2.11}
$$

The optimal consumption rate is then

$$
m = \frac{1}{1+\tau}\left[\varepsilon(\beta+\delta) + (1-\varepsilon)\left(\mu(e)'\theta - \frac{\gamma}{2}\theta'\Sigma\theta\right)\right]. \tag{2.12}
$$

Since J is increasing in a, the optimal effort is the maximizer of (2.11). There is a trade-off here. Increasing e will lower the utility from leisure $h(e)$, whereas increase the expected return $\mu(e)$.

2.3.2 Equilibrium conditions

Substituting the optimal consumption rule (2.12) into the budget constraint (2.7), individual wealth evolves according to

$$
\mathrm{d}w_i/w_i = \left(\varepsilon(\mu(e)'\theta - \beta - \delta) + (1-\varepsilon)\frac{\gamma}{2}\theta'\Sigma\theta\right)\mathrm{d}t + \theta'\mathrm{d}X. \tag{2.13}
$$

The aggregate wealth obeys a similar diffusion. The only difference is that individuals that are alive receive annuity δ on the fraction of wealth $1-\theta_2-\theta_3=\theta_1$ invested in the stock and the risk-free asset, but the annuity cancel out with the amount the annuity company

collects from deceased agents. Therefore the aggregate wealth evolves according to

$$\mathrm{d}W/W = \left(\varepsilon(\mu(e)'\theta - \beta - \delta) + (1-\varepsilon)\frac{\gamma}{2}\theta'\Sigma\theta - \delta\theta_1\right)\mathrm{d}t + \theta'\mathrm{d}\bar{X}, \tag{2.14}$$

where $\mathrm{d}\bar{X} = (\sigma_m \mathrm{d}B_m, \sigma_p \mathrm{d}B_p, \sigma_c \mathrm{d}B_c)'$ is the vector of aggregate shocks. Note that (2.14) differs from (2.13) in only two aspects: first, in (2.14) $\mathrm{d}X$ is replaced by $\mathrm{d}\bar{X}$, in which the idiosyncratic risk disappears by the law of large numbers; second, (2.14) contains the term $-\delta\theta_1$, which accounts for the annuity payment on the stock and the risk-free asset.

Since consumption is proportional to wealth, individual consumption also obeys the diffusion (2.13), and aggregate consumption obeys (2.14). Therefore (2.14) must be consistent with (2.5). Comparing coefficients, it must be the case that

$$\mu_c = \varepsilon(\mu(e)'\theta - \beta - \delta) + (1-\varepsilon)\frac{\gamma}{2}\theta'\Sigma\theta - \delta\theta_1, \tag{2.15a}$$

$$\sigma_c \mathrm{d}B_c = \theta'\mathrm{d}\bar{X} \iff \sigma_c \mathrm{d}B_c = \frac{1}{1-\theta_3}(\theta_1\sigma_m \mathrm{d}B_m + \theta_2\sigma_p \mathrm{d}B_p). \tag{2.15b}$$

Finally, since in equilibrium fraction θ_3 of wealth is in government bond and there is a unit supply of the bond, its price must satisfy

$$P = \theta_3 W.$$

Therefore the dividend yield is

$$d = \frac{D}{P} = \frac{\tau m W}{\theta_3 W} = \frac{\tau m}{\theta_3}, \tag{2.16}$$

where m is the optimal consumption rate given by (2.12).

We can obtain the equilibrium as follows.

1. Let $\mathbf{z} = (z_1, z_2, z_3, z_4) = (\mu_g, e, \theta_1, \theta_2)$ be the equilibrium objects, where μ_g is the

expected total return on the government bond, e is the effort, and θ_1, θ_2 are the fraction of wealth invested in the stock market and private equity.

2. Let $\Sigma(\mathbf{z})$ be the variance matrix of $\mathrm{d}X = (\sigma_m \mathrm{d}B_m, \sigma_p \mathrm{d}B_p + v\mathrm{d}B_i, \sigma_c \mathrm{d}B_c)'$ with $\sigma_c \mathrm{d}B_c$ replaced by (2.15b) with

$$\theta = (\theta_1, \theta_2, 1 - \theta_1 - \theta_2) = (z_3, z_4, 1 - z_3 - z_4).$$

3. Let $\mu(\mathbf{z}) = (\mu_m + \delta, \mu_p(e), \mu_g)'$ be the vector of expected returns (note that $\mu_g = z_1$ and $e = z_2$), $m(\mathbf{z})$ be the optimal consumption rate in (2.12) with $\mathbf{z}$ substituted, and $a(\mathbf{z})$ be right-hand side of (2.11) with $\mu(e)$, Σ replaced by $\mu(\mathbf{z})$ and $\Sigma(\mathbf{z})$, respectively. Define $F : \mathbb{R}^4 \to \mathbb{R}^4$ by

$$F_1(\mathbf{z}) = \varepsilon(\mu(\mathbf{z})'\theta - \beta - \delta) + (1 - \varepsilon)\frac{\gamma}{2}\theta'\Sigma\theta - \delta\theta_1 + \frac{\tau m(\mathbf{z})}{\theta_3}, \tag{2.17a}$$

$$F_2(\mathbf{z}) = \arg\max_{e = z_2} a(\mathbf{z}), \tag{2.17b}$$

$$F_3(\mathbf{z}) = \theta_1, \tag{2.17c}$$

$$F_4(\mathbf{z}) = \theta_2, \tag{2.17d}$$

where $\theta = (\theta_1, \theta_2, \theta_3)'$ is the arg max of (2.10) over $\theta \in \Delta$ with $\mu(e) = \mu(\mathbf{z})$ and $\Sigma = \Sigma(\mathbf{z})$. F_1 defines the expected total return on government bond, which is the sum of aggregate consumption growth in (2.15a) and the dividend yield in (2.16). F_2 defines the optimal effort choice. F_3 and F_4 define the optimal portfolio.

4. The equilibrium is a fixed point of the mapping $F : \mathbb{R}^4 \to \mathbb{R}^4$.

The equation $F_1(\mathbf{z}) = z_1$ is linear and therefore it can be solved by hand. After some

tedious algebra, the result is

$$z_1 = \mu_g = \left(1 - \varepsilon\theta_3 - \frac{\tau}{1+\tau}(1-\varepsilon)\right)^{-1} \left(\mu - \delta\theta_1 \right.$$
$$\left. - \left(1 - \frac{\tau}{(1+\tau)\theta_3}\right)\left(\varepsilon(\beta+\delta) + (1-\varepsilon)\left(\mu - \frac{\gamma}{2}\theta'\Sigma\theta\right)\right)\right), \quad (2.18)$$

where $\mu = (\mu_m + \delta)\theta_1 + \mu_p(e)\theta_2$.

2.3.3 Welfare

Assume that agents are ex ante identical, so they all start with initial capital normalized to 1 at $t = 0$. Letting P_0 be the initial price of the government bond and W_0 be the initial aggregate wealth (physical capital plus the market capitalization of the government bond), we have

$$P_0 = \theta_3 W_0.$$

Since the government grants 1 share of the bond to every agent, the initial aggregate wealth is

$$W_0 = 1 + P_0.$$

Combining these two equations, we get

$$W_0 = \frac{1}{1-\theta_3}. \quad (2.19)$$

The value function of an agent with wealth w is $J(w) = \frac{1}{1-\gamma}(aw)^{1-\gamma}$, where a is given by (2.11). Since $J(w)$ is a monotonic transformation of aw, the welfare of an agent at $t = 0$ is

$$V_0 = aW_0 = \frac{a}{1-\theta_3} \quad (2.20)$$

in consumption equivalent.

The welfare criterion for the future generations are more complicated. Let w_t be the initial wealth of an agent born at t, and W_t be the aggregate wealth. Since fraction of wealth θ_1 of a deceased agent goes to the annuity industry, we have $w_t = (1-\theta_1)W_t$. Therefore the certainty equivalent of the generation t welfare is

$$V_t = \mathrm{E}\left[(1-\gamma)\frac{1}{1-\gamma}(aw_t)^{1-\gamma}\right]^{\frac{1}{1-\gamma}} = a(1-\theta_1)\,\mathrm{E}[W_t^{1-\gamma}]^{\frac{1}{1-\gamma}}. \tag{2.21}$$

Since W_t is a geometric Brownian motion with drift μ_c and volatility σ_c, by Itô's lemma we have

$$\mathrm{d}\log W = \left(\mu_c - \frac{\sigma_c^2}{2}\right)\mathrm{d}t + \sigma_c \mathrm{d}B_c.$$

Hence $\log W_t - \log W_0 \sim N((\mu_c - \sigma_c^2/2)t, \sigma_c^2 t)$, so using the moment generating function of the normal distribution, we have

$$\mathrm{E}[W_t^{1-\gamma}] = W_0 \exp\left(\left(\mu_c - \frac{\sigma_c^2}{2}\right)t(1-\gamma) + \frac{\sigma_c^2 t}{2}(1-\gamma)^2\right). \tag{2.22}$$

Combining (2.19), (2.21), and (2.22), we obtain

$$V_t = \frac{a(1-\theta_1)}{1-\theta_3}\exp\left(\left(\mu_c - \frac{\gamma\sigma_c^2}{2}\right)t\right). \tag{2.23}$$

Thus there is a conflict of interest across generations. All generations care about $\frac{a}{1-\theta_3}$, which is essentially the coefficient of the value function, taking initial aggregate capital as given. The initial generation cares only about this quantity. However, the future generations also care about the proportion of wealth they inherit, $1-\theta_1$, and the risk-adjusted economic growth rate $\mu_c - \frac{\gamma\sigma_c^2}{2}$. The generations born right after $t = 0$ care only about $1-\theta_1$ (on top of $\frac{a}{1-\theta_3}$), while the generations in the far distant future care only about $\mu_c - \frac{\gamma\sigma_c^2}{2}$ because the growth dominates any difference in the initial endowment.

Therefore if the government wishes to make a Pareto improvement, it must simultaneously increase the three quantities

$$\frac{a}{1-\theta_3}, \quad \frac{a(1-\theta_1)}{1-\theta_3}, \quad \mu_c - \frac{\gamma\sigma_c^2}{2}. \tag{2.24}$$

2.3.4 Stationary distribution

To compute the stationary distribution of the economy, using (2.5), rewrite (2.13) and (2.14) as

$$\begin{aligned} \mathrm{d}w_i/w_i &= (\mu_c + \delta\theta_1)\mathrm{d}t + \sigma_c \mathrm{d}B_c + \theta_2 v \mathrm{d}B_i, \\ \mathrm{d}W/W &= \mu_c \mathrm{d}t + \sigma_c \mathrm{d}B_c. \end{aligned}$$

By Itô's formula and $\mathrm{d}B_c \mathrm{d}B_i = 0$, we obtain

$$\begin{aligned} \mathrm{d}\log w_i &= \left(\mu_c + \delta\theta_1 - \frac{\sigma_c^2 + \theta_2^2 v^2}{2}\right)\mathrm{d}t + \sigma_c \mathrm{d}B_c + \theta_2 v \mathrm{d}B_i, \\ \mathrm{d}\log W &= \left(\mu_c - \frac{\sigma_c^2}{2}\right)\mathrm{d}t + \sigma_c \mathrm{d}B_c. \end{aligned}$$

Taking the difference, we obtain

$$\mathrm{d}\log(w_i/W) = \left(\delta\theta_1 - \frac{\theta_2^2 v^2}{2}\right)\mathrm{d}t + \theta_2 v \mathrm{d}B_i. \tag{2.25}$$

Therefore the logarithm of individual wealth relative to aggregate wealth evolves according to a Brownian motion.

