The Dissertation Committee for Jijun Tan Certifies that this is the approved version of the following dissertation:

Essays on Economics of Taxation

Committee:

Don Fullerton, Supervisor

Li Gan

Jennifer Huang

Kim Ruhl

Roberton C. Williams

Essays on Economics of Taxation

by

Jijun Tan, B. E.; M. E.; M. A.

Dissertation

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

The University of Texas at Austin

May, 2006

Dedication

To my wife, Dr. Xianmei Yang

To my parents, Kuannian Tan and Xiaojin Wang

And to my children, Ruiyang Tan and Xinyang Tan

Acknowledgements

I thank my supervisor Dr. Don Fullerton for his academic guidance and generous support throughout the years. He introduced me to the field of public economics and environmental economics that I have so enjoyed. My sincere appreciation goes to members of my dissertation committee, Drs. Li Gan, Jennifer Huang, Kim Ruhl, and Roberton C. Williams, for their constructive criticisms and encouragement. I acknowledge the help I received from Drs. Guan Gong, Daniel Hamermesh, Yingyao Hu, Shuyun Li, and Daniel Slesnick. Throughout my thesis work, I am grateful for the friendship and support of my classmates, Shutao Cao, Ye Feng, Laura Juarez, and Ping Yan. Finally, I would like to thank my wife Dr. Xianmei Yang, my parents, parents-in-law, and sisters and brothers for their strong support during the days of my graduate study.

Essays on Economics of Taxation

Publication No.______________

Jijun Tan, Ph. D.

The University of Texas at Austin, 2006

Supervisor: Don Fullerton

The first chapter is an introduction and overview. Using the CES utility function, Chapter 2 numerically examines the relationship between the optimal tax-and-transfer systems and inequality of earnings under major alternative social welfare functions. In a one-bracket linear tax system, both the optimal income tax rate and the government transfer increase when earning inequality expands. In the two-bracket case, the optimal lower bracket rate and income threshold do not change in a way that is monotonic. The optimal upper bracket rate and government transfer increase with the wage spread. The lower bracket rate is greater than the upper bracket one when the spread is small, but it is larger when the spread is large. With a large elasticity of substitution between consumption and leisure, the two-bracket tax structure converges to the one-bracket case when the wage spread becomes large.

Chapter 3 examines the distributional effects of an environmental tax on the price system in a spatial model of a closed city. "Social welfare" is defined over identical residents and is affected both by environmental quality and by the rent paid to absentee landlords. The tax improves environmental quality everywhere. However, it reduces

the equilibrium wage. As the tax rate rises from low to high, the tax first improves social welfare until it hits the optimal level, and then it reduces welfare. The tax first makes the city boundary shrink, but then makes it grow. Initial increases of the tax pull down the rent for any particular location in the city, and further increases pull the rent back up.

Chapter 4 empirically tests the effect of the extra corporate tax on the choice of organizational form between corporate and non-corporate form. The overall extra corporate tax rate has significantly negative effects on the corporate share of economic activity such as capital stock and investment. My estimated effect on the corporate share of capital stands just between two major estimates in the literature; while the effect on the corporate share of investment is much larger than that on capital. The results are consistent with transaction costs in two respects: first, the corporate share of capital stock does show adjustment lags; second, the extra tax shows larger effect on the corporate share of investment than on that of capital stock.

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction

Tax performs very important roles in an economy. For example, the personal and corporate income taxes are used to redistribute income so that an equal economy can be achieved; the environmental tax is used to control pollutions so that a clean economy can be achieved. Thus, my dissertation focus on tax related but still deficiently explored topics such as the optimal two-bracket personal income taxation, distributional effects of an environmental tax in a closed city, and effects of the extra corporate tax on the choice of organizational forms, which are sequentially and respectively investigated in chapter 2, 3, and 4.

In the past two decades, both the American economy and many other economies have seen rising inequality of earnings. During almost the same period, eligibility requirements for welfare recipients are tightened in many countries. For example, the U. S. in 1996 replaced AFDC (Aid to Families with Dependent Children) by TANF (Temporary Aid for Needy Families). Nevertheless, the U. S. cuts the top personal income tax rate to a relatively very low level. Thus, both the enlarging of inequality and the shrinking of redistribution programs have occurred since the 1980's. Naturally, one might think that welfare programs should be designed so that when earning dispersion rises, support for those at the bottom would also rise.

The purpose of chapter 2 "How does the Optimal Two-Bracket Income Tax Depend on Wage Inequality" is to analyze how a tax-and-transfer system featuring a lump-sum transfer and a one-bracket linear income tax or a two-bracket piecewise-linear income tax might optimally be altered in response to changes in the underlying inequality of earnings. Using the same constant elasticity of substitution (CES) utility function that is commonly used in the literature, I numerically examine the relationship between the optimal tax-and-transfer systems and inequality of earnings under major alternative social welfare functions (SWF) such as the Bentham SWF and the Nash SWF. If the tax-and-transfer system has two income brackets, comparison of the magnitudes of both marginal tax rates is of interest.

In the one-bracket case, I find that both the optimal income tax rate and the negative intercept (government transfer) both become larger when earning inequality becomes more serious. In the two-bracket case, I find that the marginal rate of the lower bracket is greater than that of the upper bracket when the spread of the wage is relatively small, but it is larger when the spread is relatively large. Beyond that, surprisingly, I find that with a relatively large elasticity of substitution between consumption and leisure in the consumer's utility function, the two-bracket tax structure converges to the one-bracket structure when the wage spread becomes relatively quite large. Furthermore, though the optimal lower bracket rate and income threshold do not show monotonicity, the optimal upper bracket rate and government transfer are increasing with the wage spread.

A particular environmental tax may have distributional effects across different factor owners, such as workers and landowners. Compared to the other taxes, environmental taxes are less studied in the tax incidence literature. Yet the environmental tax is introduced to cure the distortion of an economy rather than to be a distorting factor of the economy like other taxes. So, the distributional effect of an environmental tax is of interest itself. Furthermore, when spatial characteristics are considered, environmental quality can be treated as both a private good and a public good. This makes the distributional effect of an environmental tax more interesting.

Chapter 3 "The Distributional Effects of a City's Environmental Tax" examines the distributional effects of an environmental tax on pollution emissions on the price system in a general equilibrium spatial model of a closed linear city where living location is a choice of the residents. Pollution emissions are generated in the central business district (CBD) of the city and then spread over the whole city. “Social welfare” of the city is defined over the many identical residents, and it is affected both by the quality of the environment and by the rent they pay to absentee landlords.

I find that an environmental tax in a closed city targeting polluting emissions can reduce the pollution emissions in the city’s CBD, and therefore improves environmental quality of the city everywhere. However, the tax reduces the wage received by labor. Overall, as the environmental tax rate rises from quite low to quite high, the tax first improves the social welfare of the city until it hits the optimal level, and then it reduces welfare as the marginal benefit falls below the marginal cost of the tax. Furthermore, for a fixed number of residents, the tax first makes the city boundary shrink, but then makes it grow as residents try to balance between better housing and shorter commuting. Initial increases in the tax pull down the rent for any particular location in the city and further increases pull the rent back up.

The environmental tax has two possible effects on the whole rent gradient because it does not affecting the rent of every location uniformly. In the fist case, when the tax increases from small to large, all rents first decrease, forming a downward shift of the rent gradient, then the rents close to the CBD turn back to increase while those away from the CBD still decrease, forming a clockwise pivoting, and at last, all rents increase, forming an upward shift. In the second case, all rents first decrease, forming a downward shift of the rent gradient, then the rents away from the CBD turn back to increase while those close to the CBD still decrease, forming a counter-clockwise pivoting, and at last, all rents increase, forming an upward shift.

In a quite polluted city with an inadequate environmental tax, the increase of the tax shifts down the rent gradient, decreasing the rent of land everywhere in the city. In contrast, for a very clean city with a high environmental tax, an increase of the tax shifts up the rent gradient, increasing the rent of land everywhere. In one type of theoretically defined city, as the environmental tax increases from quite low to quite high, it first improves the social welfare of the city, shrinks the size of the city , and cuts the rent of any particular location in the city; and then it reduces the social welfare, expands the city, and raises the rent anywhere. In another type of theoretically defined city, furthermore, a Pareto improvement is possible, which means that the tax may increase both the social welfare of the city residents and the gradient of rent received by landlords.

The corporate income tax has been criticized for a long time because of the double taxation on corporation shareholders. Nevertheless, the previous literature has conflicting conclusions on the effect of the extra tax on the distribution of economic activity between the corporate and non-corporate sectors. Gravelle and Kotlikoff (1989 and 1993) theoretically find huge deadweight loss (DWL) caused by the extra corporate tax, because it distorts the choice of organizational forms. Mackie-Mason and Gordon (1997) and Goolsbee (1998), however, empirically find that the effect of the extra corporate tax on the corporate share of economic activity, especially on the corporate share of capital, is statistically significant but every small in magnitude.

Chapter 4 “The Effect of Corporate Income Tax on Organizational Forms” empirically tests the effect of the extra corporate tax on the choice of organizational form between corporate and non-corporate form. It follows the works of Mackie-Mason and Gordon (1997) and Goolsbee (1998) by reexamining the effect of the extra corporate tax on the corporate share of capital, by using new data with a longer time span than used in previous investigations, by studying the role of transaction costs played in the effect of the extra corporate tax on changes of organizational forms, and by estimating the effect via new investment data instead of capital stock data used in the literature.

I find that the overall extra corporate tax rate has significantly negative effects on the corporate share of economic activity, which is consistent with findings in the previous literature. The effect of the extra corporate tax on the corporate share of capital stock, which is estimated by data with a longer time span than those used in the literature, stands just between two major estimates in the literature. Furthermore, the effect of the extra corporate tax on the corporate share of investment is much larger than the effect on capital. These results are consistent with transaction costs in two respects: first, the corporate share of capital stock does show adjustment lags; second, the extra corporate tax shows larger effect on the corporate share of investment than on the corporate share of capital stock. Investment can react more easily than capital stock. If capital stock could be added or subtracted with less transaction costs, then the corporate share of capital would have responded to the extra corporate tax as sensitively as the corporate share of investment.

Furthermore, estimates here account for the fact that the economy tends to increase the share of corporate assets during war time. Finally, the structural changes of tax system caused by the Tax Reform Act of 1986 encouraged people to increase the share of non-corporate assets. These structural changes may affect the way that a change in corporate tax rate would affect the share of investment going into the corporate sector.

Chapter 2: How does the Optimal Two-Bracket Income Tax Depend on Wage Inequality

2.1 INTRODUCTION

Rising inequality of earnings is one of the most remarkable characteristics of both the American economy and many other economies. In the past two decades, as described by Rosen (2004) in his popular public economics text book, Americans have seen both higher relative earnings for those at the top compared to the median as well as lower relative earnings for those at the bottom compared to the median. During this period, more serious inequalities have also been seen in many other countries such as China. Also during this period, some countries have tightened up eligibility requirements for welfare recipients. For example, the U. S. in 1996 replaced AFDC (Aid to Families with Dependent Children) by TANF (Temporary Aid for Needy Families). At the other end of the income spectrum, the U. S. has cut the top personal income tax rate to a relatively very low level. In essence, both the enlarging of inequality and the shrinking of redistribution programs have occurred since the 1980's. These two changes seem at odds with each other. One might think that welfare programs should be designed so that when earning dispersion rises, support for those at the bottom would also rise.

The economics literature has not addressed this question completely. Mirrless (1971), Stern (1976) and Cooter and Helpman (1974) show that the optimal rate of the one-bracket income tax should increase when ability is more unequally distributed among people. However, the research of Helpman and Sdaka (1978) argues that the optimal rate does not always increase. With respect to the relation between growing earning inequality and the optimal multiple-bracket income tax, economists have not paid much attention.