We can solve for the entire wealth dynamics following Toda (2014).[12] For

[12] See also Gabaix (2009) for a method to solve for the stationary distribution by using the Fokker-Planck equation (Kolmogorov forward equation).

notational simplicity, let

$$\mu = \delta\theta_1 - \frac{\theta_2^2 v^2}{2},$$
$$\sigma = \theta_2 v,$$
$$m = \log(\theta_2 + \theta_3).$$

Fix time $t > 0$. For agents with age t (*i.e.*, those already born at $t = 0$), since agents are ex ante identical, the initial wealth relative to aggregate wealth is $w_i(0)/W(0) = 1$ (which is 0 in logs). Since the death rate is δ, there is a mass $\mathrm{e}^{-\delta t}$ of such agents. For agents with age $s < t$, since they receive wealth from private equity of deceased agents and one share of the government bond, their initial wealth relative to aggregate wealth must be $\theta_2 + \theta_3$ (which is $\log(\theta_2 + \theta_3)$ in logs), the fraction of wealth a typical agent invests in the private equity and the government bond. Since log relative wealth obeys the Brownian motion, the cross-sectional mean and variance increases linearly with age. Therefore the cross-sectional density of log relative wealth at t is the normal mixture

$$\begin{aligned} f(x,t) &= \mathrm{e}^{-\delta t}\frac{1}{\sigma\sqrt{\pi t}}\mathrm{e}^{-\frac{(x-\mu t)^2}{2\sigma^2 t}} + \int_0^t \delta \mathrm{e}^{-\delta s}\frac{1}{\sigma\sqrt{\pi s}}\mathrm{e}^{-\frac{(x-m-\mu s))^2}{2\sigma^2 s}}\,\mathrm{d}s \\ &= \mathrm{e}^{-\delta t}\frac{1}{\sigma\sqrt{\pi t}}\mathrm{e}^{-\frac{(x-\mu t)^2}{2\sigma^2 t}} \\ &\quad + \frac{\delta \mathrm{e}^{\frac{\mu(x-m)}{\sigma^2}}}{\kappa\sigma}\left[\mathrm{e}^{-\frac{\kappa|x-m|}{\sigma}}\Phi\left(-\frac{|x-m|}{\sigma\sqrt{t}} + \kappa\sqrt{t}\right) - \mathrm{e}^{\frac{\kappa|x-m|}{\sigma}}\Phi\left(-\frac{|x-m|}{\sigma\sqrt{t}} - \kappa\sqrt{t}\right)\right], \end{aligned}$$

where $\kappa = \sqrt{2\delta + (\mu/\sigma)^2}$ and Φ is the cumulative distribution function of the standard normal distribution.[13] As $t \to \infty$, all terms except the second converge to 0. Therefore

[13] See Proposition 17 in Toda (2014) and its proof for the details of this derivation.

the stationary distribution is

$$\begin{aligned} f(x) = \lim_{t\to\infty} f(x,t) &= \frac{\delta}{\kappa\sigma} e^{-\frac{\mu(x-m)}{\sigma^2} - \frac{\kappa|x-m|}{\sigma}} \\ &= \begin{cases} \frac{\alpha\zeta}{\alpha+\zeta} e^{-\alpha|x-m|}, & (x > m) \\ \frac{\alpha\zeta}{\alpha+\zeta} e^{-\zeta|x-m|}, & (x \leq m) \end{cases} \end{aligned} \tag{2.26}$$

where

$$\alpha, \zeta = \sqrt{\frac{2\delta}{\sigma^2} + \frac{\mu^2}{\sigma^4}} \mp \frac{\mu}{\sigma^2}.$$

Note that $-\alpha$ and ζ are solutions to the quadratic equation

$$\varphi(\lambda) := \frac{\sigma^2}{2}\lambda^2 - \mu\lambda - \delta = \frac{\theta_2^2 v^2}{2}\lambda^2 - \left(\delta\theta_1 - \frac{\theta_2^2 v^2}{2}\right)\lambda - \delta = 0.$$

Since $\varphi(0) = -\delta < 0$ and $\varphi(-1) = \delta(\theta_1 - 1) < 0$, we have

$$-\alpha < -1 < 0 < \zeta \iff \alpha > 1, \zeta > 0.$$

The distribution (2.26) is known as *Laplace* (Kotz et al., 2001). Consequently, the distribution of relative wealth w_i/W, which is the exponential of log relative wealth, is *double Pareto* (Reed, 2001; Toda, 2012) with power law exponents α, ζ and location parameter (mode if $\zeta > 1$) $e^m = \theta_2 + \theta_3$. The precise density function of relative wealth is

$$f_{w/W}(x) = \begin{cases} \frac{\alpha\zeta}{\alpha+\zeta}(\theta_2+\theta_3)^{\alpha} x^{-\alpha-1}, & (x > \theta_2+\theta_3) \\ \frac{\alpha\zeta}{\alpha+\zeta}(\theta_2+\theta_3)^{-\zeta} x^{\zeta-1}. & (x \leq \theta_2+\theta_3) \end{cases}$$

Note that since $\alpha \to 1$ as $\delta \to 0$, we recover Zipf (1949)'s law when the death probability gets smaller (agents live longer). The fact that the stationary cross-sectional distribution becomes double Pareto is not surprising because Toda (2014) shows that it arises robustly

in a large class of dynamic general equilibrium models. However, the magnitude of the tail exponent α is model-dependent. Here we have $\alpha > 1$ and $\alpha \to 1$ as $\delta \to 0$, consistent with Zipf's law. In Benhabib et al. (2016), where there are inheritance and means-tested government subsidies, the tail exponent (in their notation β_2) is > 2.

Toda (2012) calculates the Gini coefficient for the double Pareto distribution explicitly. Using

$$\begin{aligned}
\alpha\zeta &= \frac{2\delta}{\sigma^2} = \frac{2\delta}{\sigma^2}, \\
\zeta - \alpha &= \frac{2\mu}{\sigma^2} = \frac{2\delta\theta_1}{\sigma^2} - 1, \\
\alpha + \zeta &= 2\sqrt{\frac{2\delta}{\sigma^2} + \frac{\mu^2}{\sigma^4}} = \sqrt{1 - \frac{4\delta}{\sigma^2}(2-\theta_1) + \frac{4\delta^2\theta_1^2}{\sigma^4}},
\end{aligned}$$

the result is

$$\begin{aligned}
G_{w/W} &= \frac{2\alpha^2 + 2\alpha\zeta + 2\zeta^2 + \zeta - \alpha}{(\alpha+\zeta)(2\alpha-1)(2\zeta+1)} \\
&= \frac{2(\zeta-\alpha)^2 + 4\alpha\zeta + (\zeta-\alpha)}{(\alpha+\zeta)(4\alpha\zeta - 2(\zeta-\alpha) - 1)} \\
&= \frac{1 + \frac{\delta}{\sigma^2}(8 - 6\theta_1) + \frac{4\delta^2\theta_1^2}{\sigma^4}}{(1 + \frac{4\delta}{\sigma^2}(2-\theta_1))\sqrt{1 + \frac{4\delta}{\sigma^2}(2-\theta_1) + \frac{4\delta^2\theta_1^2}{\sigma^4}}},
\end{aligned}$$

where $\sigma = \theta_2 v$. The Gini coefficient approaches 1 (extreme inequality) as $\delta \to 0$.

2.4 Numerical example

Since in our model the government has an ability to offer an insurance through taxation against the risk for which there is no private market (the idiosyncratic risk of private equity), if the moral hazard problem did not occur, the welfare of agents would be maximized by full insurance: the optimal consumption tax rate would be

infinite, which is equivalent to 100% income tax. The main interest of current study is quantitative evaluation for the implication of moral hazard problem: with a small elasticity of entrepreneur's effort, whether it is still possible to improve the welfare by risk sharing through taxation and, if possible, how much the optimal tax rate is affected. Our model is solvable and suitable for numerical simulations while the model is elaborate enough to answer questions regarding welfare, portfolio allocation (risk-taking), economic growth, and wealth distribution in the presence of moral hazard problem. The results from numerical simulations of our model show the moral hazard problem has a significant impact on the usefulness of risk sharing through fiscal policy.

2.4.1 Functional forms

We assume the effort takes values in $0 \leq e \leq 1$, the disutility from effort is $h(e) = (1-e)^{\eta}$ with $0 < \eta < 1$. If we interpret $l = 1 - e$ as "leisure", then the elasticity of the utility from leisure with respect to leisure is η. The private equity return is

$$\mu_p(e) = \underline{\mu} + (\bar{\mu} - \underline{\mu})e$$

with $\underline{\mu} < \bar{\mu}$. This specification means that the private equity return is $\bar{\mu}$ with maximal effort ($e = 1$) and $\underline{\mu}$ with minimal effort ($e = 0$). Then the first-order condition for the maximization of (2.11) (after substituting $\mathbf{z}$ and log-differentiating) is

$$-\frac{\eta}{1-e} + \frac{(\bar{\mu} - \underline{\mu})\theta_2}{\varepsilon(\beta+\delta) + (1-\varepsilon)\left(\mu(e)'\theta - \frac{\gamma}{2}\theta'\Sigma\theta\right)} = 0. \tag{2.27}$$

This equation is linear in e and therefore can be solved by hand. After some tedious algebra, the solution is

$$e = \frac{1}{1+\eta(1-\varepsilon)}\left(1-\eta\frac{\varepsilon(\beta+\delta)+(1-\varepsilon)\left(\mu(0)'\theta-\frac{\gamma}{2}\theta'\Sigma\theta\right)}{(\bar{\mu}-\underline{\mu})\theta_2}\right), \tag{2.28}$$

where $\mu(0) = (\mu_m+\delta, \underline{\mu}, \mu_g)'$ is the vector of expected returns with zero effort.

2.4.2 Calibration

We calibrate the parameters at annual frequency. The relative risk aversion coefficient $\gamma = 4$, which is relatively common in the macro literature. The elasticity of intertemporal substitution is $\varepsilon = 0.7$. We are aware that the vast majority of the empirical estimates of EIS is less than 1, and most macro papers assume so but most macro-finance papers choose a value larger than 1 (Schmidt and Toda, 2015). The results are not sensitive to this choice, and all the graphs below look very similar whether $\varepsilon = 0.1$ or $\varepsilon = 1.5$. The effective discount rate is $\beta+\delta = 0.04$, which is standard. We set the death rate to $\delta = 0.01$, which implies an average lifespan of $1/\delta = 100$ years. We choose such a low rate because we interpret agents as dynasties, not necessarily as households. The stock market return and volatility are $\mu_m = 0.04$ and $\sigma_m = 0.08$, which are arguably low. We choose these numbers because we interpret the first technology as a mature industry. We set the maximum and minimum return on technology and the aggregate and idiosyncratic volatility to be $\bar{\mu} = 0.1$ (10%), $\underline{\mu} = 0.02$ (2%), $\sigma_p = 0.15$ (15%), and $v = 0.1$ (10%), which seem reasonable. The correlation between the stock market and private equity is $\rho = 0.8$. As for the most important parameter, the elasticity of effort, we use two values of 0.05 and 0.075 for η because we want to see how much effect those small elasticities have on the optimal tax policy. Table 2.1 summarizes the parameter values. We then solve for the equilibrium by changing the consumption tax rate from 0

to 100%.