The purpose of this paper is to analyze how a tax-and-transfer system featuring a lump-sum transfer and a one-bracket linear income tax or a two-bracket piecewise-linear income tax might optimally be altered in response to changes in the underlying inequality of earnings. Using the same constant elasticity of substitution (CES) utility function that is commonly used in the literature, I numerically examine the relationship between the optimal tax-and-transfer systems and inequality of earnings under major alternative social welfare functions (SWF) such as the Bentham SWF and the Nash SWF. If the tax-and-transfer system has two income brackets, comparison of the magnitudes of both marginal tax rates is of interest.

In the one-bracket case, I find that both the optimal income tax rate and the negative intercept (government transfer) both become larger when earning inequality becomes more serious. In the two-bracket case, I find that the marginal rate of the lower bracket is greater than that of the upper bracket when the spread of the wage is relatively small, but it is larger when the spread is relatively large. Beyond that, surprisingly, I find that with a relatively large elasticity of substitution between consumption and leisure in the consumer's utility function, the two-bracket tax structure converges to the one-bracket structure when the wage spread becomes relatively quite large. Furthermore, though the optimal lower bracket rate and income threshold do not show monotonicity, the optimal upper bracket rate and government transfer are increasing with the wage spread.

My paper is different from the literature in the following respects. First of all, I examine continuous changes of earning inequality in the one-bracket case, while the existing literature has investigated only a few discrete cases of earning inequality. My research thus clarifies any ambiguity in the comparison among those discrete cases. Second, I also focus on effects of enlarging earning inequality on the optimal two-bracket income tax, effects not addressed in the literature. I find some new and interesting results.

The paper proceeds as follows. Section 2 reviews the optimal income tax literature; Section 3 investigates the relationship between earning inequality and the one-bracket optimal income tax using theoretical distributions of ability; Section 4 similarly studies the relationship between earning inequality and the two-bracket optimal income tax, finally, Section 5 presents my conclusions.

2.2 LITERATURE REVIEW

Early literature calculates the optimal income tax rate assuming a linear income tax, a log utility functional form, a lognormal distribution of ability, and a particular social welfare function (SWF) to be maximized. For example, Mirrless (1971) shows that the optimal one-bracket income tax rate is increased when ability is more unequally distributed. He uses only two discrete cases of ability dispersion to address this tendency, however, one representing a moderate level of dispersion and the other representing an extremely high level of dispersion. Also, he does not investigate this problem based on a mean preserving process, so the change in mean wage may have confounded his result.

Stern (1976) reexamines Mirrless' simulation results for the optimal one-bracket income tax using the same distributions of ability as used by Mirrless (1971) but with a different utility functional form, a CES utility function with a smaller elasticity of substitution between consumption and leisure. He also shows that the optimal one-bracket income tax rate tends to increase when the spread of the ability distribution rises. He still does not control for the mean, and he uses exactly the same two discrete cases used by Mirrless.

Cooter and Helpman (1974) also use a CES utility function to examine the relationship between the optimal one-bracket income tax and ability dispersion. They use three different types of ability distributions to represent low, medium, and high ability inequality. Thus they have three cases instead of two, and they do keep the mean unchanged in their simulations. Still, however, these distributions differ from each other not only in ability dispersion, but also in many other respects. Their simulation results show that the optimal one-bracket income tax rate tends to increase under any of the seven social welfare functions used in their paper.

All three of the above studies find that the optimal one-bracket income tax rate increases when ability is more widely distributed, but Helpman and Sdaka(1978) argue that theoretically this rate does not always increase and is not determined in general. Slemrod and Bakijia (2000) survey the papers mentioned above and tend to discount the conclusion of Helpman and Sdaka (1978). They agree that the optimal one-bracket income tax should rise when earning inequality becomes more serious.

Besides the one-bracket income tax, economists have also studied more complicated taxation structures, such as the multiple-bracket income tax. Simulation results of Mirrless (1976) find that rates of the optimal marginal income tax including the rate of the top bracket should be greater than zero. In theoretical work, however, Phelps (1973), Sadka (1976), and Seada (1977) argue that the optimal income tax rate of the very top person should be zero (because to change the rate from any positive number to zero is a Pareto improvement). Stiglitiz (1982, 1987) also agrees that the person with highest ability should have a zero marginal income tax rate, and the person with lowest ability should have a positive marginal income tax. Basically, economists suggest a zero marginal tax rate of the top person because such a tax rate can encourage the richest person to work more and thus to improve total social welfare. The debate is not over.

Sheshinski (1989) stands out, saying that a smaller upper bracket tax rate with a larger lower bracket tax rate is not optimal. However, Slemrod, Yitzhaki and Mayshar, and Lundholm (1994) point out that Sheshinski's proof is not reasonable. Furthermore, they show in their many simulations that the optimal marginal income tax rate of the upper bracket is smaller than that of the lower bracket.

Whereas that body of work employs two or three levels of ability dispersion, this paper investigates the whole spectrum to see how the optimal tax rate is affected by each increment to the variance of wages (holding the mean constant). In addition, whereas that body of work looks at the effect of wage dispersion on the one-bracket rate, this paper looks at effects on both rates of a two-bracket income tax. Whereas Slemrod *et al* (1994) consider only one level of wage dispersion and find that the second-bracket rate is lower than the first-bracket rate, this paper shows that the reverse pattern occurs for higher wage dispersion. Moreover, evidence suggests that the higher level of wage dispersion is now more relevant for the U. S., and especially other countries. Thus the optimal second-bracket rate is likely higher than the low-bracket rate.

2.3 EARNING INEQUALITY AND THE OPTIMAL ONE-BRACKET LINEAR INCOME TAX

2.3.1 The Model

Consider a simple model with N agents who have identical preferences given by the utility function:

(2.1) $U(c_i, 1\text{-}h_i)$

where c_i is the consumption of individual i, and h_i is her labor supply. This function is nicely behaved in the sense that $U_1 > 0$, $U_2 > 0$, $U_{11} < 0$, and $U_{22} < 0$. Each individual has a time endowment of one and may split it between leisure and labor.

Each individual randomly gets ability and corresponding wage w according to a probability distribution with $f_w(w)$ as its p.d.f. and $F_w(w)$ as its C.D.F.. Wages accepted by individuals are independent of one another. A government maximizes a particular social welfare function using a one-bracket linear tax-and-transfer system that has a lump-sum benefit b to all individuals and a constant marginal tax rate t. I assume that each individual uses all her income to consume and does not save, no matter whether she receives wage income or government transfer.

The individual's budget constraint is $b+(1-t)\, w_i h_i = c_i$. Thus, given w_i, b, and t, individual i maximizes:

(2.2) $U\,[b+(1-t)\, w_i h_i,\, 1-h_i]$

by choosing her labor supply, h_i. This generates her labor supply function $h_i(w)$. It is straightforward to see that individual i participates in the labor market as long as:

(2.3) $U\,[b+(1-t)w_i\, h_i(w_i),\, 1-h_i(w_i)] \geq U(b,1)$

where $U(b,1)$ is the utility that individual i can get if she does not provide any labor. Let $w_i{*}$ be the wage at which individual i is indifferent between working and not working , i.e.:

(2.4) $U\,[b+(1-t)w_i{*}\, h_i(w_i{*}),\, 1-h_i(w_i{*})] = U(b,1)$

This $w_i{*}$ indicates the no-envy wage developed by Foley (1967) and Varian (1974). Let:

(2.5) $P_w = \Pr\,(w_i \geq w_i{*})$

be the probability that individual i works. Given that w_i has a C.D.F., $F_w(w_i)$, the probability can be written as:

(2.6) $P_w = 1 - F_w\,(w_i{*})$

Because individuals are *ex ante* identical and their wages are independent of one another, subscript i in (2.1) to (2.6) can be ignored, so that all individuals have exactly the same equations.

By choosing b and t, the government maximizes a particular social welfare function (SWF) subject to a balanced government budget constraint:

(2.7) $\Sigma_i\, t\, w_i\, h(w_i) = N\, b$

where the left side of the equation is the revenue of the government and the right side is the expenditure of the government. When the population is big enough, equation (2.7) can be written as:

(2.8) $\mathrm{E}\,[t\, w\, h(w)] = b$

Substituting w^* into (2.8), I get:

(2.9) $P_w\, \mathrm{E}\,[t\, w\, h(w) |\, w > w^*] = b$

Given $F_w(w)$ and $f_w(w)$, the balanced government budget constraint can be rewritten as:

(2.10) $t\int_{w^*}^{\infty} w h(w)\, f_w(w)\, dw = b$

Social welfare functions of the government could include the Bentham SWF and the Nash SWF, both of which are utilitarian social welfare functions. Under each different SWF, the optimizing problem of the government is different:

1. The Bentham SWF. Under this criterion, the government maximizes the un-weighted sum of everybody's utility. CES utility $U\,(c_i, 1\text{-}h_i)$ is homothetic, but the marginal utility of consumption c declines with the amount of consumption, and so even the un-weighted sum of utilities can be raised by redistribution from a person with high c to a person with low c.[1] So, the government is averse to unequal

[1]If income were used to buy two goods X and Y, where $U(X, Y)$ is a CES or other homothetic utility functions, then the marginal utility of income is constant, and redistribution of income cannot raise the un-weighted sum of utilities. In my case, however, no redistribution ($t = b = 0$) would mean that each person uses endowment $w_i{\cdot}1$ to maximize $U(w_i h_i, 1\text{-}h_i)$. To see that some redistribution can increase

consumption. The government maximizes the expected utility of a single person when the population is big enough, because individuals are *ex ante* identical and their wages are independent of one another. So, the government chooses t and b to maximize:

(2.11) $E\,\{[U[(1\text{-}t)wh(w),\, 1\text{-}h(w)\}$

subject to (2.8). Given $F_w(w)$ and $f_w(w)$, the government chooses t and b to maximize:

(2.12) $\int_{w^*}^{\infty} U[(1\text{-}t)wh(w),\, 1\text{-}h(w)]\, f_w(w)\, dw + F_w(w^*)\, U(b, 1)$

subject to (2.10).

2. The Nash SWF. Under this criterion, the government is averse to unequal utility itself; it maximizes the un-weighted product of the utility of all individuals. When the population is big enough, the government maximizes the expectation of the log of utility of a single person, choosing t and b to maximize:

(2.13) $E\,(\log\{U[b+(1\text{-}t)wh(w),\, 1\text{-}h(w)]\})$

subject to (2.8). Given $F_w(w)$ and $f_w(w)$, the government chooses t and b to maximize:

(2.14) $\int_{w^*}^{\infty} \log\{U[b+(1\text{-}t)wh(w),\, 1\text{-}h(w)]\}\, f_w(w)\, dw + F_w(w^*) \log[U(b, 1)]$

subject to (2.10).

2.3.2 The Utility Functional Form

For comparability to the literature, I choose the CES utility function following Cooter and Helpman (1974), Stern (1976), and Slemrod, *et al* (1994). Let the functional form of (2.1) be:

(2.15) $[\alpha c_i^{\,(\sigma\text{-}1)/\sigma} + (1\text{-}\alpha)(1\text{-}h_i)^{\,(\sigma\text{-}1)/\sigma}\,]^{\,\sigma/(\sigma-1)}$

---

welfare in this case, consider the simple example where preferences involve inelastic demand for leisure $(1\text{-}h_i)$. Then $U(w_ih_i, 1\text{-}h_i)$ can mean every unequal distribution of consumption $c_i = w_ih_i$, and concavity in c means that $b > 0$ can help raise total welfare.

where σ is the elasticity of substitution between consumption and leisure, and α is the weight on consumption. Given this function, the individual chooses h_i to maximize:

(2.16) $\{\alpha[b+(1-t)w_i h_i]^{(\sigma-1)/\sigma} + (1-\alpha)(1-h_i)^{(\sigma-1)/\sigma}\}^{\sigma/(\sigma-1)}$

By solving (16), I get:

(2.17) $h_i(w_i) = \{1-b[(1-\alpha)/\alpha]^{\sigma}[(1-t)w_i]^{-\sigma}\}/\{1+[(1-\alpha)/\alpha]^{\sigma}[(1-t)w_i]^{1-\sigma}\}$

Inequality (2.3) becomes:

(2.18) $\{\alpha[b+(1-t)w_i h_i(w_i)]^{(\sigma-1)/\sigma} + (1-\alpha)[1-h_i(w_i)]^{(\sigma-1)/\sigma}\}^{\sigma/(\sigma-1)} \geq U(b, 1)$

where $U(b, 1) = [\alpha\ b^{(\sigma-1)/\sigma} + (1-\alpha)]^{\sigma/(\sigma-1)}$ is individual i's utility when she stays outside the labor market. The wage rate that makes inequality (2.18) into an equation is:

(2.19) $w_i{}^* = \{b^{1/\sigma}[(1-\alpha)/\alpha]\}/(1-t)$

Individual i will work if and only if $w_i \geq w_i{}^*$. Again, since individuals are *ex ante* identical and their wages are independent of one another, subscript i in (2.15) to (2.19) can be ignored, which means that all individuals can have exactly the same equations.