Table 2.1: Parameter values

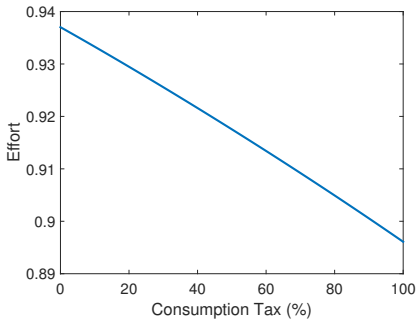
Description	Symbol	Value
Discount rate	β	0.03
Risk aversion	γ	4
Elasticity of substitution	ε	0.7
Death rate	δ	0.01
Elasticity of effort	η	0.05; 0.075
Stock market return	μ_m	0.04
Stock market volatility	σ_m	0.08
Maximum investment return	$\bar{\mu}$	0.1
Minimum investment return	$\underline{\mu}$	0.02
Aggregate volatility	$\bar{\sigma}_p$	0.15
Correlation	ρ	0.8
Idiosyncratic volatility	v	0.1

2.4.3 Results

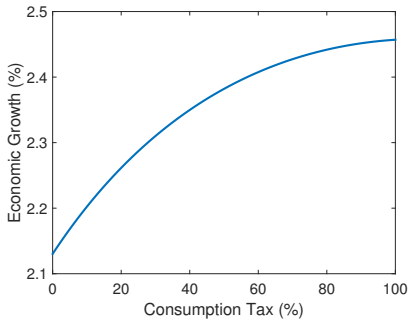
Figure 2.1 reports the effects of consumption tax on the variables of interest for the case of $\eta = 0.05$. Figure 2.1a shows the effort exerted by entrepreneurs. Since the risk sharing through taxation provides an insurance, it exacerbates the moral hazard problem as expected. But the reduction in effort is moderate because of our choice of small η; when consumption tax is newly introduced at the rate of 1%, the reduction in effort makes the expected return of private equity go down from 9.50 to 9.49 percentage point. If the elasticity of the expected return with respect to net-of-tax rate is calculated based on these figures, the elasticity is quite small at $\tau = 0$, equal to 0.03.[14]

Despite the lower expected return on private equity caused by moral hazard, Figure 2.1b shows that the economy grows faster (the graph shows the expected aggregate consumption growth rate, μ_c). This is caused by three forces. First, higher tax allows

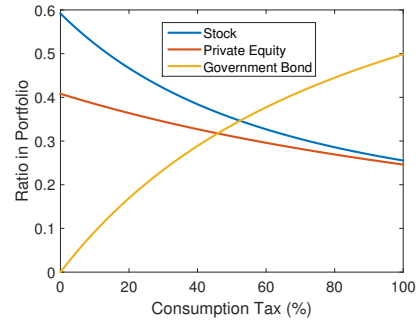
[14]We define the net-of-tax rate τ^N as $\tau^N = 1/(1+\tau)$. The elasticity of the expected return ζ is defined as $\zeta = (\Delta\mu_p(e)/\Delta\tau^N)(\tau^N/\mu_p(e))$. ζ increases as the tax rate goes up, but not much. At the optimal tax rate for the initial generation, 63% as shown below, ζ is still small and equal to 0.06.



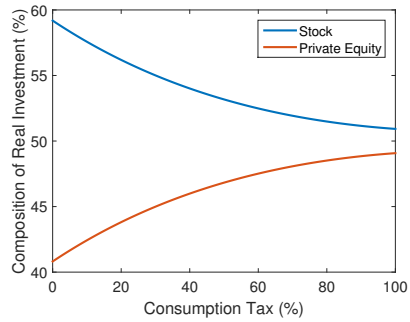
(a) Effort



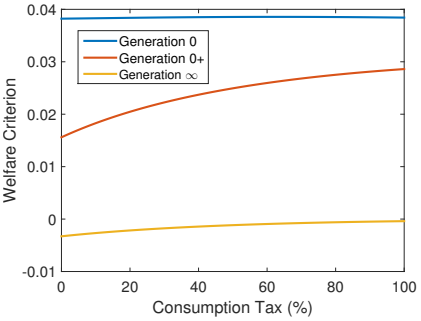
(b) Economic growth



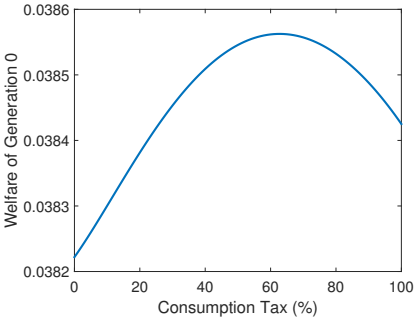
(c) Portfolio



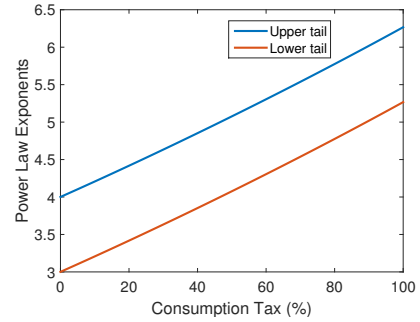
(d) Fraction of real investment



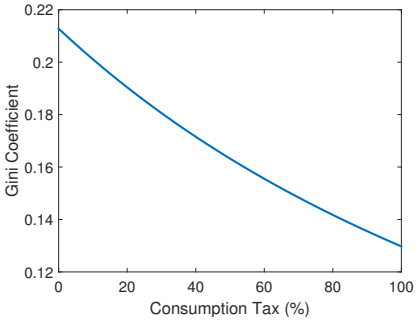
(e) Welfare criterion of each generation



(f) Welfare of initial generation



(g) Power law exponents



(h) Gini coefficient

Figure 2.1: Comparative statics with respect to the consumption tax rate ($\eta = 0.05$)

better risk sharing and induces the agents to allocate more capital to the high-risk, high-return private equity. Second, this reallocation of capital increases the total aggregate risk in the economy (because the aggregate volatility of private equity is higher than that of the stock market) and induces agents to save more. Third, due to better risk sharing, the idiosyncratic risk declines, which induces the agents to save less. As discussed in Schmidt and Toda (2015), precautionary savings occur if and only if $\varepsilon < 1$; the sign is reversed if $\varepsilon > 1$. In either case, the capital reallocation effect dominates. In other words, taxation encourages risk-taking, being consistent with the conclusion of traditional literature, in this case of very small degree of moral hazard problem, and the economy grows faster.

Figure 2.1c shows the portfolio share of the stock market (θ_1), private equity (θ_2), and the government bond (θ_3). Because increasing the tax rate raises the market capitalization of the bond, θ_3 monotonically increases. Both θ_1 and θ_2 declines, but private equity decreases relatively slower, which means that agents allocate more capital to private equity. In fact, within real investment, the fraction of private equity $\frac{\theta_2}{\theta_1+\theta_2}$ increases (Figure 2.1d). (Of course, increasing the tax further causes too much moral hazard and the private equity will collapse to zero eventually, but that does not happen until a consumption tax of around 270%.)

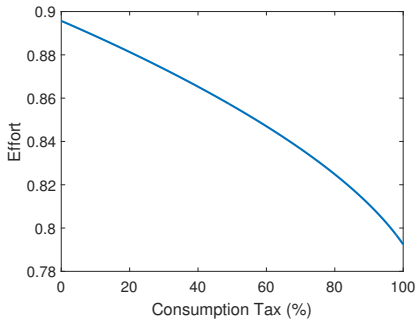
Figure 2.1e shows the welfare criteria of each generation: for the initial generation ($t = 0$), $\frac{a}{1-\theta_3}$; for generations born right after ($t = 0+$), $\frac{a(1-\theta_1)}{1-\theta_3}$; for generations in the far distant future ($t = \infty$), $\mu_c - \frac{\gamma\sigma_c^2}{2}$; see (2.24). Since the welfare of the initial generation is fairly flat, Figure 2.1f magnifies the graph. According to these graphs, all generations gain up to tax rate around 63% (39% tax burden out of income). For generations 0 and 0+, the unit is consumption equivalent. At $\tau = 63\%$, the welfare gain of the initial generation is 0.89%, while that of the generation born right after is 68%. Future generations gain even more since the risk-adjusted growth is higher. Therefore risk sharing through taxation

can achieve a Pareto improvement in this case, and most benefit is enjoyed by the future generations. The reason why the generation born right after $t = 0$ gains so much is because they inherit a larger fraction of wealth since private equity investment increases, which by assumption cannot be pledged for annuity. If some part of private equity can be used for annuity, the conclusion might be different.

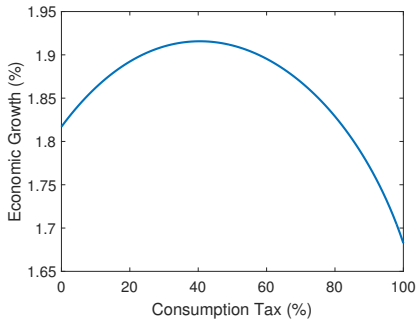
Figure 2.1g shows the power law exponent of the cross-sectional consumption distribution. With no tax, the exponents are 4 for the upper tail and 3 for the lower tail, which are roughly the values for the U.S. consumption distribution (Toda and Walsh, 2015). As the tax rate increases, so do the power law exponents because better risk sharing reduces the idiosyncratic volatility, which determines inequality. In fact, the Gini coefficient in (2.1h) monotonically decreases.

We show the result of the same exercise for the other parameter choice, $\eta = 0.075$, in Figure 2.2. Since the elasticity of effort is higher in this case, entreprener's effort level goes down more rapidly than the previous case as the tax rate increases as shown in Figure 2.2a. Because of this change in effort, the expected return of private equity decreases from 9.17% to 9.16% when consumption tax is introduced by 1%, which implies the elasticity of the expected return with respect to net-of-tax rate equals 0.06 at $\tau = 0$. Even though the elasticity is still quite low, just slightly higher than the previous case, the remaining variables respond to consumption tax very differently.