2.3.3 Simulation Results with a Relatively Small σ

In order to find how the optimal tax-and-transfer system depends on earning inequality, I first show how the values of the optimal one-bracket income tax rate t and government transfer b change with a mean-preserving spread of earning inequality. My interpretation is that the increase of the standard deviation of a particular wage distribution describes an increased dispersion of earnings only, with no other changes (such as the mean wage). So, each particular value of the standard deviation has at least one corresponding pair of values for the optimal tax rate and transfer. By investigating those values, I may see the relationship between the tax-and-transfer program and earning inequality.

Table 2.1: Key Elasticities for Labor Supply of the Mean Person[2]

(evaluated at t = 0.224, b= 0.057, and α= 0.6136)

	σ = 0.4	σ = 1.0
Uncompensated Labor Supply Elasticity	-0.141	0.133
Compensated Labor Supply Elasticity	0.232	0.649
Income Elasticity	-0.373	-0.517

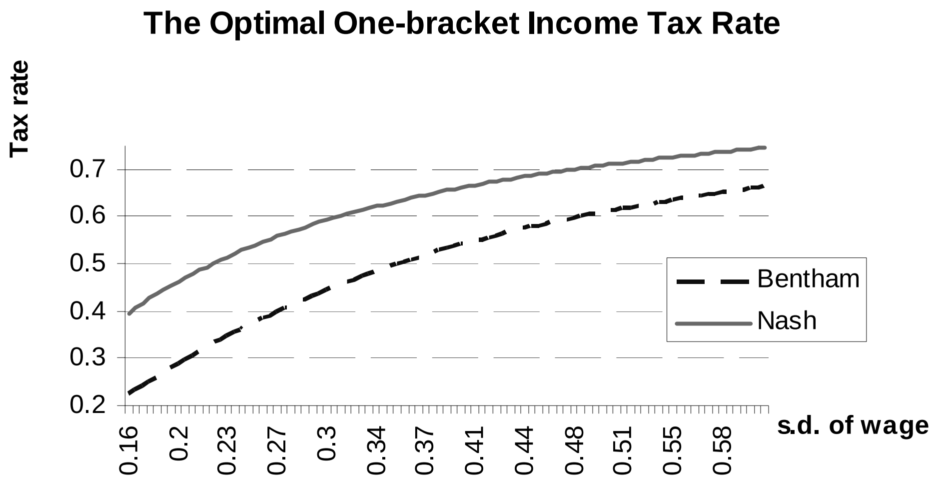
In my simulation, I first assume that the wage distribution is lognormal with a mean of 0.3969 as found by Lydall (1968) and used by Mirrless (1971) and Stern (1976) in their simulations.[3] This mean wage rate represents the labor income of the person with mean wage who uses all her time endowment to work and does not rest. Mirrless (1971) says that the lognormal distribution is "intended to represent a realistic distribution of skills within the population". Following Stern (1976), I set the elasticity of substitution between consumption and leisure in the CES utility function at σ = 0.4, and the consumption weight at α = 0.6136. Stern (1976) argues that σ = 0.4 is a more realistic value than σ = 1 used by Mirrless (1971).[4] Changes of σ cause changes of the elasticities for labor supply as shown by Table 2.1.

[2]The income elasticity is calculated by $(\partial h/\partial b)\{[b+(1-t)wh]/h\}$. The uncompensated labor supply elasticity is not zero when σ is set to 1.0 (Cobb-Douglas utility) is because my model has non-labor income (the government transfer). The compensated labor supply elasticity is calculated by Slutsky equation: compensated elasticity = uncompensated elasticity – income elasticity.

[3]The 0.3969 is the mean of the lognormal distribution used by Mirrless (1971), Stern (1976), and Slemrod *et al* (1994), the corresponding normal distribution of which has a mean of -1 and a variance of 0.39. Mirrless (1971) uses this value first. He, however, does not indicate what the real meaning of the values is and only says it is derived from a table of Lydall (1968).

[4]The reason that 0.6136 is chosen as the value of α by Stern (1976) is because when σ and α are set to 0.5 and 0.6136, a person facing no tax and transfer would to use two thirds of her time endowment to work.

Figure 2.1: The Optimal One-Bracket Income Tax Rate (σ = 0.4)



In my simulation, I change the standard deviation (s.d.) of the wage gradually from 0.1609 to 0.6109 by increments of 0.005, and for each value we calculate the optimal tax rate and transfer, keeping the mean wage constant at 0.3969. So, the coefficient of variation (c.v.) of wage changes from 0.405 to 1.539.[5] The 0.1609 represents very moderate earning inequality, as used by Mirrless (1971) and Stern (1976), while the 0.6109 represents quite serious earning inequality.[6]

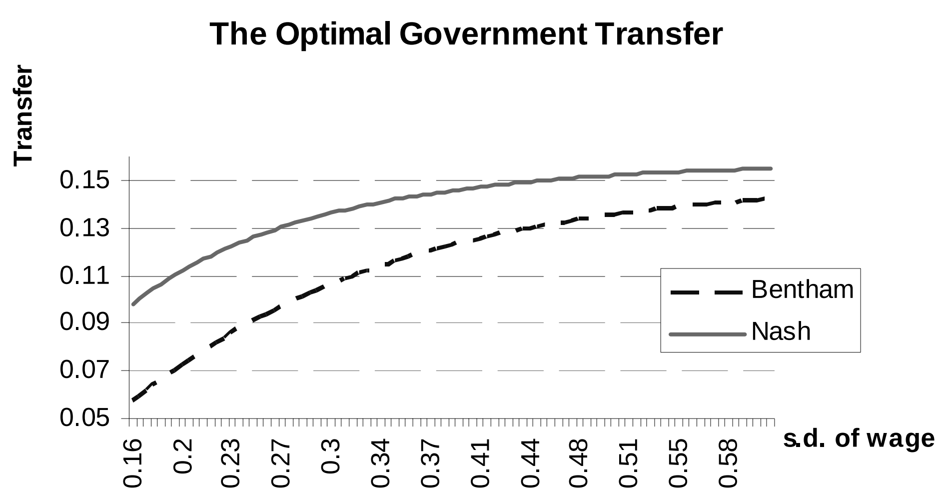
Figure 2.1 shows how the optimal one-bracket income tax rate reacts under both the Bentham SWF and the Nash SWF when earning inequality changes from the moderate level to the serious level. I find that under both SWFs, the optimal rate is

[5]The coefficient of variation of the wage rate in the U. S. varies from 0.590 to 0.888 during the period between 1979 and 2004 (by data from CPS MORG 1979-2004, NBER). That of Mexico varies from 1.561 to 2.721 during the period between 1995 and 1999 (by data from INEGI). Assuming that income inequality is highly correlated with wage inequality, I expect to see even larger values from most other developing countries due to the famous Kuznets Curve (Kuznets, 1955) that says income inequality increases when a country starts to be industrialized but finally decreases when it becomes a developed country. Glaeser (2005) confirms this relationship. An updated Kuznets Curve with 1998 data from the World Bank can be found in his paper.

[6]The 0.1609 is the s.d. of the lognormal distribution used by Mirrless (1971), Stern (1976), and Slemrod *et al* (1994), the corresponding normal distribution of which has a mean of -1 and a variance of 0.39.

strictly increasing with the standard deviation of wage with no exceptions. When the spread is relatively low, such as 0.1609, the optimal rate is 0.224 under the Bentham SWF (0.397 under the Nash SWF).[7] When the spread is extremely large, such as 0.6109, the optimal rate is as big as 0.664 under the Bentham SWF (0.745 under the Nash SWF). Intuition here is straightforward. When earning inequality becomes more serious, more individuals drop into the low income class and depend on government transfer to live. Thus, the government needs to collect more revenue from those working to subsidize the others.

Figure 2.2: The Optimal Government Transfer (σ = 0.4, one-bracket)



Beyond this, the optimal rates under the Bentham SWF are always lower than under the Nash SWF. This is because the Nash SWF puts more weight on the utility of the poor than does the Bentham. Hence, the government needs to have higher tax rates that can collect more revenue to finance more transfers to the poor. With respect to the government transfer b, it is also strictly increasing with the wage spread under both

[7]Stern (1976) gets an optimal tax rate of 0.223 under the Bentham SWF. The rest of the horizontal axis in Figure 1 is new.

SWFs as shown in Figure 2.2. The optimal transfer grows from 0.057 (roughly 14.4% of the mean wage) to 0.142 (35.8%) under the Bentham SWF (from 0.098 (24.7%) to 0.155 (39.1%) under the Nash SWF) when the standard deviation increases from 0.1609 to 0.6109. As expected, the optimal transfers under the Nash SWF are all bigger than under the Bentham SWF.

2.3.4 Simulation Results with a Relatively Large σ

Though Stern (1976) believes that the small value of the elasticity of substitution between consumption and leisure ($\sigma = 0.4$) is more realistic than larger values, the value of $\sigma = 1$ used by Mirrless (1971) is still of interest at least for comparison. As shown in Figure 2.3 and 2.4, we repeat the simulations above with $\sigma = 1.0$, while holding other parameters unchanged. Basically, a larger σ means a larger uncompensated labor supply elasticity as shown by Table 2.1.

I find that the change of σ from 0.4 to 1.0 does not affect my conclusion that the optimal one-bracket linear income tax rate and government transfer are strictly increasing with the wage spread. However, the increase of σ shifts down both the optimal tax rate and the optimal transfer. For the s.d. = 0.1609 used by Mirrless (1971) and Stern (1976), as shown in Figure 2.3, the optimal rate drops from 0.224 to 0.126 under the Bentham SWF. As shown in Figure 4, the optimal transfer drops from 0.057 to 0.029 under the Bentham SWF (The optimal rate drops from 0.397 to 0.229, while the optimal transfer drops from 0.098 to 0.050 under the Nash SWF). Obviously, increases of the uncompensated labor supply elasticity force the government to implement smaller and smaller tax rates. By using lower tax rates, the government encourages elastic workers to work, so that enough revenue can be collected from them to finance government transfers.