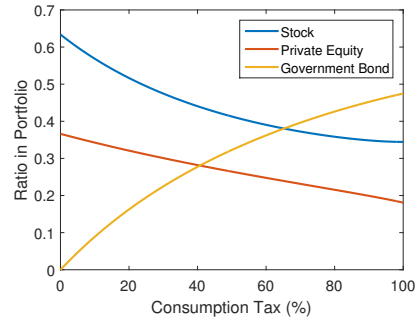
While economic growth is spurred by taxation for low tax rates, the growth rate turns to decrease as the tax rate goes up beyond 40% (Figure 2.2b). Figures 2.2c and 2.2d suggest the main reason for this movement of economic growth rate. As in the previous case, the portfolio share of government bond increases consistently since increasing the tax rate raises the market capitalization of the bond. As for the two investment technologies, agents raises the share of private equity as the tax rate increases when the tax rate is low, but then they start to reduce its share once the tax rate goes beyond certain



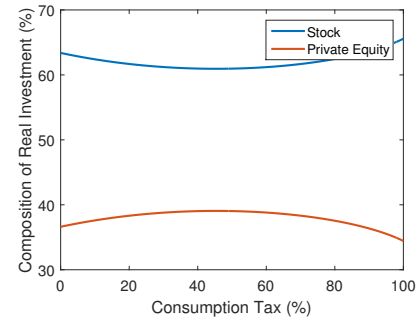
(a) Effort



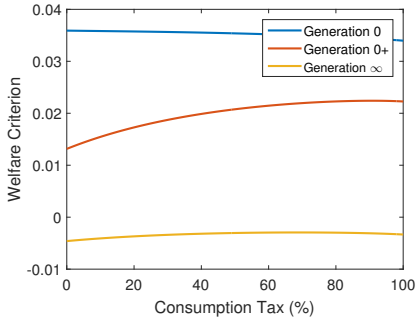
(b) Economic growth



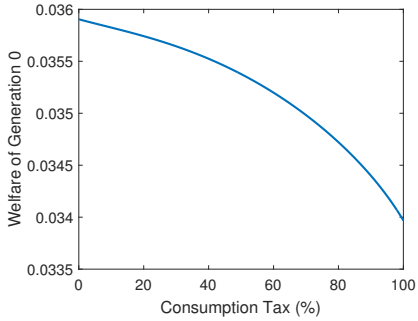
(c) Portfolio



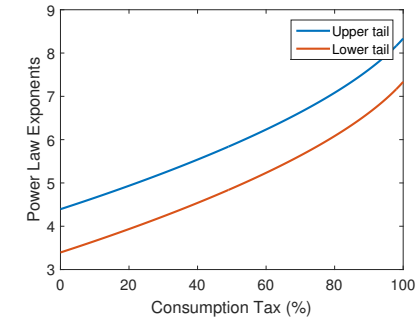
(d) Fraction of real investment



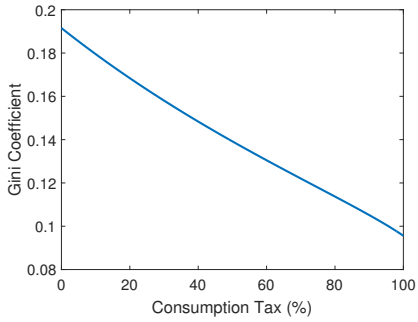
(e) Welfare criterion of each generation



(f) Welfare of initial generation



(g) Power law exponents



(h) Gini coefficient

Figure 2.2: Comparative statics with respect to the consumption tax rate ($\eta = 0.075$)

level. This result is caused by the following mechanism: taxation alters the attractiveness of private equity as a financial asset in two ways. First, a higher tax rate makes private equitly more attractive by reducing idiosyncratic risk through risk sharing. Second, since the moral hazard problem gets more severe as the tax rate increases, it makes private equily less attractive through the decrease in expected return. Figure 2.2d shows the former effect outweighs the latter at first until the tax rate of 47%, then the relation is reversed. This effect of taxation on portfolio allocation is reflected in the hump-shaped graph of economic growth rate.

Figure 2.2e shows it is possible to improve the welfare of future generations by a low tax rate, but Figure 2.2f indicates the welfare of initial generation cannot be improved by the fiscal policy tool we are considering. It may look puzzling that even the lowest tax rate harms the welfare of initial generation at first glance; when the tax rate is low, risk-taking is encouraged, economic growth is accelerated, and the effort level is reduced, which contributes to a higher utility. In fact, consumption tax has its own negative effect on utility through the wedge on the choice between consumption and effort (or leisure consumption).[15] Since the positive gain from risk sharing through taxation is small for the initial generation, the loss from the intratemporal wedge exceeds the gain and reduces the welfare in this case. This result sheds light on the important issue that the previous studies have ignored. Even if taxation successfully encourages risk sharing, it does not necessarily imply improvement of welfare when the performance of risky investment depends on entrepreneur's effort.

Figure 2.2g and 2.2h show a higher tax rate invariably contributes to mitigating inequality in the economy better even when the responses of other variables to a increase in the tax rate are very different between the cases.

[15]This type of effect is not specific to our choice of consumption tax. For example, capital income tax also brings the intratemporal wedge.

2.5 Concluding remarks

Economists have argued that taxation may well encourage risk-taking. Although the studies with that conclusion are still relevant, the implication of moral hazard problem has been largely ignored in the literature. This theoretical study aims to answer how the moral hazard problem affects the optimal consumption tax rate, employing a tractable model that can describe welfare, risk-taking, economic growth, and inequality. The results of our numerical simulations show that the effect of moral hazard problem is significant. If the return of risky asset depends on entrepreneur's effort, risk sharing through taxation cannot improve the welfare even under an arguably low elasticity of effort.

To derive relevant implications for policy debates, the key parameter is the degree of moral hazard, for which we look at the elasticity of the expected return with respect to net-of-tax rate in this study. Although it seems difficult to obtain reliable estimation for this parameter by an empirical study, our results suggest understanding the moral hazard problem is critical for the study of taxation and risk-taking.

2.6 Appendix

2.6.1 Solution algorithm

This section explains how to numerically compute the equilibrium. As explained in the preceding sections, the equilibrium is a fixed point of a nonlinear mapping $F : \mathbb{R}^4 \to \mathbb{R}^4$. Let $z = (\mu_g, e, \theta_1, \theta_2)$ be the initial guess of the equilibrium objects. We use the Matlab `fsolve` command, except that we use a few tricks to improve the numerical stability.

Updating the government bond return We use (2.17a) to update the government bond return μ_g. One caveat is that since $F_1(z)$ contains $\theta_3 = 1-\theta_1-\theta_2$ in the numerator, which is zero when the tax rate is zero, it introduces a discontinuity. To deal with this, we solve the autarky case ($\tau = 0$) separately, and put a reasonable bound on μ_g (say $[-1,1]$) so that the updated μ_g will stay in the bound even if $F_1(z)$ is infinite.

Updating the effort We use (2.28) to update the effort. If $e < 0$, we replace by 0. If $e > 1$, we replace by 1.

Updating the portfolio To update the portfolio, we first solve the optimal portfolio problem with no shortselling constraints. For generality let J be the number of assets, $\mu \in \mathbb{R}^J$ be the vector of expected returns, and Σ be the $J \times J$ variance-covariance matrix. Then the optimal portfolio problem is

$$\text{maximize} \quad \mu'\theta - \frac{\gamma}{2}\theta'\Sigma\theta \quad \text{subject to} \quad \theta \geq 0, 1'\theta = 1. \tag{2.29}$$

Clearly, our portfolio problem (2.10) is a special case of (2.29) with $J = 3$. The gradient of the objective function of (2.29) is $\mu - \gamma\Sigma\theta$. Hence it is straightforward to solve for the optimal portfolio problem numerically using the `fmincon` command in Matlab.

However, numerical optimization introduces instability in the subsequent equilibrium computation. To deal with this issue, we note that most of the time the nonnegativity constraint $\theta \geq 0$ does not bind. Then the problem becomes equality constrained, which we can solve analytically. Let $\phi = (\theta_1, \ldots, \theta_{J-1})'$ be the portfolio share of all assets except the last. Using the accounting constraint $1'\theta = 1$, it follows that $\theta = a + A\phi$, where

$$\underbrace{a}_{J\times 1} = \begin{bmatrix} 0_{J-1} \\ 1 \end{bmatrix} \quad \text{and} \quad \underbrace{A}_{J\times(J-1)} = \begin{bmatrix} I_{J-1} \\ -1'_{J-1} \end{bmatrix}.$$

Here $0_K, 1_K, I_K$ denote the K-vector of zeros, K-vector of ones, and $K \times K$ identity matrix. Substituting $\theta = a + A\phi$, the optimal portfolio problem (2.29) (without the nonnegativity constraint) becomes

$$\text{maximize} \quad \mu'(a + A\phi) - \frac{\gamma}{2}(a + A\phi)'\Sigma(a + A\phi).$$

By the first-order condition, the solution is

$$A'(\mu - \gamma\Sigma a) - \gamma A'\Sigma A\phi = 0 \iff \phi = \frac{1}{\gamma}(A'\Sigma A)^{-1}A'(\mu - \gamma\Sigma a).$$

Using this expression, we can compute the optimal portfolio $\theta = a + A\phi$. In practice, we first compute this θ, and if it satisfies $\theta \geq 0$, then we know that it is the (unique) solution. If some element of θ is negative, then we switch to the numerical solution using `fmincon`.

One caveat is that when the tax rate is very high (so moral hazard is severe) and the minimum return $\underline{\mu}$ is low, zero investment in private equity ($\theta_2 = 0$) may be optimal. Then the government bond and stock market will be identical from individual's point of view, so the optimal portfolio is indeterminate. In this case we need to pin down θ_1 and $\theta_3 = 1 - \theta_1$ from (2.18) using the no-arbitrage condition $\mu_g = \mu_m + \delta$.

Chapter 2, in full, is currently being prepared for submission for publication of the material. Miyoshi, Yoshiyuki; Toda, Alexis Akira. The dissertation author was the primary investigator and author of this material.

Chapter 3

Rental Housing Investment and Tax Reforms: An Empirical Study of Tax Clientele Model

Abstract

This paper examines the relationship between household marginal tax rates and the probability of owning rental housing. A simple model of tax clienteles predicts a positive association between the marginal tax rate and the ownership of rental housing in general, but the passive loss provisions introduced by the tax reform in 1986 is expected to limit this association to households with income less than the threshold value. The empirical results based on 1983, 1989, 1992, 1995, and 1998 Surveys of Consumer Finances show the marginal tax rate has a positive association with rental housing ownership only for the group of households that can deduct passive losses on rental housing investment, implying the tax clientel model is relevant for rental housing investment.

3.1 Introduction

The effects of taxation on housing investment has been one of the major topics in public finance.[1] In the literature, owner-occupied housing has drawn considerably more interest than rental housing.[2] Although the homeownership rate in the US is relatively high compared to the other developed countries, one third of the households in the nation live in rental housing. Thus expanding our understanding of how taxation affects rental housing investment is important.

The interaction between progressive income tax and holding of rental properties is an aspect of this subject that has not been studied extensively. It has been known since Modigliani and Miller (1958) that investors are divided into "clienteles" of different types of financial assets if the marginal tax rate varies across investors and the financial assets are subject to different rules of taxation. For example, if stocks are taxed more lightly than bonds, possibly due to a low capital gains tax rate, stocks attract investors with a high marginal tax rate. The goal of this study is to provide empirical evidence about whether the tax clientele model can be applied to the portfolio choice including rental housing investment. Examining the relevance of tax clientele model for rental housing is meaningful because the tax rate applied to (marginal) investors of rental housing affects the cost of rental housing provision and eventually the market rent.[3]

The literature of emprical studies of taxation and portfolio choice is not large because of data availability, complicated incentives caused by the tax code, and potential endogeneity of tax rates. I take advantage of the provisions regarding passive losses that was introduced as part of the Tax Reform Act (TRA) in 1986 to identify the effect of

[1] Rosen (1985) surveys the early economic studies on this topic.

[2] For example, the studies of taxation and owner-occupied housig include Glaeser and Shapiro (2002), Hilber and Turner (2014), Poterba (1984) and Poterba and Sinai (2008), Rosen and Rosen (1980).