Figure 2.3: The Optimal One-Bracket Income Tax Rate ($\sigma = 1.0$)

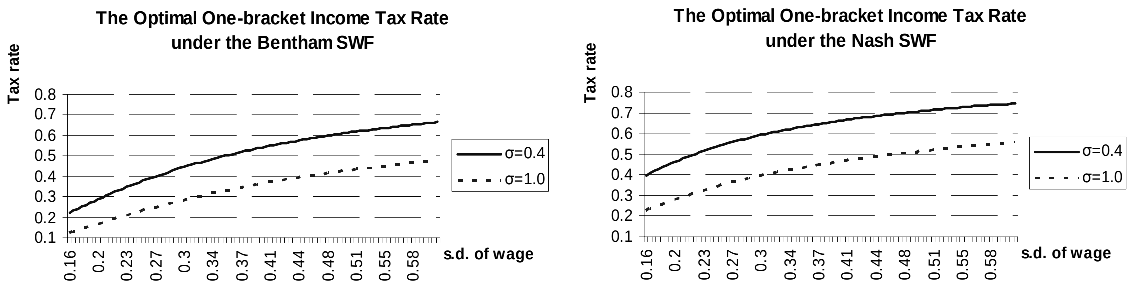
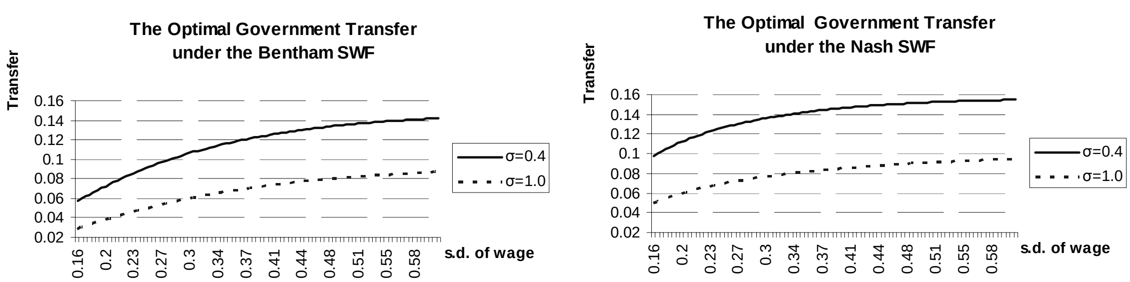


Figure 2.4: The Optimal Government Transfer ($\sigma = 1.0$, one-bracket)



2.4 EARNING INEQUALITY AND THE OPTIMAL TWO-BRACKET LINEAR INCOME TAX

2.4.1 The Model

I continue to use the same SWFs and lognormal distributions of ability as used in the previous section, and in this section I investigate the relationship between earning inequality and the optimal two-bracket linear income tax. Compared to the one-bracket case, the two-bracket case is more complicated. Though the preferences of individuals have not changed, the budget constraint has changed because of the introduction of the second marginal tax rate. Explicitly, individual i now chooses h_i to maximize:

(2.20) $U\,[b+(1-t_1)\min(w_ih_i, \hat{Y})+(1-t_2)\max(w_ih_i-\hat{Y}, 0), 1-h_i]$

where h_i is the labor supply of individual i, t_1 is the marginal tax rate of the first bracket, t_2 is the marginal tax rate of the second bracket, and $\hat{Y}$ is the threshold between the first income bracket and the second income bracket. I still use the CES utility function (2.15) as the utility functional form. Then (2.20) becomes:

(2.21) $\{\alpha[b+(1-t_1)\min(w_ih_i, \hat{Y})+(1-t_2)\max(w_ih_i-\hat{Y}, 0)]^{(\sigma-1)/\sigma}+(1-\alpha)(1-h_i)^{(\sigma-1)/\sigma}\}^{\sigma/(\sigma-1)}$

Again, because all individuals are *ex ante* identical, and their wages are independent of one another, subscript i can be ignored in (2.20) and (2.21). The government now has four policy tools instead of two: one government transfer, one income threshold and two marginal income tax rates. Therefore, under Bentham's additive SWF, the government chooses t_1, t_2, b, and $\hat{Y}$ to maximize:

(2.22) $\int_{w*}^{\infty} U\{b+(1-t_1)\min[wh(w),\hat{Y}]+(1-t_2)\max[wh(w)-\hat{Y},0],1-h(w)\}f_w(w)dw$

$+ F_w(w^*)U[b,1]$

subject to the balanced budget constraint:

(2.23) $\int_{w*}^{\infty}\{t_1 \min[wh(w), \hat{Y}] + t_2 \max[wh(w)-\hat{Y}, 0]\}\, f_w(w)\, dw = b$

where $f_w(w)$ is the p.d.f. of ability, $F_w(w)$ is the C.D.F of ability, and w^* is the labor market participation condition that fulfills:

(2.24) $U\,\{b+(1-t_1)\min[w^*h(w^*), \hat{Y}]+(1-t_2)\max[w^*h(w^*)-\hat{Y}, 0], 1-h(w^*)\} = U(b,1)$

If and only if $w_i < w^*$, individual i stays outside the labor market. Under the multiplicative Nash SWF, the government chooses t_1, t_2, b, and $\hat{Y}$ to maximize:

(2.25) $\int_{w*}^{\infty}\log\{U[b+(1-t_1)\min(wh(w), \hat{Y})+(1-t_2)\max(wh(w)-\hat{Y}, 0),1-h(w)]\}f_w(w)dw$

$+ F_w(w^*)\log[U(b,1)]$

subject to (2.23).

Since both the individual's and the government's problem are highly non-differentiable, I follow Slemrod, *et al* (1994) by using approximating methods to simulate the relationship between the optimal two-bracket linear income tax and earning inequality. I draw 2000 points from each lognormal distribution of ability used in this section to represent a wage distribution. Each point accounts for a 0.0005 increase in the cumulative frequency of the wage. The lowest cumulative frequency is 0.0005 while the highest is 0.9995. Without losing generality, I assume only 2000 individuals live in the economy, and each is exclusively assigned a wage from the 2000 wages drawn. In the approximation, the individual's problem does not change at all. Each individuals still maximizes (2.20) by choosing labor supply h. However, the government's problem changes a little bit. Under the Bentham SWF, the government now chooses t_1, t_2, b, and $\hat{Y}$ to maximize:

(2.26) $\Sigma_i\, U_i\, \{b+(1-t_1)\min[w_i h_i(w_i), \hat{Y}]+(1-t_2)\max[w_i h_i(w_i)-\hat{Y}, 0], 1-h(w_i)\}$

subject to the balanced government budget constraint:

(2.27) $\Sigma_i\, \{t_1 \min[w_i h(w_i), \hat{Y}]+ t_2 \max[w_i h(w_i)-\hat{Y}, 0]\} = 2000\, b$

where i ranges from 1 to 2000, and w_i is the wage of individual i. Under the Nash SWF, the government chooses t_1, t_2, b, and $\hat{Y}$ to maximize:

(2.28) $\Sigma_i \log \{U_i\, [b+(1-t_1)\min(w_i h_i(w_i), \hat{Y})+(1-t_2)\max(w_i h_i(w_i)-\hat{Y}, 0), 1-h(w_i)]\}$

subject to (2.27). Actually, the government has only three free choices from the four tools, because the fourth tool can be solved out by the balanced government constraint (2.27). In my simulations, government transfer b is solved out, leaving t_1, t_2, and $\hat{Y}$ as the chosen variables.

2.4.2 Simulation Results with a Relatively Small σ

In this section, I simulate a case where the elasticity of substitution between consumption and leisure (σ) is set to as small as 0.4.[8] This value is from Stern (1976) and followed by Slemrod, *et al* (1994). In this simulation, α is set to 0.6136

[8] I change σ around 0.4 from 0.3 to 0.7 by 0.1 to check the sensitivity of my simulation. Figure A.1 in Appendix A shows results when σ is set to 0.3, 0.5, 0.6 and 0.7. These alternatives do not change my conclusion at all.

following Stern (1976) as in the one-bracket cases.[9] Table 2.2 shows the key elasticities for labor supply of the mean person when she faces a two-bracket income tax. Figure 2.5 shows the effect of earning inequality on optimal tax rates under both the Bentham SWF and the Nash SWF.

Table 2.2: Key Elasticities for Labor Supply of the Mean Person[10]

(evaluated at $t_1 = 0.234$, $t_2 = 0.200$, $b = 0.059$, $\hat{Y} = 0.315$, and $\alpha = 0.6136$)

	$\sigma = 0.4$	$\sigma = 1.0$
Uncompensated Labor Supply Elasticity	-0.139	0.140
Compensated Labor Supply Elasticity	0.236	0.666
Income Elasticity	-0.375	-0.526

When the wage standard deviation is relatively small, I find the result of Slemrod *et al* (1994) that the optimal lower bracket rate is greater than the upper bracket rate.[11] When the standard deviation is relatively large, however, then I find the opposite result.[12] This "switchover point" appears when the standard deviation is somewhere between 0.3109 and 0.3609, for both SWFs, roughly twice as big as the 0.1609 used by Slemrod, *et al* (1994) as their only earning inequality level. The coefficient of variation at the

[9]Slemrod *et al* (1994) use varied α in their simulations. Particularly, they use 0.41 in their $\sigma = 0.4$ case, which put a less-than-half weight to consumption.

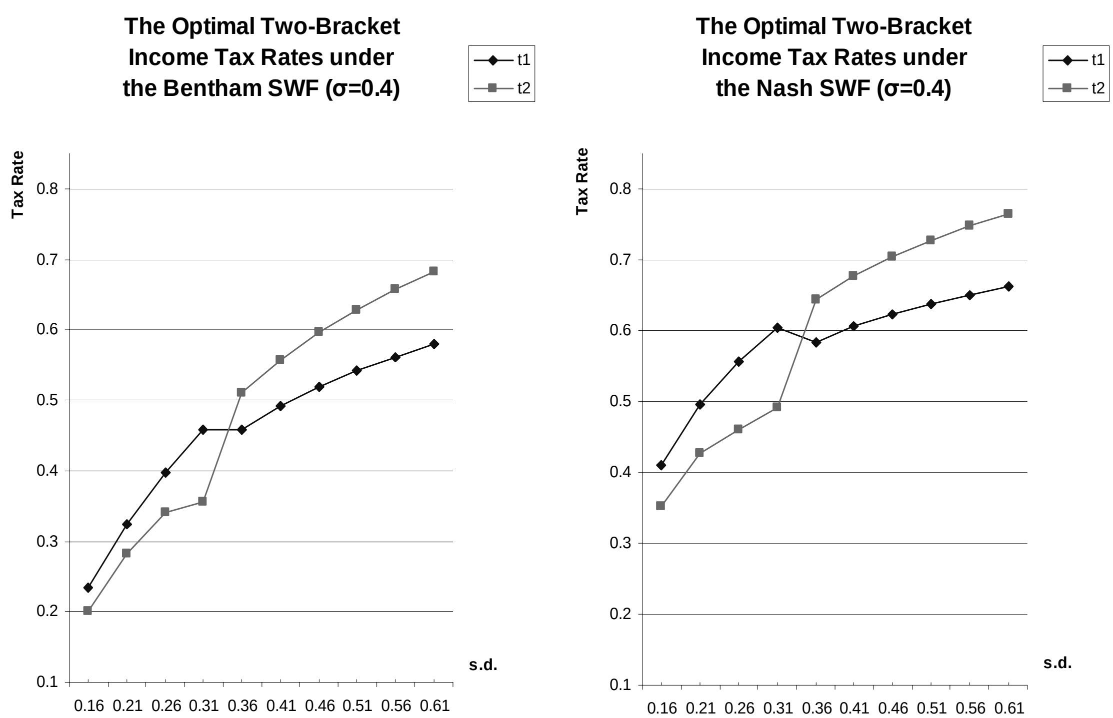
[10]The income elasticity is calculated by $(\partial h/\partial b)\{[b+(1-t_1)\min(wh, \hat{Y})+(1-t_2)\max(wh-\hat{Y}, 0)]/h\}$. The uncompensated labor supply elasticity is not zero when σ is set to 1.0 (Cobb-Douglas utility) is still because of non-labor income (the government transfer). The compensated labor supply elasticity is calculated by Slutsky equation: compensated elasticity = uncompensated elasticity – income elasticity

[11]In the case where $\sigma = 0.1609$ and the SWF is the Bentham SWF, Slemrod *et al* (1994) find that t_1, t_2, b, and $\hat{Y}$ equal 0.234, 0.202, 0.058 and 0.300 respectively. I find that they are 0.234, 0.200, 0.059 and 0.315. Those two groups of values differ from each other slightly because Slemrod *et al* use $\alpha = 0.41$ and I use $\alpha = 0.6136$ following Stern (1976).

[12]In all ten cases including σ is set to 0.3, 0.4, 0.5, 0.6 and 0.7, I find this switchover appearing between 0.2109 and 0.4109. Moreover, six out of ten times, the switchover appears between 0.3109 and 0.3609.

switchover point is between 0.783 and 0.909 (for both SWFs).[13] In contrast, because Slemrod *et al* (1994) use only s.d. = 0.1609, they find that the optimal lower bracket rate is always greater than the upper bracket one. Surprisingly, when I allow for greater possible wage inequality, I show that their result does not always hold.

Figure 2.5: The Optimal Two-Bracket Income Tax Rate (σ = 0.4).

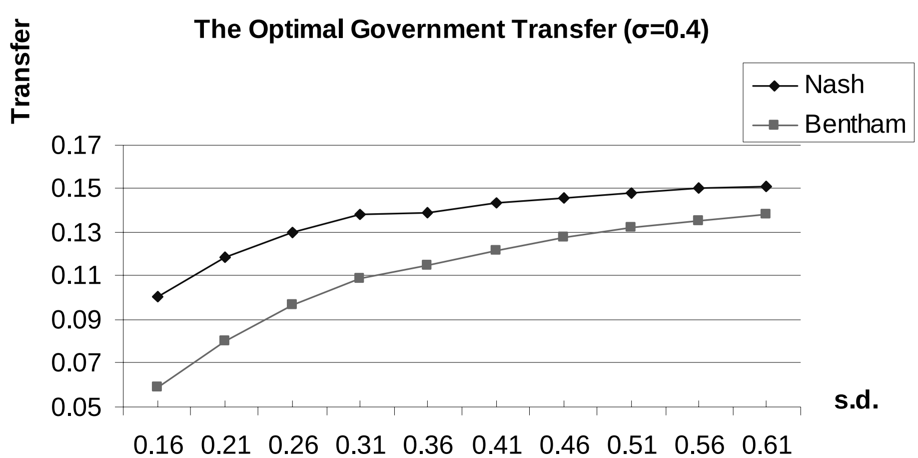


Furthermore, under both SWFs, the optimal upper bracket rate (t_2) is always increasing with the wage spread (for my parameters), whereas the optimal lower bracket rate (t_1) is increasing overall but not around the switchover point. For example, when earning inequality changes from a mild level where the wage spread is 0.1609 used

[13]The highest three coefficients of variation of the wage rate of the U. S. between 1979 and 2004 are 0.888 (1993), 0.802 (2004), and 0.793 (1992), values that are in this interval. The s.d. = 0.1609 used by Slemrod et al (1994) yields a coefficient of variation equal to 0.405, which is outside the range of 0.590 to 0.888 witnessed in the U. S. from 1979 to 2004.

by Mirrless (1971), Stern (1976) and Slemrod *et al* (1994) to the extreme level where the spread 0.6109, the optimal upper bracket rate increases monotonically from 0.200 to 0.682 under the Bentham SWF (from 0.351 to 0.764 under the Nash SWF). Surprisingly, the optimal lower bracket rate does not increase monotonically. Overall, it increases from 0.234 to 0.579 under the Bentham SWF (from 0.410 to 0.662 under the Nash SWF). In the Nash case, the lower bracket rate falls a bit when the standard deviation rises from 0.3109 to 0.3609 (where the rate decreases slightly from 0.604 to 0.583). Though the lower bracket rate does not have a setback in the Bentham case when σ = 0.4, it is quite flat in the switching area, changing from 0.45788 to 0.45849 (and it does fall near the switchover point when σ is set to 0.3 or 0.5).

Figure 2.6: The Optimal Government Transfer (σ = 0.4, two-bracket).

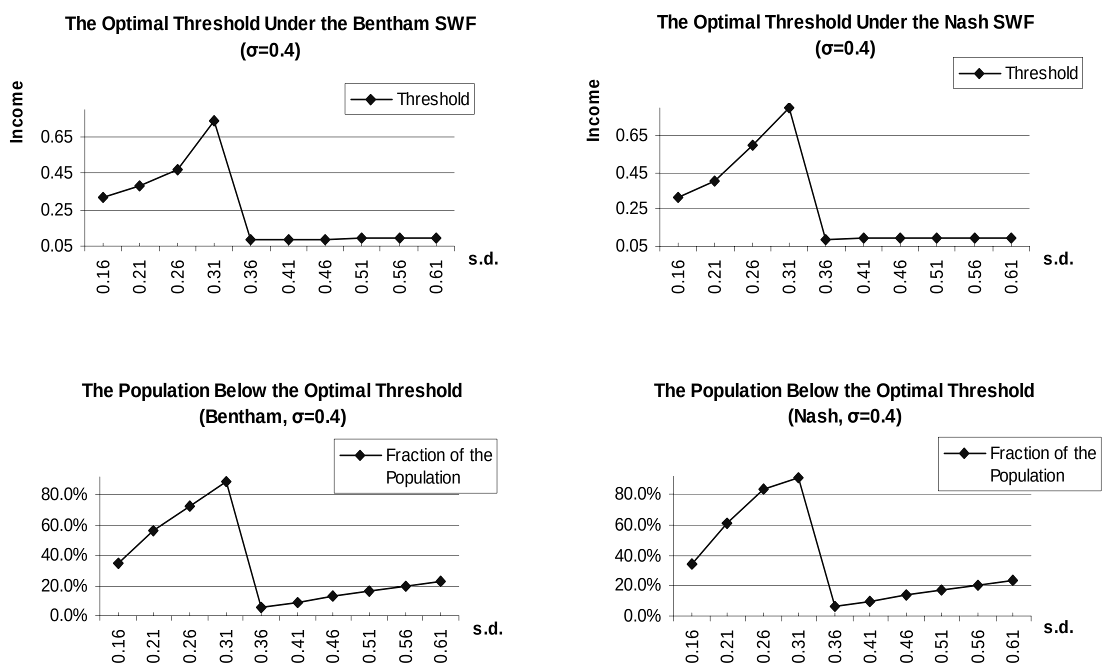


Regarding the optimal government transfer, as shown in Figure 2.6, it is strictly increasing with the wage spread under both SWFs. [14]. When the wage standard deviation changes from 0.1609 to 0.6109, the optimal transfer grows from 0.059 (roughly 14.9% of the mean wage) to 0.138 (34.8%) under the Bentham SWF (from 0.101 (25.4%)

[14]Please also see simulation results for other values of σ in Figure A.2 of Appendix A.

to 0.151 (38.0%) under the Nash SWF). As shown in Figure 2.7, the optimal income threshold $(\hat{Y})$ that divides the two brackets does not show a monotone property. Moreover, both of the optimal rates $(t_1$ and $t_2)$ under the Bentham SWF are larger than under the Nash SWF, while the optimal transfer (b) under the Bentham SWF is less than under the Nash SWF, a result that is similar to the one-bracket case. It is still because the Nash SWF puts more weight on the poor.

Figure 2.7: The Optimal Threshold and the Population below the Threshold (σ = 0.4)



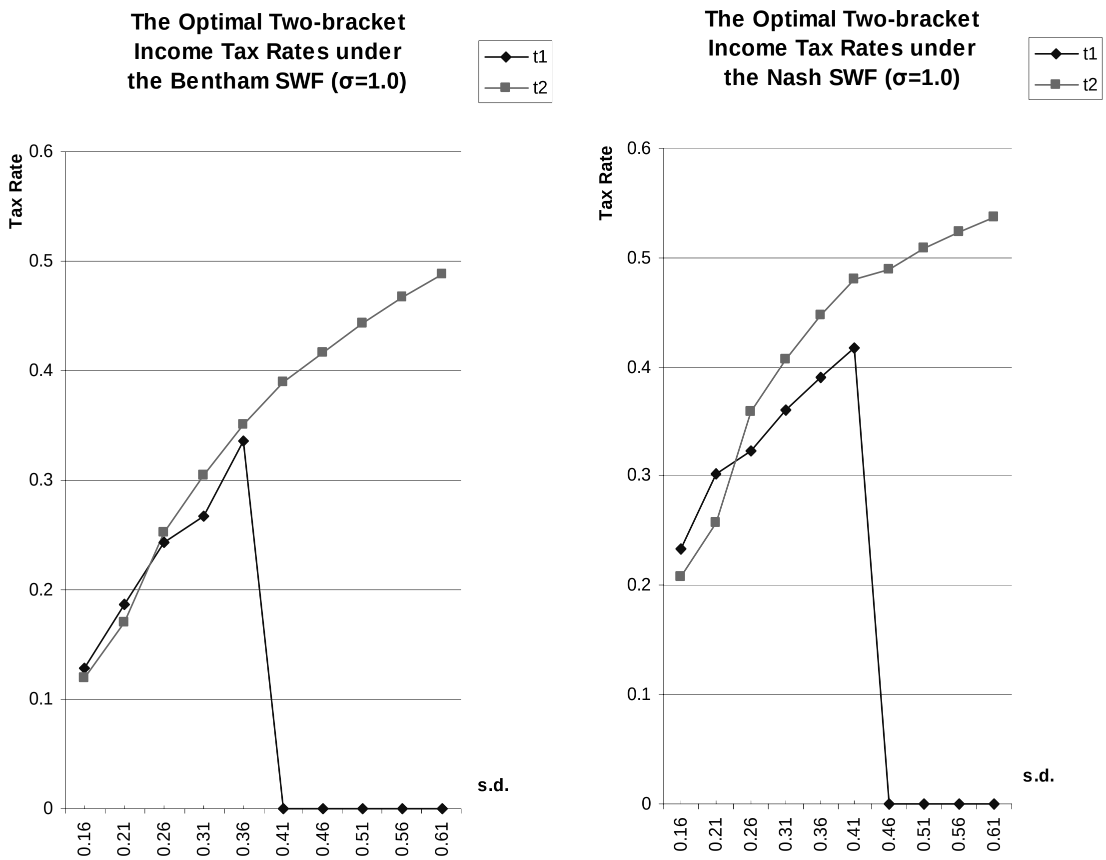
Several reasons together explain my conclusions regarding an increase in wage spread. First, the reason for the overall increases of both optimal rates is that the government has to increase both rates to collect necessary revenue to help support the poor when earning inequality become more serious. This is comparable to the one-bracket case. Second, when earning inequality is relatively mild, the population of the "middle class" that pays positive net taxes but whose income is still less than or equal the threshold $\hat{Y}$ is relatively large. For example, the optimal threshold, as shown in Figure 7, can be as high as 0.737 under the Bentham SWF (0.796 under the Nash SWF), which are almost twice as big as the mean wage, 0.3969. The result is that 91.2% of the population does not pay the upper bracket tax under the Bentham SWF (89.3% under the Nash SWF). The large population of the middle class means that the taxable income of this class is also large. Thus the government is able to raise a substantial amount of revenue from the middle class in the first bracket (facing t_1). As a result, the government does not need to raise substantial revenue from the rich who earn more than the threshold. Without losing revenue, the government can implement smaller upper bracket rates, to encourage labor supply of the rich, those who are the most productive workers in the economy. However, as earning inequality rises to a high level, the middle class shrinks rapidly. As shown in Figure 7, the optimal threshold rises gradually and then decreases dramatically from 0.737 to 0.082 under the Bentham SWF (from 0.796 to 0.088 under the Nash SWF). The population below the threshold cut from 91.2% to 5.0% under the Bentham SWF (from 89.3% to 6.0% under the Nash SWF). The majority of the tax base is then shifted from the lower bracket to the top bracket. Finally, since the government is then treating the rich as the major target, it is able to give those with low income a smaller lower bracket rate that encourages them to work and improve their welfare and therefore total social welfare as well.

2.4.3 Simulation Results with a Relatively Large σ

In addition to the simulation with a small σ, we repeat the simulation approach, but changing σ from 0.4 used by Stern (1976) to 1.0 used by Mirrless

(1971).[15] Again, as shown in Table 2.2, the increase of σ means the increase of the uncompensated labor supply elasticity.