[3] As a related figure, Poterba (1990) estimates TRA will increase the steady state rent by 10-15 percent assuming the marginal investors remain in the top bracket. If the tax rate for the marginal investors declines due to the passive loss provisions, the estimate for the rent increase should be larger.

taxation on holding of rental housing. Before TRA, households could deduct passive losses from their ordinary income. Since at that time it was common for rental housing investment to generate tax losses especially for the first several years,[4] the deduction of passive losses was valuable for investors with high marginal tax rates. However, TRA restricted households from deducting passive losses as a general rule in an attempt to deal with tax sheltering activities. The interesting feature that accompanied this general rule is that households whose adjusted gross income is below a certain level are allowed to deduct tax losses for rental housing investment from the ordinary income. Therefore the taxation may affect the two groups of housholds—households with income above or below the threshold income—differently.

I first present a simple model of tax clienteles, in which a household with a high marginal tax rate is more likely to own residential rental properties in its portfolio if there is no restriction on passive losses. But the model implies that in the presence of the restriction on passive losses, the marginal tax rate does not correlate with the ownership of residental rental properties for households whose income is above the threshold while the positive correlation still holds for the other group of households.

To test the prediction of the model, I use data from five years of the Surveys of Consumer Finances (SCF), which span before and after TRA. I estimate the effect of marginal tax rate on the probability of ownership of residential rental properties. The econometric issue is the potential endogeneity of marginal income tax rate. The theory predicts the effect of marginal tax rate on portfolio choice, but in turn portfolio choice affects income and then the income tax rate. I use the first-dollar marginal tax rate, which is a standard remedy in the empirical public finance literature, to address this issue.

The results of probit estimation are modestly consistent with the predictions of the model although some of the results make the interpretation ambiguous. For three

[4] See Follain et al. (1987) for a formal analysis.

years out of four years after TRA, the estimates for the marginal tax rate are (marginally) statistically significant for the group of households that are allowed deduction of passive losses. On the contrary, the estimate for the other groups of households are never statistically significant for the same years. Thus I conclude the tax clientele model is relevant for rental housing investment as well as for the portfolio choice among financial assets although the standard errors are high, possibly due to the idiosyncratic nature of rental housing investment.

This study aims to contribute to two strands of literature. First, in the literature on rental housing investment the interaction between the progressive tax schedule and portfolio choice has been mostly ignored although there are a couple of exceptions. Berkovec and Fullerton (1992) construct and estimate a general equilibrium model in which the supply of rental housing is endogenously determined and households choose their portfolio considering a progressive income tax schedule. While the model of Berkovec and Fullerton (1992) is essentially static, Chambers et al. (2009) use a dynamic model and solve it for steady state equilibrium. Both studies provide interesting implication for policy proposals by simulations, but they presume household should follow the solution to the standard maximization problem in the presence of progressive taxation without empirical evidence.[5] This study provides the first empirical evidence that households actually follow the prediction of the tax clientele model.

Second, this study also contributes to the empirical literature on taxation and portfolio choice. Among the previous empirical studies on this subject,[6] the current study is closest to Poterba and Samwick (2002) in that they use cross-section data from the several years of SCF and try to remedy the endogeneity problem regarding the marginal tax rate by the first dollar variable. The contribution of this study is the test of

[5]In addition, neither Berkovec and Fullerton (1992) nor Chambers et al. (2009) take into accout the restriction on passive losses.

[6]Poterba (2002) provides an extensive survey of the literature.

the tax clientele model for rental housing investment with an additional dimension of identification that utilizes the passive loss provisions for rental housing.

The strucure of the rest of the paper is as follows: section 2 present a simple model for taxation and portfolio choice including rental housing investment, in which the passive loss provisions can be incorporated. Section 3 summarizes the changes in the tax code in 1980s that are relevant to rental housing investment. Section 4 presents the dataset and the empirical specification. Section 5 reports the estimation results. Section 6 is the brief conclusion.

3.2 A tax clientele model for rental housing investment

Although the previous studies have proposed various tax clientele models that provide theoretical tools to analyze how taxation affects households' portfolio choice among financial assets, such as bonds, stocks, and tax-exempt and tax-deferred assets, few studies include real assets in their scope.[7] I provide a simple clientele model to obtain some insights as to the relation between taxation and incentive for rental housing investment. In the model, agents aim to maximize their return, and they are heterogeneous in terms of income tax rates. There are three types of assets: taxable bond, tax-exempt bond, and residential rental property. All the assets are riskless in this economy; taxable and tax-exempt bond pay deteministic amounts to the owners, and housing rent is determined outside the model without uncertainty. The asset prices and returns are determined in equilibrium to clear the supply of assets. Therefore if there were no rental housing, it would be the model of Miller (1977): in equilibrium, Agents with high tax rates choose to own tax-exempt bond while agents with low tax rates own

[7]The classical examples for clientele models include Auerbach and King (1983) and Miller (1977).

taxable bond. Formally,

$$\begin{aligned} &E_i = W_i,\ B_i = 0 \quad \text{if } (1-t_i)r^b < r^e, \\ &E_i = 0,\ B_i = W_i \quad \text{otherwise,} \end{aligned} \tag{3.1}$$

in which W_i is the total wealth of household i; E_i and B_i are the values of tax-exempt bond and taxable bond owned by household i, respectively; r^b and r^e are the returns to tax-exempt bond and taxable bond determined in the market, respectively; t_i is the income tax rate for household i.

In the model of this study, households can also invest in the residential rental property, and they are allowed to borrow a loan to finance the investment up to the value of property at the interest rate equal to r^b. The model assumes an exgenous debt limit on total balance for this type of loan for each household, D_i. While rent income is taxed at the same rate as income from taxable bond, households deduct depreciation of the property and interest payment for the loan. The rate of tax depreciation d may not be equal to the rate of economic depreciation δ, but both rates are assumed to be constant over the lifetime of property. The asset price P for a unit of housing is also assumed to be constant across time.[8] Household i's problem to maximize the net return is represented as

$$\begin{aligned} \max_{E,B,H,\lambda} \quad & r^e E_i + (1-t_i)r^b B_i + (1-t_i)(RH_i - r^b \lambda_i P H_i) - (\delta - t_i d)PH_i \\ \text{s.t.} \quad & W_i = E_i + B_i + (1-\lambda_i)PH_i \end{aligned} \tag{3.2}$$

in which H_i is the amount of rental housing; λ_i is the loan-to-value ratio of rental housing investment.

[8]This assumption is pluasible if the economy is in the steady state.

Let t^* be the real number such that $(1-t^*)r^b = r^e$. Then by the same argument regarding equation 3.1, households whose t_i is greater than t^* make no investment in taxable bond in this model as well. Moreover, a definite implication for loan-to-value ration can be derived for this group of households: they choose the 100% loan-to-value ratio because after-tax cost of property loan interest, $(1-t_i)r^b$, is less than the return from investment in tax-exempt bond, r^e, for this group. On the contrary, households with tax rates less than t^* invest no wealth in tax-exempt bond. If a household in this group wants to invest in rental housing, it is indifferent about how much to borrow to finance the investment.

First, consider the rates of tax and economic depreciation are equal: $d = \delta$. Then equation 3.2 can be rewritten as

$$r^e E_i + (1-t_i)r^b B_i + (1-t_i)(RH_i - dPH_i - r_b PH_i).$$

In this case, the equilibrium gross return for rental housing is equal to the sum of depreciation rate and market interest rate: $R/P = d + r_b$. Households with any income tax rates are indifferent about how much to invest in rental housing since there is no excess return due to heterogeneity in tax rates.

Second, let the tax depreciation rate be greater than the economic depreciation.[9] Then the return for rental housing investment, which is represented as follows, is an

[9]This is a common assumption to analyze the tax incentive for housing investment, especially investment behavior in early 80s, when tax sheltering activities that take advantage of residential rental properties were rampant (Ozanne 2012, Poterba 1993). In practice, however, although the tax depreciation rate is larger only in early years of lifetime of rental properties, the rate is smaller (usually zero) than economic depreciation rate in later years. To incorporate this feature of tax code, the model should have a housing asset price as a function of its vintage, and it makes the model significantly more complicated. Follain et al. (1987) disucuss the relation between the "front-loading" of depreciation allowance and the passive loss provision.

increasing function of t_i:

$$(1-t_i)\left(\frac{R}{P}-r^b\right)-(\delta-t_i d).$$

In equilibrium, households invest in rental housing up to its debt limit D_i if its tax rate is greater than the cutoff tax rate t^{**}. The cutoff tax rate and housing asset price are determined by the system of equations:

$$\begin{aligned}(1-t^{**})\left(\frac{R}{P}-r^b\right)-(\delta-t^{**}d)&=0,\\ \int_{t_i>t^{**}}\frac{D_i}{P}\mathrm{d}i&=\bar{H},\end{aligned} \tag{3.3}$$

in which $\bar{H}$ is the aggregate demand for housing, which is determined outside the model.

The taxable income from rental housing investment is non-positive for any households in equilibrium in this model. If a household owns H_i unit of rental housing and $t_i \geq t^{**}$, it can be confirmed that the taxable income from rental housing, $RH_i - dPH_i - r^b PH_i$, is less than or equal to zero and that the equality holds only for the marginal household whose tax rate equals t^{**}. Therefore the tax rules regarding losses on rental housing investment are critical for the incentive for housing investment. Consider the tax provision that does not allow households whose tax rate is greater than some rate $\bar{t}$ to offset their income from financial investment against losses on housing investment, which is similar to the anti-shelter provision introduced by 1986 Tax Reform Act. Then for any households that cannot deduct tax losses from rental housing from the other income, the after-tax return from housing investment is negative:

$$\frac{R}{P}-r^b-\delta=\frac{t^{**}(\delta-d)}{1-t^{**}}<0.$$

Thus households that are not qualified for deduction of losses on rental housing do not

own any rental property. In addition, note that this expression for the return does not depend on the individual tax rate for households, t_i. The equilibrium under this rule of tax losses is described as follows: there exists the cutoff tax rate $\tilde{t}^{**}$, and households own rental property to their debt limit if their tax rate satisfies $\tilde{t}^{**} \leq t_i \leq \bar{t}$; equation 3.3, the second equation for equilibrium conditions, is modified as

$$\int_{\tilde{t}^{**} \leq t_i \leq \bar{t}} \frac{D_i}{P} \mathrm{d}i = \bar{H}.$$

This model is too simple and includes some bold assumptions, including no consideration of uncertainty. In particular, if a strict borrowing constraint for low and middle income households is present in reality, it may well severely distort the prediction of the model, concentration of ownership of rental housing below the threshold. However, it offers some intuitions about the effect of income taxation on households' portfolio choice among financial assets and rental housing. In general, households with higher marginal tax rates are more likely to own rental properties if the rules of tax depreciation is more favorable than economic depreciation. If the tax code does not allow some households with high marginal tax rates to deduct losses on rental housing, the marginal tax rate should affect the group of households that are qualified for deduction significantly even after controlling the income and wealth effects. The empirical analysis to follow examines the latter intuition using the data set of household portfolio before and after the tax reform in 1986.