Figure 2.8: The Optimal Two-Bracket Income Tax Rate (σ = 1.0)

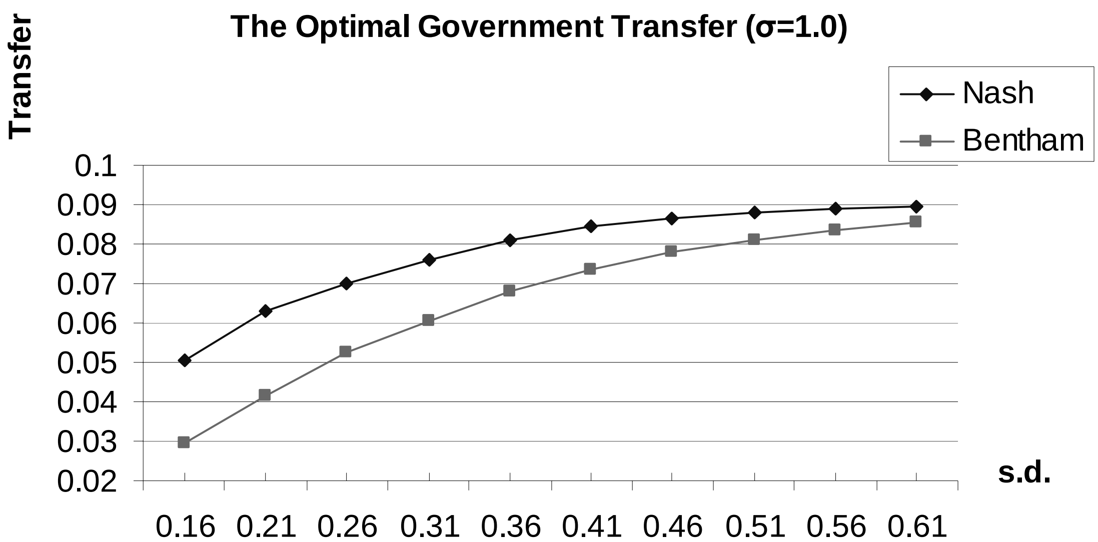


As shown in Figure 2.8 and 2.9, I find that the optimal upper bracket rate and government transfer are increasing with the wage spread under both SWFs. The optimal lower bracket rate, however, is not monotonic. Figure 2.10 shows the optimal threshold is not monotonic either. In addition, I still find a switchover point where t_2 rises above t_1 under both SWFs (when the standard deviation is between 0.2109 and

[15] Other values of σ, such as 0.9 and 1.1 that are around 1.0, are also simulated to check the sensitivity of my simulation. To set σ to be 0.9 or 1.1 does not change my conclusion at all. Please see the simulation results in Table B.1 and Table B.2 in Appendix B.

0.2609 as shown in Figure 2.8).[16] The coefficient of variation of the switchover point is between 0.531 and 0.657.[17] Before the switch, the optimal lower bracket rate is greater than the optimal upper bracket one. After the switch, the upper bracket rate is higher.

Figure 2.9: The Optimal Government Transfer (σ = 1.0, two-bracket)

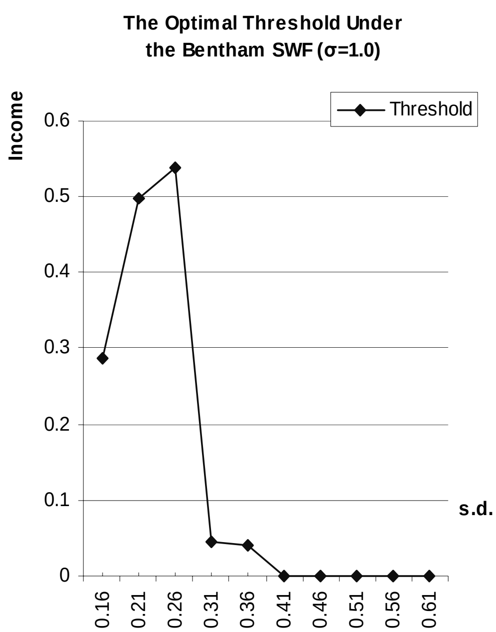


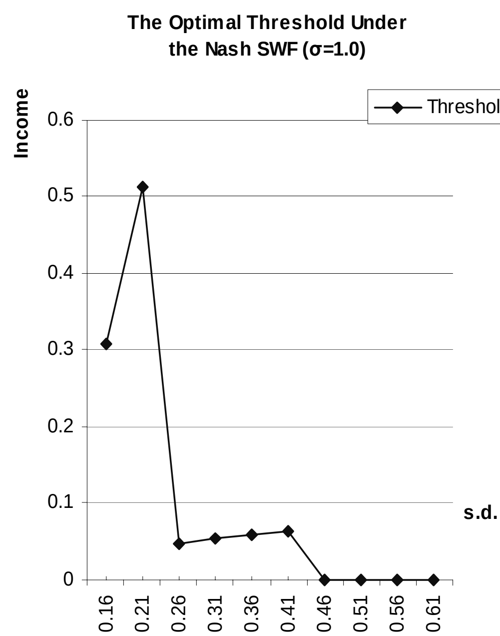
Furthermore, I find that the two-bracket tax structure converges completely to the one-bracket case under both SWFs when earning inequality becomes quite serious.[18] As shown in Figure 2.8, the optimal upper bracket rate is always greater than zero and is increasing with the wage spread under both SWFs, whereas the optimal lower bracket rate stays positive only before the point where s.d. = 0.4109 and c.v. = 1.035 under the Bentham SWF (s.d. = 0.4609 and c.v. 1.161 under the Nash SWF).[19] It then drops to

[16]In all six cases including σ = 0.9, 1.0, and 1.1, I find this switchover appearing between 0.2109 and 0.3609. Please see these switchovers in Table B.1 and B.2 of Appendix B.
[17]Most coefficient of variations of wage of the U.S. ranging from 0.590 to 0.888 from 1979 to 2004 are included in this interval, while 0.405 generated by s.d. = 0.1609 is still not included.
[18]I find this in all six cases where σ = 0.9, 1.0, and 1.1. Please see Table B.1 and B.2 of Appendix B.
[19]The coefficients of variation of Mexico from 1995 to 1999 are all larger then the 1.035 of the Bentham case (larger then the 1.161 of the Nash case also). Given that most developing countries have similar c.v.

zero after that. In addition, when the lower bracket rate drops to zero, the threshold drops to zero also. This means that the two-bracket structure converges to one-bracket.[20]

Figure 2.10: The Optimal Threshold (σ = 1.0)





This convergence is actually a special case of the switchover, in which the optimal lower bracket rate falls to as small as zero, and the threshold drops to zero also. First, as explained in the previous section, when earning inequality is quite serious, the optimal lower bracket rate is smaller than the optimal upper bracket rate to help the middle class. Second, workers of the middle class are very elastic because large σ means a large uncompensated labor supply elasticity. This forces the government to use an even smaller lower bracket rate to keep the middle class working and to prevent them

with that of Mexico, and most developed countries have similar c.v. with that of the U. S., a one-bracket income tax could be more suitable for developing countries than developed countries.

[20]Moreover, the two-bracket structure converges earlier as σ becomes larger from 0.9 to 1.1 as shown in Table B.1 and Table B.2 of Appendix B.

from becoming net welfare recipients. Last, the pressure of raising revenue drives the government to decrease the threshold, and to enlarge the population in the higher bracket, to collect more taxes that can be used to finance government transfers. All these causes interacting together imply that the middle class disappears while the two-bracket structure of taxation becomes one bracket.

2.5 CONCLUSIONS

Consistent with the results of Mirrless (1971), Stern (1976), and Cooter and Helpman (1974), my simulations generally favor the conclusion for the one-bracket case that both the optimal income tax rate and the government transfer increase when earnings become more unequally distributed. Moreover, I go on to show that the tax rate and transfer are strictly increasing with the wage spread. This conclusion does not depend on whether a relatively small or large elasticity of substitution between consumption and leisure is used in the simulation. A larger value of the elasticity changes only the magnitude but not the trend.

In the two-bracket case, I similarly find that the optimal upper bracket rate and government transfer are also always increasing with the wage spread. When the substitution elasticity is relatively small, the optimal lower bracket rate is increasing with wage disparity overall, but not in the area near the switchover point. It is not monotonic when the elasticity is large. I confirm results of Slemrod *et al* (1994) for a relatively low wage disparity that the upper bracket rate is less than the lower bracket rate. With a wage spread close to that of the U. S. in recent years, however, the result is reversed. Beyond this, I also find an interesting phenomenon. With a relatively large elasticity of substitution between consumption and leisure in the individual’s utility function, the optimal two-bracket income tax structure converges to the one-bracket case when earning inequality becomes serious. Though this can be treated as a special case of the switchover, it is still surprising that the lower bracket rate and the income threshold can be as low as zero. Furthermore, this theoretically simulated result may indicate that developing countries with serious income inequality may need to implement the one-bracket income tax structure instead of the multiple-bracket structure.

Chapter 3: The Distributional Effects of a City's Environmental Tax

3.1. INTRODUCTION

A particular environmental tax may have distributional effects across different factor owners, such as workers and landowners. One category of tax incidence research focuses on the static comparative analysis of a particular tax. These general equilibrium models feature highly abstract analyses and abundant results providing useful intuition.[21] Compared to the other taxes, environmental taxes are less studied in the tax incidence literature. Partially, it is because an environmental tax could be treated as an example of a particular tax form that already has been examined adequately by tax incidence research. Yet the environmental tax is introduced to cure the distortion of an economy rather than to be a distorting factor of the economy like other taxes. So, the distributional effect of an environmental tax is of interest itself. Furthermore, when spatial characteristics are considered, environmental quality can be treated as both a private good and a public good. Private features of environmental quality come from individual choice of residential location, while public features come from the fact that the area's total pollution level eventually influences everybody. This makes the distributional effect of an environmental tax more interesting.

Henderson (1977) investigates the distributional effect of a pollution tax in a system of cities, and he focuses on population migrations. Baumol and Oates (1988) also review the distributional effects of an environmental tax, using intuitive analysis. Rapanos (1992, 1995) and Fullerton and Heutel (2004a, 2004b) investigate environmental tax incidence using the framework of Harberger (1962) without spatial characteristics. Frankel (1987) studies a pollution tax and spatial amenity together in a

[21]Among those making significant contributions studying this area are Harberger (1962), Mieszkowski (1972), Pechman and Okner (1974), Henderson (1977), and Fullerton and Rogers (1993).

closed city. He finds that landlords may bear negative, partial or no burden of the pollution tax. However, his finding is based on particular functional forms for utility and locational amenity.

In this paper, I examine the distributional effects of an environmental tax on the price system in a spatial model of a city where living location is a choice of the residents. Unlike other environmental tax researchers just mentioned, I use a general equilibrium model of urban economics with the quality of the environment included in residents' preferences. Thus environmental quality is ultimately a choice, as when individuals decide where to live in the city. "Social welfare" is defined over the many identical residents, and it is affected both by the quality of the environment and by the rent they pay to absentee landlords. In contrast, other researchers treat the environment as a pure Samuelsonian public good that affects all residents equally. Also, these other urban economics papers focus on problems such as property tax incidence, city size, urban congestion and population migration. In contrast, my paper examines the incidence of an environmental tax, this is, the changes in factor and output prices.

In a general equilibrium analysis, I find that an environmental tax in a closed city targeting polluting emissions can reduce the pollution emissions in the city's central business district (CBD), and therefore improves environmental quality of the city everywhere. This is because the tax increases the price of pollution emissions (as a factor input) so that less pollution emissions are used in the production of the composite good. In addition to the reducing of pollution, the tax reduces the wage received by labor. This is because firms try to shift the burden of the environmental tax onto labor, which is inelasticly supplied. Lower pollution as an input means a lower marginal product of labor. Overall, as the environmental tax rate rises from quite low to quite high, the tax first improves the social welfare of the city until it hits the optimal level, and then it reduces welfare as the marginal benefit falls below the marginal cost of the tax. Furthermore, for a fixed number of residents, the tax first makes the city boundary shrink but then makes it grow as residents try to balance between better housing and shorter commuting. Initial increases in the tax pull down the rent for any particular location in the city and further increases pull the rent back up. The lower bids of residents that result from lower income are offset by the higher bids that result from the pursuit of better environmental quality.

The environmental tax has two possible effects on the whole rent gradient because it does not affecting the rent of every location uniformly. In the fist case, when the tax increases from small to large, all rents first decrease, forming a downward shift of the rent gradient, then the rents close to the CBD turn back to increase while those away from the CBD still decrease, forming a clockwise pivoting, and at last, all rents increase, forming an upward shift. In the second case, all rents first decrease, forming a downward shift of the rent gradient, then the rents away from the CBD turn back to increase while those close to the CBD still decrease, forming a counter-clockwise pivoting, and at last, all rents increase, forming an upward shift.

In a quite polluted city with an inadequate environmental tax, the increase of the tax shifts down the rent gradient, decreasing the rent of land everywhere in the city. In contrast, for a very clean city with a high environmental tax, an increase of the tax shifts up the rent gradient, increasing the rent of land everywhere. In one type of theoretically defined city, as the environmental tax increases from quite low to quite high, it first improves the social welfare of the city, shrinks the size of the city , and cuts the rent of any particular location in the city; and then it reduces the social welfare, expands the city, and raises the rent anywhere. In another type of theoretically defined city, furthermore, a Pareto improvement is possible, which means that the tax may increase both the social welfare of the city residents and the gradient of rent received by landlords.

The paper proceeds as follows. Part 2 reviews both the tax incidence literature and the rent gradient theory in urban economics; Part 3 presents and solves the model; Part 4 provides the comparative static analysis in the closed city case; finally, Part 5 discusses conclusions and possible problems.

3.2. LITERATURE REVIEW

The fundamental importance of tax incidence study is that it reveals the difference between statutory incidence and the economic tax burden distributed across different groups of individuals, regions, industries, and owners of the factors of production. The difference may arise if the tax causes changes in the equilibrium prices of factors (the sources side) or products that people buy (the uses side). Arnold Harberger (1962) provides a framework to analyze tax incidence using a general equilibrium model. Later on, Pechman and Okner (1974) and Fullerton and Rogers (1993) extend tax incidence analysis by capturing income heterogeneity among individuals. The former authors use annual income, while the latter authors use lifetime income. Bull, Hassett and Metcalf (1994) extend tax incidence analysis to find distributional effects across jurisdictions.

With respect to environmental tax incidence, Baumol and Oates (1988) have addressed the distribution of benefits and costs of environmental taxes in their book. They illustrate two opposite academic attitudes towards environmental quality or pollution: a pure public good opinion following Paul Samuelson (1954) and a pure private good view described by Charles Tiebout (1956). Baumol and Oates use general equilibrium ideals to explain the different attitudes towards environment quality and the corresponding tax incidence analysis, and they focus on the distributional effects across different income groups on the uses side. The effect across different factor owners on the sources side is addressed by Rapanos (1992, 1995), and Fullerton and Heutel (2004a, 2004b), who investigate environmental tax incidence through a classical Harberger framework. They all use general equilibrium models without spatial characteristics.

Most economists view environmental quality a public good as mentioned above. When thinking about all individuals living in a suburban or rural area, however, location choice can readily affect the consumption of environmental quality. Urban economists have well-developed general equilibrium models to deal with location choice, and this research vehicle is widely used in property tax incidence studies. The original framework with location in a general equilibrium model is developed by Von Thunen (1826), Isard (1956), Beckmann (1957), Muth (1961), and Alonso (1964). Their framework is the foundation of the model used in this paper.

William Alonso (1964) presents a general equilibrium urban model with location as an individual choice. To solve the urban model more easily, Stull (1973, 1974) and Wheaton (1974) develop a new method different from that of Alonso. To do comparative static analysis, Alonso's method is also extended later by Brown (1985). In general, in these models, mobility within the city means that utility is equalized, so those living further away from the CBD with higher commuting costs must be compensated, in equilibrium, by lower rental costs. Stull (1974) also introduces an externality of a city's size and points out that a positive rent gradient is possible. Henderson (1977) investigates the distributional effect of a pollution tax in a system of cities. He emphasizes population migration among cities and does not discuss the distributional effect on different production factor owners. Hockman (1978) introduces a dispersion function of pollution into the utility function in the urban model.[22] Using such a model, he analyzes the shape of the land rent gradient and the optimal pollution tax, but he does not pay attention to tax incidence. Frankel (1985) investigates the relationship between property values and amenity changes in a closed city. He includes environmental quality in his model and shows how dis-amenity affects the land rent gradient. However, he does not mention tax incidence either.

Later, Frankel (1987) investigates a tax and spatial amenity together in a closed city while assuming no commuting cost, Cobb-Douglas utility functions, and particular amenity functional forms. He integrates an amenity that is a function of distance and pollution into residents' utility function and examines the relationships between the positive land rent gradient and two forms of tax, an income tax and an excise tax. He finds that landlords may bear negative, partial or no burden of an excise tax. The incidence is decided by a parameter of the Cobb-Douglas utility function. He mentions that under some particular conditions, his conclusion on the effect of an excise tax on the land rent gradient also applies to a pollution tax, but he does not provide detail.

3.3. THE BASIC MODEL

The model has N identical resident renters who are owners of all labor and a fixed number of identical landlords who are owners of all residential land. The residents work at a central business district (CBD) and live in a mono-centric city that spreads over a featureless horizontal straight line with a fixed width and a variable length from the CBD at point zero to a boundary at point b. The landlords live elsewhere. The city is "open" if residents can move in or out of it, while it is "closed" otherwise.

[22] If E is the pollution emissions at the CBD, and D is the distance from the CBD, then the ambient environmental quality at location D is represented by $Q(D, E)$, where $\partial Q/\partial D > 0$ and $\partial Q/\partial E < 0$.

Thus, the population of residents can change if the city is open, but not if the city is closed. The line of the city provides land space to users for any purpose. Without loss of generality, I assume the width of the line is one, so that each point on the line provides a fixed amount of land that is normalized to be one unit of land. All residents live to the right of the CBD point, so that one residential area is formed from point zero to the boundary b. Land in the residential area is used only for residential purposes. In addition, the city's border can vary according to the city's population and the demand of its residents for land. The area outside the city is an agricultural area, which produces a fixed output equivalent to R^A dollars per unit of land.

Without loss of generality, I assume the CBD is only a small point at the left end of the straight line and does not occupy any positive square unit of land. Competitive firms located in the CBD use a constant return to scale (CRTS) production technology to produce a composite good X, produced using labor L and pollution emissions E as the inputs.[23] They sell the composite good directly to consumers without transportation cost. Obviously, this composite good is produced using a dirty process with variable polluting emissions. In order to control pollution, the local government levies an environmental tax on emissions at a rate of t dollars per unit, as suggested by Pigou (1932). The tax revenue is used to finance a per-capita government transfer g to every resident. Mathematically, the production function of firms is:

(3.1) $X = (L, E)$

where $X_L > 0$, $X_E > 0$ and $X_{LL} < 0$, $X_{EE} < 0$, and X_{LE} ($\equiv X_{EL}$) > 0. With perfect competition and the CRTS production technology, firms earn zero profit:

(3.2) $X = \omega L + t E$

[23]Environmental economists are accustomed to treat pollution as an input to production rather than output from it. Almost all processes of modern production are dirty, so pollution has been necessary to produce anything, as essential as any other factor such as land, labor or capital.

where ω is the wage rate, and the price of the composite good is normalized to one.

The pollutant spreads from the CBD at a declining rate across the city. It does not influence behavior of the firms, but it damages environmental quality of the residential area. These damages diminish with distance. Thus, as in Hockman (1978), the environmental quality of a particular location can be expressed by a function $Q(D, E)$, which has $Q_D \equiv \partial Q/\partial D > 0$ and $Q_E \equiv \partial Q/\partial E < 0$. So, living in such a city, a resident has a utility function:

(3.3) $U = U\,[C, H, Q(D, E)]$

where C is her consumption of the composite good, H is her consumption of land space, and $Q(D, E)$ is the environmental quality of her chosen location.[24] The first and second derivatives of the utility function are $U_C > 0$, $U_H > 0$, and $U_Q > 0$ and $U_{CC} < 0$, $U_{HH} < 0$, and $U_{QQ} < 0$. Each resident provides one unit of labor in the CBD inelastically and earns the wage ω. The resident commutes to the CBD daily to work from her living location that is D miles from the CBD. So, she bears a daily commuting cost $T(D)$, where $dT/dD \equiv T_D > 0$. The road system in this city is extremely well developed and has no traffic congestion at all. Thus, the budget constraint of the resident is:

(3.4a) $C + R(D)\,H + T(D) = Y$

(3.4b) $Y \equiv \omega + g$

where $R(D)$ is the rent of unit land at location D, and Y is income of the resident, which consists of wage and government transfer. As a rational individual, the resident chooses C, H, and D to maximize her utility function (3.3) subject to her budget constraint (3.4a). When maximizing her utility, she treats Y ($\equiv \omega + g$) and E as given. Particularly, she also treats $R(D)$ as given in the sense that

[24]Following Alanso (1964), a land space represents "housing" and directly enters individuals' utility function.

though she is able to choose her living location D, she cannot affect the rent there, which is given by $R(D)$. The consumer's optimization yields choices for C and H that depend on the choice of D, so these can be written as $C(D)$ and $H(D)$.[25]

Though landlords do not live in this city and also do not commute, they use their rental income to buy the composite good for their consumption:[26]

$$(3.5)\quad c = \int_0^b R(D)\, H(D)\, dD$$

where c is the consumption of the composite good of all landlords, and the right hand side (RHS) of (3.5) is the total rental income of landlords owning city land.

Furthermore, the government budget constraint is:

$$(3.6)\quad N\, g = t\, E$$

The market clearing conditions of the composite good and labor are:

$$(3.7)\quad X = c + \int_0^b [C(D)+T(D)]\, dD$$

$$(3.8)\quad N = L$$

where $\int_0^b [C(D)+T(D)]\, dD$ is the consumption and transportation costs of all residents.

The land market determines the rent of land, which is also called the bid price of land in the literature.[27] To find the bid price of land, I need first to find the family of rent gradients, each element of which represents a rent gradient for certain utility level. I can solve for the family of rent gradients by following Stull (1973, 1974) and Wheaton

[25]Even though all residents have the same income and preference, they may live in different locations in the city. This means the effective income $[Y - T(D)]$ of each individuals differ. Therefore, the choices for C and H of each individual depend on her living location D.

[26]To keep the model simple, I assume landlords of the agriculture area don't participate into the economy of the city until their land is included in the city.

[27]According to Alonso (1964), the bid price of land is the highest offer of bidders when landlords auction their land to residents. It is equivalent to a mathematical problem that landlords maximizes the rent of each unit of their land subject to a constraint that ensures residents are able to keep a certain utility level with which they are satisfied.

(1974). First, I rearrange the consumer budget constraint (3.4a) at location D to get a rent function:

(3.9) $R(D) = [1/H(D)]\,[Y - C(D) - T(D)]$

Then, given distance D, environmental quality of the location $Q(D, E)$, and individual's income Y ($\equiv \omega + g$), residents and landlords interact in an auction that serves to maximize (3.9) with respect to $C(D)$ and $H(D)$ subject to a constraint of an exogenous utility level of residents:

(3.10) $V = U\,[C(D), H(D), Q(D, E)]$

where, V is the possible equilibrium utility level. By solving this problem (Appendix C), I get:

(3.11a) $C = C(D, Y, E, V)$

(3.11b) $H = H(D, Y, E, V)$

Substituting C and H into the rent function, I get the family of rent gradients (one for each V):

(3.12) $R = R(D, Y, E, V)$

Among the arguments of R, one is chosen by individuals (D), and the others are determined endogenously in equilibrium (Y, E, and V). Furthermore, after Y ($\equiv \omega + g$), E, and V are determined in the equilibrium, (12) will then be the equilibrium rent gradient, and (3.11a) and (3.11b) will be the equilibrium consumption of the composite good and land space of residents at each location, since C and H satisfy the equilibrium utility function (3.10) and the budget constraint (3.4a) simultaneously.