3.3 Tax reforms in 1980s

The US federal income tax system experienced two major changes during 1980s: the Economic Recovery Tax Act (ERTA) in 1981 and TRA in 1986. These reforms had very different effects on the incentives for rental housing investment. The provisions

that seemed to affect rental housing investment include ones regarding depreciation, capital gains, and passive losses. ERTA increased the tax advantage of rental housing by shortening the tax lifetime for those property from 32 to 15 years. On the contrary, TRA couteracted this effect by extending the liftime to 27.5 years. A similar pattern is found for capital gains tax rates. While ERTA lowered the marginal tax rate on long-term capital gains for the top income bracket from 28 percent to 20 percent, TRA put the top tax rate back to 28 percent. Since there were no special provisions for captail gains from rental property, unlike for owner-occupied properties, the changes in capital gains tax rate should have had real impacts on the incentive of investors.

In addition to these changes, TRA had some provisions to prevent tax payers from using rental housing as a tax shelter vehicle. It has widely been recognized that rental housing was one of the most active shelter vehicles in the early 1980s partly because of the generous depreciation rule introduced by ERTA (Follain et al. 1987, Ozanne 2012, Poterba 1993). Investors could usually report taxable losses even though they did not have any economic losses. TRA aimed to deter this practice by the provision regarding passive losses. The provision made it no longer possible to offset ordinary taxable income with passive losses, as which taxable losses on rental housing are usually counted, but it allowed only passive income to be offset. However, TRA also had the special exemption for this anti-shelter provision: if landlord's modified adjusted gross income is below $100,000, the landlord may deduct up to $25,000 in passive losses against other income.[10] The maximum amount of deduction starts to phase out beyond $100,000 of modified adjusted gross income and becomes zero at $150,000. Therefore landlords whose income satisfies the requirement of exemption could still have the benefit of passive losses after TRA. Although there have been a considerable volume of studies for

[10]The modified adjusted gross income is defined as adjusted gross income figured without taking into accout the taxable amount of social security, the deductible contributions to individual retirement accounts and some pension plans, any passive activity income and loss, and other income components that are usually in small amounts.

the effects of changes in depreciation rule and capital gains tax rate, no formal empirical study has been conducted for the effect of the anti-shelter provision of TRA.[11] The econometric analysis in the next section is focused on that effect.

3.4 Data and methodology

3.4.1 Survey of Consumer Finances

I use data from SCF conducted in 1983, 1989, 1992, 1995, and 1998.[12] SCF is a triennial survey that started in 1983. Those surveys cover the time period that is suitable to examine the short-term and long-term effects of ERTA and TRA. SCF has great advantages for the purpose of this study; it reports the detailed composition of financial and non-financial assets, including the information about rental real estate properties as well as owner-occupied housing. One of the important features of SCF is its oversampling of high-income households. Since the distribution of wealth is skewed towards high-income households, the oversampling helps us obtain accurate information about households' portfolio choice. SCF also reports various income components of households, which makes it possible to estimate their federal income tax liabilities and marginal tax rates.[13] Since this study examines how tax rates and rules affect affect households' incentive for rental housing investment, these types of information is essential. One weekness of the data set of this study is that it does not include data from 1986 SCF. 1986 SCF could be useful to look at households' behavior regarding rental housing just before the introduction of TRA, the period for which experts have

[11] Poterba (1993) briefly argues that the anti-shelter provition of TRA worked citing the statistics of real estate partnership sales.

[12] Poterba and Samwick (2002) use SCF of same years and finds broadly consistent evidence with the standard clientele model regarding the correlation between the marginal tax rate and the ownership probability for financial assets.

[13] Since the public data sets of SCF do not report the geographic information of households, their state tax liabilities and marinal tax rates cannot be estimated.

Table 3.1: Probability of ownership (percent) of residential rental properties

Income range ($000, 1995 dollars)	1998	1995	1992	1989	1983
0-15	1.91	2.64	2.18	2.04	1.63
15-25	4.86	5.11	5.03	3.41	3.82
25-50	8.25	6.75	7.52	9.61	6.05
50-75	13.54	10.07	12.40	10.79	11.56
75-100	16.80	13.69	16.64	15.13	17.55
100-250	24.62	23.24	27.66	29.45	19.94
250 +	39.98	27.27	43.98	34.50	35.32
Total	8.92	7.39	8.27	8.67	6.85

Notes: Households are weighted by sample weights in each year.

said that tax sheltering activities with rental housing investment was rampant. Unlike the surveys in the other years, however, 1986 SCF does not report the information about income components, and it is impossible to estimate the marginal tax rates of households accurately enough.

3.4.2 Summary information on owners of rental housing

Based on the information of SCF, I define owners of rental housing as households that own any type of residential properties, including single family houses and apartments. Households are defined as owners no matter if they own the properties independently, jontly with others, or through partnerships.

Table 3.1 presents summary information on the ownership probabilities of rental housing by income group. There is a strong positive correlation between income and rental housing ownership. Moreover, this correlation is fairly constant over time, even after TRA. Since the marginal income tax rate is correlated with income, it is not clear whether this pattern is caused by the income effect or the effect of marginal tax rate on portfolio choice.

3.4.3 Estimating marginal tax rates

I obtain the estimates of marginal tax rates for the SCF households from the National Bureau of Economic Research's TAXSIM program. Although it is possible to estimate the marginal tax rate by making full use of the information about income components reported in SCF, there is an econometric issue when such estimates for marginal tax rates are used to study the tax effects on portfolio choice as mentioned above. The marginal tax rate is an endogenous variable for households because it is affected by households' portfolio choice; for example, if a household that receives high financial income takes advantage of tax sheltering activities heavily, its resulting marginal tax rate may be as low as middle income households. Exsiting empirical studies avoid this issue by constructing the "first-dollar" marginal tax rate, which is the marginal tax rate if income from financial assets is assumed to equal zero.[14]

I basically follow the procedure of Poterba and Samwick (2002) to construct the first-dollar marginal tax rates. Let $\mathbf{Y}_{it}^{B}$ be the vector of income components and deductions for household i in time t when the household's income from dividends, interest, and capital gains is assumed to be zero. Then $T(\mathbf{Y}_{it}^{B})$ is the federal income tax liability calculated by TAXSIM based on $\mathbf{Y}_{it}^{B}$. The first dollar marginal tax rate is defined as $\tau_{it} = [T(\mathbf{Y}_{it}^{B} + \Delta) - T(\mathbf{Y}_{it}^{B})]/\Delta$, in which Δ is an increment in ordinary income. The increment is chosen as the maximum of 5% of the total value of household's financial assets and \$100. Poterba and Samwick (2002) justify this choice for increment because 5% can be considered as the normal return for the financial assets.

[14] In a recent study, Kawano (2014) uses this type of variable to test the dividend clientele hypothesis with data from SCF.

3.4.4 Empirical specification

This study is focused on how the rules regarding passive losses that was introduced by TRA affected households' incentive for rental housing investment. I estimate probit models for rental housing ownership as a function of the marginal tax rate and the modified aggregate gross income (MAGI) year by year. I also control for a variety of socioeconomic and demographic variables.[15]

The estimating equation is given by

$$y_{it}^{*} = \gamma_L \mathbb{I}(\tilde{Y}_{it} \leq C)\tau_{it} + \gamma_H \mathbb{I}(\tilde{Y}_{it} > C)\tau_{it} + \eta \mathbb{I}(\tilde{Y}_{it} \leq C) + \mathbf{x}_{it}\beta + \varepsilon_{it},$$

$$\mathrm{P}(y_{it} = 1 | \tilde{Y}_{it}, \mathbf{x}_{it}) = \Phi(y_{it}^{*}),$$

in which y_{it} is the binary variable for ownership of residential rental properties; y_{it}^{*} is the latent variable; $\mathbb{I}(\cdot)$ is the indicator function, $\tilde{Y}_{it}$ is MAGI and C is the cutoff value of MAGI for the passive losses rule exemption, which is set as \$100,000 for all the survey years; $\mathbf{x}_{it}$ is the vector of control variables including the constant term; γ_L, γ_H, η, and β are the parameters to be estimated and ε_{it} is the error term.

The parameters of interest are γ_L and γ_H. η, the coefficient on the binary variable that indicates MAGI is no more than the cutoff, is included for the technical reason: it allows the effect of marginal tax rate to have different intercepts across the two MAGI groups. If the tax code does not treat losses from retal housing investment differently across income groups, households with a higher marginal tax rate has a stronger incentive to participate in rental housing investment for the purpose of offsetting their positive income. After the negative losses rule was put in effect in 1986, however, the argument in section 3.2 predicts that γ_L is positively larger in magnitude than γ_H because the

[15] I have also estimated tobit models for the ratio of loan balance collateralized by residential rental properties to the total financial value for each household. The results are quite similar to the ones from probit models, arguably reflecting very small probability of ownership for residential rental properties.

high-income households can no longer offset their income with passive losses from rental housing.

Since the portfolio choice is usually expected to be affected by income and wealth effects, I include a set of binary variables that indicate the income and net worth categories of households. Those categories are reported in thousands of 1995 dollars. The control variables include demographic variables: educational attainment, age, gender of household head, marital status, and the size of household. SCF asks households about their subjective attitude towards taking risk. The variable "risk averse" takes unity if the household is not willing to take any financial risks, and it takes zero otherwise.

3.5 Empirical findings

Table 3.2 reports the probit coefficients and standard errors for the ownership of residential rental properties by years. For the years after 1986 except for 1992, the coefficients for the marginal tax rate for households from the lower MAGI group are marginally statistically significant: the figures fail to satisfy the 5% significance level, but the p-values range between 0.053 and 0.060. On the contrary, the marginal tax rate coefficient for the higher MAGI group are cosistently statistically insignificant, and it is even negative for year 1998. The results provide modest support for the prediction made for the effect of passive losses rule introduced by 1986 TRA for these three years. However, the result for 1992 is an anomaly. The marginal tax rate coefficients for both MAGI groups are negative although they are not statistically significant. I cannot provide a plausible explanation for this anomaly. But there may be systematic inaccuracy in the first-dollar marginal tax rate calculation for this year because the results from the probit estimation for ownership of another asset category, tax-deferred accounts, also shows an anomaly only for this year, as I will present in the following.

Table 3.2: Probit estimates for the ownership of residential rental properties

Year	1998		1995		1992		1989		1983	
	Coeff.	S.E.	Coeff.	S.E.	Coeff.	S.E.	Coeff.	S.E.	Coeff.	S.E.
Constant	-1.313	0.363	-2.236	0.429	-2.184	0.451	-2.122	0.594	-2.408	0.814
MAGI<=100	-0.388	0.314	-0.312	0.328	-0.152	0.303	-0.135	0.438	-0.352	0.751
MTR										
MAGI<=100	0.939	0.485	0.920	0.490	-0.313	0.503	0.959	0.511	0.813	0.582
MAGI>100	-0.762	0.798	0.554	0.834	-0.361	0.826	0.257	1.336	0.164	1.386
Income										
15-25	-0.055	0.157	0.074	0.139	0.174	0.150	0.044	0.169	0.151	0.158
25-50	0.188	0.144	0.064	0.137	0.242	0.149	0.266	0.155	0.156	0.182
50-75	0.101	0.171	-0.047	0.165	0.319	0.178	0.264	0.187	0.276	0.238
75-100	0.137	0.187	-0.041	0.176	0.246	0.192	0.162	0.212	0.272	0.287
100-250	0.265	0.190	-0.212	0.185	0.226	0.200	0.527	0.207	0.227	0.320
250 +	0.418	0.202	-0.083	0.194	0.492	0.209	0.633	0.228	0.504	0.358
Net worth										
50-100	0.501	0.138	0.271	0.144	0.384	0.136	0.342	0.146	0.224	0.122
100-250	0.850	0.119	0.347	0.124	0.781	0.114	0.611	0.125	0.730	0.108
250-1000	1.167	0.120	0.938	0.117	1.284	0.119	1.001	0.128	0.996	0.126
1000 +	1.554	0.137	1.340	0.128	1.603	0.135	1.304	0.154	1.002	0.176
Education										
High school	-0.112	0.114	0.046	0.113	-0.104	0.118	-0.061	0.114	-0.147	0.105
Some college	0.023	0.113	0.084	0.115	0.024	0.118	-0.108	0.123	-0.043	0.113
College degree	-0.067	0.117	0.054	0.117	-0.095	0.119	-0.109	0.123	-0.112	0.126
Post college	-0.259	0.120	0.180	0.119	-0.083	0.121	-0.033	0.126	-0.170	0.124
Age										
25-34	-0.622	0.207	0.071	0.273	0.497	0.327	0.335	0.402	0.751	0.343
35-44	-0.407	0.194	0.241	0.268	0.538	0.324	0.346	0.400	0.685	0.347
45-54	-0.240	0.192	0.302	0.268	0.603	0.324	0.487	0.400	0.660	0.348
55-64	-0.091	0.195	0.593	0.270	0.495	0.327	0.446	0.401	0.798	0.347
65 +	-0.200	0.196	0.506	0.270	0.362	0.327	0.269	0.402	0.665	0.351
Risk averse	-0.341	0.078	-0.139	0.070	-0.150	0.068	-0.148	0.073	-0.103	0.077
Female	-0.288	0.103	-0.041	0.106	-0.223	0.110	-0.035	0.128	-0.121	0.132
Married	-0.049	0.085	0.176	0.088	-0.068	0.093	-0.058	0.110	-0.026	0.119
HH Size	-0.022	0.027	0.002	0.025	-0.037	0.026	-0.065	0.028	-0.044	0.029

How the marginal tax rate affected the incentive for rental housing investment in the run-up to 1986 TRA is another topic of interest. The argument in section 3.2 predicts that the marginal tax rate will be associated with the ownership of rental housing and the association will be particularly stronger for the high MAGI group because 1981 ERTA introduced the favorable depreciation rule, allowed households to report large tax losses on rental housing investment. However, the marginal tax rate coefficients are not statistically significant although both coefficients are positive. This result might be due to the fact that the data is just one year after the enactment of ERTA since the income variables of 1983 SCF is based on tax filing in 1982. Households might have taken some time to take advantage of new provisions in the tax code. At the same time, this result may well be caused by weak identification of variation in the marginal tax rate, especially if the large standard error for the high MAGI group is considered. This hypothesis seems plausible because the equivalent coefficient in the probit estimation for ownership of tax-deffered accounts turns out to be statistically insignificant for this year.

The result for year 1983 makes it difficult to derive a definite conclusion about another interesting topic: the adjustmet speed for households' portfolio. Scholz (1994) conducts the descriptive analysis of porfolio structure using the 1983 and 1989 Surveys of Consumer Finances and finds basically no change in real estate investment between these two years, which this study confirms in part by Table 3.1. Gordon (1994) suggests a potential explanation that the portfolio adjustment may be slow because of the long life of real property. If the marginal tax rate is positively associated with a high probability of rental housing ownership just before 1986, which could be tested more clearly if a data set of good quality for around 1985 were available, the combination of marginal tax rate coefficients for 1989 would suggest households adjusted their portfolio quite quickly in response to the tax reform, even including the real asset. If that hypothesis is true, the cross-sectional pattern of ownership of rental housing across the income groups,

Table 3.3: Average marginal effects of marginal tax rate on the probability of ownership

Year	1998	1995	1992	1989	1983
Residential rental properties					
MAGI<=100	0.186*	0.182*	-0.063	0.194*	0.109
	(1.94)	(1.88)	(-0.62)	(1.88)	(1.40)
MAGI>100	-0.151	0.109	-0.073	0.052	0.022
	(-0.96)	(0.66)	(-0.44)	(0.19)	(0.12)
Tax-deferred accounts					
MAGI<=100	0.092	0.396**	0.283**	0.413**	0.361**
	(0.89)	(3.73)	(2.65)	(3.82)	(3.72)
MAGI>100	0.590**	0.892**	0.256	0.850**	0.076
	(2.83)	(4.06)	(1.20)	(2.51)	(0.28)

Notes: Z-statistic for the original probit estimation are in the parentheses. Asterisks indicate statistical significance at the 5% (**) and 10% (*) significance levels. See text for further discussion.

which is presented in Table 3.1, is caused by the income effect, implying the income effect dominates the effect of progressive tax schedule based on the tax clientele model. However, only with the current resutls, no definite conclusion about the speed of portfolio adjustment can be drawn.

The coefficients for income and net worth variables are estimated as predicted in most of the years. Especially, high net worth has a strong association with the probablity of rental housing ownership while high income usually are associated with a higher probability. As a whole, none of the demographic control variables have an explanatory power. This result contrasts with Poterba and Samwick (2002), in which at least a college-level degree are associated with a higher probability of any categories of financial assets.

The anomalies found in the probit results may make someone doubt the reliability of identification power of the first-dollar marginal tax variable used in this study. Although Poterba and Samwick (2002) use the variable constructed by the same procedure and find the predicted associations between the marginal tax rate and the ownership probability

for some financial asset categories, they do not estimate the coefficient separately by income groups as this study does. Thus, as a reference, I pick tax-deferred accounts, for which the standard clientele model predicts positive association between the marginal tax rate and the ownership probability and estimate the probit model for this financial asset with the same variables as in Table 3.2.[16] Table 3.3 presents the average marginal effects calculated from the probit estimations for the ownership probabilities for residential rental properties and tax-deferred accounts, along with the Z-statistics in the original probit estimations. For 1989, 1995, and 1998, the coefficient for the marginal tax rate for the high MAGI group, which is consistent with the theory, although, as I mentioned earlier, the estimation fails to find statistically significant evidence for 1983 and 1992. From this result, the first-dollar marginal tax rate appears to be a relevant variable for an empirical study on the tax clientele model.

3.6 Conclusion

This study estimates the probit model for the probability of owning rental housing to test the prediction of the tax clientele model in the presence of the passive loss provisions introduced by TRA in 1986. The results offer modest support for the predictions although some of the results make interpretation difficult. The marginal tax rate is positively associated with the probability of rental housing ownership for the group of households to whom the exception for restriction on passive losses is applied for most of the years after TRA. The marginal tax rate indicates no association with rental housing for the other group, which is subject to the restriction.

The most diffecult part of this empirical study is the high correlation and endogeneity between the marginal tax rate and income. There are convincing arguments for

[16]The tax-deferred accounts variable in this study takes unity if the household has an asset of positive value in IRA account or Keogh account.

us to expect the probability of rental housing ownership is correlated with income itself, not necessarily through the marginal tax rate. Households with high income may be more willing to invest in risky assets, including real estates. Since purchasing real estates is commonly financed by debt to a considerable extent, the existence of credit constraint for low and middle income households will hinder them from rental housing investment. SCF provides detailed information needed to estimate tax liabilities and makes it possible to identify the variation in the marginal tax rate after controlling income. However, more detailed data on taxation would allow us to generate a more precise variable for the marginal tax rate and reduce the standard errors in the estimation.

References

Albanesi, Stefania (2011). "Optimal Taxation of Entrepreneurial Capital with Private Information". In: *Working Paper*.

Anand, Bharat N. and Richard Sansing (2000). "The Weighting Game: Formula Apportionment as an Instrument of Public Policy". In: *National Tax Journal* 53.2, pp. 183–199.

Angeletos, George-Marios (2007). "Uninsured idiosyncratic investment risk and aggregate saving". In: *Review of Economic Dynamics* 10.1, pp. 1–30.

Aoki, Shuhei and Makoto Nirei (2014). "Zipf's Law, Pareto's Law, and the Evolution of Top Incomes in the U.S."

Asea, Patrick K. and Stephen J. Turnovsky (1998). "Capital Income Taxation and Risk-Taking in a Small Open Economy". In: *Journal of Public Economics* 68, pp. 55–90.

Atkinson, Anthony B. and Joseph E. Stiglitz (2015). *Lectures on Public Economics*. Updated. Princeton University Press.

Auerbach, Alan J. and Mervyn A. King (1983). "Taxation, Portfolio Choice, and Debt-Equity Ratios: A General Equilibrium Model". In: *Quarterly Journal of Economics* 98.4, pp. 587–610.

Benabou, Roland (2002). "Tax and Education Policy in a Heterogeneous Agent Economy: What Levels of Redistribution Maximize Growth and Efficiency?" In: *Econometrica* 70.2, pp. 481–517.

Benhabib, Jess, Alberto Bisin, and Shenghao Zhu (2011). "The Distribution of Wealth and Fiscal Policy in Economies with Finitely Lived Agents". In: *Econometrica* 79.1, pp. 123–157.

— (2016). "The Distribution of Wealth in the Blanchard-Yaari Model". In: *Macroeconomic Dynamics* 20, pp. 466–481.

Berkovec, James and Don Fullerton (1992). “A General Equilibrium Model of Housing, Taxes, and Portfolio Choice”. In: *Journal of Political Economy* 100.2, pp. 390–429.

Blanchard, Olivier J. (1985). “Debt, Deficits, and Finite Horizons”. In: *Journal of Political Economy* 93.2, pp. 223–247.

Chambers, Matthew, Carlos Garriga, and Don E. Schlagenhauf (2009). “Housing Policy and the Progressivity of Income Taxation”. In: *Journal of Monetary Economics* 56.8, pp. 1116–1134.

Conesa, Juan Carlos and Dirk Krueger (2006). “On the optimal progressivity of the income tax code”. In: *Journal of Monetary Economics* 53.7, pp. 1425–1450.

Conesa, Juan Carlos, Sagiri Kitao, and Dirk Krueger (2009). “Taxing Capital? Not a Bad Idea After All!” In: *American Economic Review* 99.1, pp. 25–48.

Cullen, Julie Berry and Roger H. Gordon (2007). “Taxes and Entrepreneurial Risk-taking: Theory and Evidence for the U.S.” In: *Journal of Public Economics* 91.7-8, pp. 1479–1505.

Diamond, Peter A. and Emmanuel Saez (2011). “The Case for a Progressive Tax: From Basic Research to Policy Recommendations”. In: *Journal of Economic Perspectives* 25.4, pp. 165–190.

Domar, Evsey D. and Richard A. Musgrave (1944). “Proportional Income Taxation and Risk-Taking”. In: *Quarterly Journal of Economics* 58.3, pp. 388–422.

Duffie, Darrell and Larry G. Epstein (1992). “Stochastic Differential Utility”. In: *Econometrica* 60.2, pp. 353–394.

Eaton, Jonathan (1981). “Fiscal Policy, Inflation and the Accumulation of Risky Capital”. In: *Review of Economic Studies* 48.3, pp. 435–445.

Eaton, Jonathan and Samuel Kortum (2002). “Technology, Geography, and Trade”. In: *Econometrica* 70.5, pp. 1741–1779.

Edmiston, Kelly D. and F. Javier Arze del Granado (2006). “Economic Effects of Apportionment Formula Changes: Results from a Panel of Corporate Income Tax Returns”. In: *Public Finance Review* 34.5, pp. 483–504.

Fajgelbaum, Pablo D., Eduardo Morales, Juan Carlos Suárez Serrato, and Owen Zidar (2015). “State Taxes and Spatial Misallocation”.

Follain, James R., Patric H. Hendershott, and David C. Ling (1987). “Understanding the Real Estate Provisions of Tax Reform: Motivation and Impact”. In: *National Tax Journal* 40.3, pp. 363–372.

Gabaix, Xavier (2009). “Power Laws in Economics and Finance”. In: *Annual Review of Economics* 1.1, pp. 255–293.

Glaeser, Edward L. and Jesse M. Shapiro (2002). “The Benefits of the Home Mortgage Interest Deduction”. In: *NBER Working Paper*.

Golosov, Mikhail, Narayana Kocherlakota, and Aleh Tsyvinski (2003). “Optimal Indirect and Capital Taxation”. In: *Review of Economic Studies* 70.3, pp. 569–587.

Golosov, Mikhail, Aleh Tsyvinski, and Iván Werning (2007). “New Dynamic Public Finance: A User’s Guide”. In: *NBER Macroeconomics Annual*. Ed. by Daron Acemoglu, Kenneth Rogoff, and Michael Woodford. Vol. 21. MIT Press. Chap. 5, pp. 317–388.

Goolsbee, Austan and Edward L. Maydew (2000). “Coveting Thy Neighbor’s Manufacturing: the Dilemma of State Income Apportionment”. In: *Journal of Public Economics* 75.1, pp. 125–143.

Gordon, Roger H. (1985). “Taxation of corporate capital income: Tax revenues versus tax distortions”. In: *Quarterly Journal of Economics* 100.1, pp. 1–27.

— (1994). “Comments”. In: *Tax Progressivity and Income Inequality*. Ed. by Joel Slemrod. MIT Press, Cambridge, MA, pp. 268–274.

Gordon, Roger H. and Julie Berry Cullen (2012). “Income Redistribution in a Federal System of Governments”. In: *Journal of Public Economics* 96.11-12, pp. 1100–1109.

Gordon, Roger H. and Jeffrey K. Mackie-Mason (1994). “Tax Distortions to the Choice of Organizational Form”. In: *Journal of Public Economics* 55.2, pp. 279–306.

Gordon, Roger H. and John D. Wilson (1986). “An Examination of Multijurisdictional Corporate Income Taxation under Formula Apportionment”. In: *Econometrica* 54.6, pp. 1357–1373.

Gouveia, Miguel and Robert P. Strauss (1994). “Effective Federal Individual Income Tax Functions: An Exploratory Empirical Analysis”. In: *National Tax Journal* 47.2, pp. 317–339.

Heathcote, Jonathan, Kjetil Storesletten, and Giovanni L. Violante (2014). “Consumption and labor supply with partial insurance: An analytical framework”. In: *American Economic Review* 104.7, pp. 2075–2126.

Hilber, Christian A. L. and Tracy M. Turner (2014). “The Mortgage Interest Deduction and its Impact on Homeownership Decisions”. In: *Review of Economics and Statistics* 96.4, pp. 618–637.

Jones, Charles I. (2015). “Pareto and Piketty: The Macroeconomics of Top Income and Wealth Inequality”. In: *Journal of Economic Perspectives* 29.1, pp. 29–46.

Kawano, Laura (2014). "The dividend clientele hypothesis: Evidence from the 2003 tax act". In: *American Economic Journal: Economic Policy* 6.1, pp. 114–136.

Kenc, Turalay (2004). "Taxation, Risk-Taking and Growth: A Continuous-Time Stochastic General Equilibrium Analysis with Labor-Leisure Choice". In: *Journal of Economic Dynamics and Control* 28.8, pp. 1511–1539.

Kocherlakota, Narayana (2005). "Zero Expected Wealth Tax: A Mirrlees Approach to Dynamic Optimal Taxation". In: *Econometrica* 73.5, pp. 1587–1621.

Kotz, Samuel, Tomasz J. Kozubowski, and Krzysztof Podgórski (2001). *The Laplace Distribution and Generalizations*. Boston: Birkhäuser.

Mackie-Mason, Jeffrey K. and Roger H. Gordon (1997). "How Much Do Taxes Discourage Incorporation?" In: *Journal of Finance* 52.2, pp. 477–505.

Mazerov, Michael (2001). *The Single-Sales-Factor Formula: A Boon to Economic Development or a Costly Giveaway?* Tech. rep. Center on Budget and Policy Priorities.

McLure, Charles E. (1981). "The Elusive Incidence of the Corporate Income Tax: the State Case". In: *Public Finance Review* 9.4, pp. 395–413.

Melitz, Marc J. (2003). "The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity". In: *Econometrica* 71.6, pp. 1695–1725.

Merton, Robert C. (1971). *Optimum Consumption and Portfolio Rules in a Continuous-Time Model*.

Miller, Merton H. (1977). "Debt and Taxes". In: *Journal of Finance* 32.2, pp. 261–275.

Mirrlees, James A., ed. (2011). *Tax by Design*. Institute for Fiscal Studies. Oxford.

Modigliani, Franco and Merton H. Miller (1958). "The Cost of Capital, Corporation Finance and the Theory of Investment". In: *American Economic Review* 48.3, pp. 261–297. arXiv: arXiv:1011.1669v3.

Mossin, Jan (1968). "Taxation and Risk-Taking: An Expected Utility Approach". In: *Economica* 35.137, pp. 74–82.

Ozanne, Larry (2012). "Taxation of Owner-Occupied and Rental Housing". In: *Congressional Budget Office Working Paper*. Working Paper Series.

Piketty, Thomas and Emmanuel Saez (2013). "Optimal Labor Income Taxation". In: *Handbook of Public Economics*. Vol. 5. Elsevier B.V., pp. 392–474.

Piketty, Thomas, Emmanuel Saez, and Stefanie Stantcheva (2014). "Optimal Taxation of Top Labor Incomes: A Tale of Three Elasticities". In: *American Economic Journal: Economic Policy* 6.1, pp. 230–271.

Pinto, Santiago M. (2007). “Corporate Profit Tax, Capital Mobility, and Formula Apportionment”. In: *Journal of Urban Economics* 62.1, pp. 76–102.

Poterba, James M. (1984). “Tax Subsidies to Owner-Occupied Housing: An Asset-Market Approach”. In: *Quarterly Journal of Economics* 99.4, pp. 729–752.

— (1990). “Taxation and Housing Markets: Preliminary Evidence on the Effects of Recent Tax Reforms”. In: *NBER Working Paper*. arXiv: arXiv:1011.1669v3.

— (1993). “Tax Reform and Housing Market in the Late 1980s: Who Knew What, and When Did They Know It?” In: *Real estate and the credit crunch: Proceedings of a conference held at Melvin Village, New Hampshire, September 1992*. Federal Reserve Bank of Boston, pp. 230–251.

— (2002). “Taxation, Risk-Taking, and Household Portfolio Behavior”. In: *Handbook of Public Economics*. Ed. by Alan J. Auerbach and Martin Feldstein. Vol. 3. Elsevier, pp. 1109–1171.

Poterba, James M. and Andrew A. Samwick (2002). “Taxation and Household Portfolio Composition: US Evidence from the 1980s and 1990s”. In: *Journal of Public Economics* 87.1, pp. 5–38.

Poterba, James M. and Todd Sinai (2008). “Tax expenditures for owner-occupied Housing: Deductions for Property Taxes and mortgage interest and the exclusion of imputed rental income”. In: *American Economic Review* 98.2, pp. 84–89.

Reed, William J. (2001). “The Pareto, Zipf and Other Power Laws”. In: *Economics Letters* 74.1, pp. 15–19.

Rosen, Harvey S. (1985). “Housing subsidies: Effects on housing decisions, efficiency, and equity”. In: *Handbook of Public Economics*. Ed. by Alan J. Auerbach and Martin Feldstein. Vol. 1, pp. 375–420.

Rosen, Harvey S. and Kenneth T. Rosen (1980). “Federal taxes and homeownership: Evidence from time series”. In: *Journal of Political Economy* 88.1, pp. 59–75.

Runkel, Marco and Guttorm Schjelderup (2011). “The Choice of Apportionment Factors under Formula Apportionment”. In: *International Economic Review* 52.3, pp. 913–934.

Saez, Emmanuel (2001). “Using Elasticities to Derive Optimal Income Tax Rates”. In: *Review of Economic Studies* 68.1, pp. 205–229.

Saito, Makoto (1998). “A Simple Model of Incomplete Insurance the Case of Permanent Shocks”. In: *Journal of Economic Dynamics and Control* 22.5, pp. 763–777.

Sandmo, Agnar (1985). “The Effects of Taxation on Savings and Risk Taking”. In: *Handbook of Public Economics*. Ed. by Alan J. Auerbach and Martin Feldstein. Vol. 1, pp. 265–311.

Schmidt, Lawrence D. W. and Alexis Akira Toda (2015). "Do You Save More or Less in Response to Bad News? A New Identification of the Elasticity of Intertemporal Substitution".

Scholz, John Karl (1994). "Tax Progressivity and Household Portfolios: Descriptive Evidence from the Surveys of Consumer Finances". In: *Tax Progressivity and Income Inequality*. Ed. by Joel Slemrod. MIT Press, Cambridge, MA, pp. 219–267.

Stiglitz, Joseph E. (1969). "The Effects of Income, Wealth, and Capital Gains Taxation on Risk-Taking". In: *Quarterly Journal of Economics* 83.2, pp. 263–283.

Suárez Serrato, Juan Carlos and Owen Zidar (2014). "Who Benefits from State Corporate Tax Cuts? A Local Labor Markets Approach with Heterogeneous Firms".

Svensson, Lars E. O. (1989). "Portfolio Choice with Non-Expected Utility in Continuous Time". In: *Economics Letters* 30.4, pp. 313–317.

Toda, Alexis Akira (2012). "The Double Power Law in Income Distribution: Explanations and Evidence". In: *Journal of Economic Behavior and Organization* 84.1, pp. 364–381.

— (2014). "Incomplete Market Dynamics and Cross-Sectional Distributions". In: *Journal of Economic Theory* 154, pp. 310–348.

— (2015). "Asset Prices and Efficiency in a Krebs Economy". In: *Review of Economic Dynamics* 18.4, pp. 957–978.

Toda, Alexis Akira and Kieran Walsh (2015). "The Double Power Law in Consumption and Implications for Testing Euler Equations". In: *Journal of Political Economy* 123.5, pp. 1177–1200.

Yaari, Menahem E. (1965). "Uncertain Lifetime, Life Insurance, and the Theory of the Consumer". In: *Review of Economics Studies* 32.2, pp. 137–150.

Zipf, George K. (1949). *Human Behavior and the Principle of Least Effort*. Addison Wesley, Cambridge, MA.