In equilibrium, the residents must be indifferent among various residential locations, willing to live anywhere in the city. Since the area outside the city is the agricultural area, the rent at the city limit b must equal the rent of agricultural land R^A, which is given exogenously. Thus:

(3.13) $R^A = R(b, Y, E, V)$

Moreover, the total land demanded by all residents must equal the total land supplied within the city limit by all landlords. This is equivalent to the condition that all residents must live within the city and exhaustively occupy all land space of the city:[28]

(3.14) $N = \int_0^b [1/H(D, Y, E, V)]\, d(D)$

where $1/H(D, Y, E, V)$ is the population density at location D.[29]

Since the production function (3.1) is CRTS and the market is perfect competitive, the equilibrium wage equals the marginal product of labor, and the environmental tax rate (the price of pollution emissions) equals the marginal product of pollution emissions. After substituting the labor market clearing condition (3.9) into the marginal products of labor and pollution emissions, I have:

(3.16a) $\omega = X_L (N, E)$

(3.16b) $t = X_E (N, E)$

From (3.16b), I can solve for pollution emissions E as a function of the environmental tax rate t and the population N:

(3.17) $E = X_E^{-1} (N, t) \equiv E(N, t)$

Substituting (3.17) into (3.16a), the wage is a function of t and N:

(3.18a) $\omega = X_L [N, E(N, t)]$

which can also be denoted as $\omega(N, t)$. By the government budget constraint (3.6), the per-capita transfer g can also be written as a function of t and N:

---

[28]This condition can also be expressed by:

(3.15) $b = \int_0^b [n(D)\, H(D, Y, E, V)]\, d(D)$

where $n(D)$ represents the number of people living at location D.

[29]Since each location provides one unit of land and the residents living at the same location choose the same amount of residential land space, the $1/H(D)$ is basically how many residents are living at location D. The integration of $1/H(D)$ from location zero to the border b is the number of residents living within the city.

(3.18b) $g = [t\, E(N, t)]/N$

which can be denoted as $g(N, t)$. Substituting (3.18a) and (3.18b) into the definition of the individual's income (4b), I have income as a function of t and N:

(3.19) $Y \equiv \omega + g = \omega(N, t) + g(N, t)$

which can be denoted as $Y(N, t)$. Substituting (3.17) and (3.19) into (3.13) and (3.14), I have:

(3.20) $R^A = R(b, Y(N, t), E(N, t), V)$

(3.21) $N = \int_0^b \{1/H[D, Y(N, t), E(N, t), V]\}\, d(D)$

If the city is *closed*, then V and b are endogenous, and N is fixed, so I can use any given environmental tax rate t to solve for V and b simultaneously by (3.20) and (3.21). If the city is *open*, then N and b are endogenous and V is fixed exogenously, and in this case I can solve for N and b simultaneously by (3.20) and (3.21).

3.4. THE INCIDENCE OF THE ENVIRONMENTAL TAX

3.4.1. Useful Derivatives and Differentiations

In order to find the incidence of the environmental tax t, I need to find the marginal effects of a positive change of t on the land rent at each point $R(D)$, the wage ω, the pollution emissions E, the population N, the city limit b, and the utility level V, which is also social welfare of the city This basically is a comparative static analysis of equation (3.17) - (3.21). To approach this objective, however, I need first to check the properties of endogenous variables such as $C(D, Y, E, V)$, $H(D, Y, E, V)$, $R(D, Y, E, V)$, $E(N, t)$, $\omega(N, t)$, $g(N, t)$, and $Y(N, t)$, which will all be useful later.

As shown in Appendix C, the first order partial derivatives of C and H with respect to their arguments can be obtained by totally differentiating the first order conditions of the maximization of the rent function (3.9) subject to the utility constraint (3.10), using m to denote the ration of marginal utilities ($m \equiv U_H/U_C$), assumed by Wheaton (1974) to have properties $\partial m/\partial C > 0$ and $\partial m/\partial H < 0$.[30]:

(3.22a) $\partial C(D)/\partial Y = -m/\{H\,[\partial m/\partial H - (\partial m/\partial C)\,m]\} > 0$

(3.22b) $\partial H(D)/\partial Y = 1/\{H\,[\partial m/\partial H - (\partial m/\partial C)\,m]\} < 0$

(3.22c) $\partial C(D)/\partial E = -U_Q Q_E\,/U_C - m\,[\partial H(D)/\partial V] >< 0$

(3.22d) $\partial H(D)/\partial E = [-U_Q Q_E\, m\,(\partial m/\partial C+1/H)]/[U_H(\partial m/\partial C-\partial m/\partial H)] > 0$

(3.22e) $\partial C(D)/\partial V = 1/U_C - m\,[\partial H(D)/\partial V] >< 0$

(3.22f) $\partial H(D)/\partial V = [m\,(\partial m/\partial C+1/H)]/[U_H(\partial m/\partial C-\partial m/\partial H)] > 0$

As shown in Appendix C, the first order partial derivatives of $R(D, Y, E, V)$ can be obtained using the envelope theorem when I maximize (3.9) subject to (3.10):

(3.23a) $\partial R/\partial D = -[1/H(D)]\, T_D - \mu(D)\, U_Q\, Q_D >< 0$

(3.23b) $\partial R/\partial Y = 1/H(D) > 0$

(3.23c) $\partial R/\partial E = -\mu(D)\, U_Q\, Q_E < 0$

(3.23d) $\partial R/\partial V = \mu(D) < 0$

where $\mu(D)$ is the Lagrange multiplier, which is the added rent for an exogenous change in utility and is less than zero, as shown in Appendix C.[31] Furthermore, I assume $\partial R/\partial D < 0$ for further investigations, which means that the increase of the environmental quality does not result in upward sloping rent gradients.[32]

---

[30]To have these properties, Wheaton assumes that utility is strictly quasi-concave, and that both the composite good and housing have positive income effects, which is consistent with my assumptions. As shown in Appendix C, I need to assume $\partial(U_H/U_C)/\partial Q = 0$ to obtain (3.22d). Two of these six properties, (3.22c) and (3.22d), are new to literature, while the others follow the results of Wheaton (1974) and Sasaki (1987).

[31]In an urban economy, given all other factors unchanged, more utility means less rent ($\mu < 0$).

[32]This assumption insures me to have a traditional decreasing rent gradient and stay outside the debate of whether or not the rent gradient should be positive.

As shown in Appendix D, the partial derivates of $E(N, t)$, $\omega(N, t)$, $g(N, t)$, and $Y(N, t)$ are:

(3.24a) $\partial E/\partial N = -X_{EL}/X_{EE} > 0$

(3.24b) $\partial E/\partial t = 1/X_{EE} < 0$

(3.24c) $\partial \omega/\partial N = [X_{LL}X_{EE} - (X_{LE})^2]/X_{EE} \leq 0$

(3.24d) $\partial \omega/\partial t = X_{LE}/X_{EE} < 0$

(3.24e) $\partial g/\partial N = -(t/N)/(X_{EL}/X_{EE}) - Et/(N^2) >< 0$

(3.24f) $\partial g/\partial t = (t/N)/X_{EE} + E/N >< 0$

(3.24g) $\partial Y/\partial N = \partial \omega/\partial N + \partial g/\partial N = [X_{LL}X_{EE}-(X_{LE})^2]/X_{EE}-(t/N)/(X_{EL}/X_{EE})-Et/(N^2) >< 0$

(3.24h) $\partial Y/\partial t = \partial \omega/\partial t + \partial g/\partial t = X_{LE}/X_{EE} + (t/N)/X_{EE} + E/N >< 0$

Particularly, (3.24h) implies that the Porter Hypothesis (Porter, 1995) might be true, since $\partial Y/\partial t$ can be no less than zero, which means the society is able to abate pollution emissions without bearing dollar costs. Though the Porter Hypothesis is still mathematically possible as indicated in (3.24h), I exclude the case of non-positive dollar costs of pollution abatement because the Hypothesis is currently considered unlikely among economists (absent some other additional market failures). So, $\partial Y/\partial t$ is assumed to be strictly less than zero ($\partial Y/\partial t < 0$).

In addition, I totally differentiate $E(N, t)$ and $\omega(N, t)$ and use equations (3.16)and (3.24):

(3.25) $dE = -(X_{EL}/X_{EE})\, dN + (1/X_{EE})\, dt$

(3.26) $d\omega = \{[X_{LL}X_{EE} - (X_{LE})^2]/X_{EE}\}\, dN + (X_{LE}/X_{EE})\, dt$

which show that dE and $d\omega$ are functions of dN and dt.

As shown in Appendix E, after totally differentiating (3.20) and (3.21), I have:

(3.27a) $-\dfrac{\partial R(b)}{\partial b}db - \dfrac{\partial R(b)}{\partial V}dV = [\dfrac{\partial R(b)}{\partial Y}\dfrac{\partial Y}{\partial N} + \dfrac{\partial R(b)}{\partial E}\dfrac{\partial E}{\partial N}]dN + [\dfrac{\partial R(b)}{\partial Y}\dfrac{\partial Y}{\partial t} + \dfrac{\partial R(b)}{\partial E}\dfrac{\partial E}{\partial t}]dt$

$$(3.27b)\ \frac{1}{H(b)}db - \left(\int_0^b -\frac{1}{H^2}\frac{\partial H}{\partial V}dD\right)dV = [1 + \left(\int_0^b \frac{1}{H^2}\frac{\partial H}{\partial Y}dD\right)\frac{\partial Y}{\partial N} + \left(\int_0^b \frac{1}{H^2}\frac{\partial H}{\partial E}dD\right)\frac{\partial E}{\partial N}]dN$$

$$+ [\left(\int_0^b \frac{1}{H^2}\frac{\partial H}{\partial Y}dD\right)\frac{\partial Y}{\partial t} + \left(\int_0^b \frac{1}{H^2}\frac{\partial H}{\partial E}dD\right)\frac{\partial E}{\partial t}]dt$$

where $R(b)$ and $H(b)$ are when $R(D)$ and $H(D)$ are evaluated at distance b. Equations in (3.27) show that db and dV are functions of dN and dt also. Furthermore, as shown in Appendix E, after totally differentiating (3.12) at a particular location D, I have:

$$(3.28)\quad dR(D) = \left(\frac{\partial R}{\partial Y}\frac{\partial Y}{\partial N} + \frac{\partial R}{\partial E}\frac{\partial E}{\partial N}\right)dN + \left(\frac{\partial R}{\partial Y}\frac{\partial Y}{\partial t} + \frac{\partial R}{\partial E}\frac{\partial E}{\partial t}\right)dt + \frac{\partial R}{\partial V}dV$$

where $R(D)$ is the rent at the particular location D. Equation (3.28) shows that $dR(D)$ is a function of dN, dt, and dV.

3.4.2. The Incidence in a Closed City

If the city is a closed one, the population cannot change, while the equilibrium utility or the social welfare level V is decided endogenously. Thus, $dN = 0$, and equation (3.25) - (3.28) can be rewritten as:

(3.29) $dE = (1/X_{EE})\ dt$

(3.30) $d\omega = (X_{LE}/X_{EE})\ dt$

$$(3.31a)\ -\frac{\partial R(b)}{\partial b}db - \frac{\partial R(b)}{\partial V}dV = [\frac{\partial R(b)}{\partial Y}\frac{\partial Y}{\partial t} + \frac{\partial R(b)}{\partial E}\frac{\partial E}{\partial t}]dt$$

$$(3.31b)\ \frac{1}{H(b)}db - \left(\int_0^b -\frac{1}{H^2}\frac{\partial H}{\partial V}dD\right)dV = [\left(\int_0^b \frac{1}{H^2}\frac{\partial H}{\partial Y}dD\right)\frac{\partial Y}{\partial t} + \left(\int_0^b \frac{1}{H^2}\frac{\partial H}{\partial E}dD\right)\frac{\partial E}{\partial t}]dt$$

(3.32) $dR(D) = [(\partial R/\partial Y)(\partial Y/\partial t) + (\partial R/\partial E)(\partial E/\partial t)]\ dt + (\partial R/\partial V)\ dV$

From (3.29) and (3.30), I have the effects of the environmental tax on the pollution emissions and the wage of labor:

(3.33) $dE/dt = 1/X_{EE} < 0$

(3.34) $d\omega/dt = X_{LE}/X_{EE} < 0$

since $X_{LE} > 0$ and $X_{EE} < 0$. Equation (3.33) implies that in a closed city, a positive change of the environmental tax does induce firms to abate their pollution emissions in the CBD so that the environmental quality of the city is improved everywhere. This is because the environmental tax raises the marginal cost (price) of pollution emissions as a factor input, and the firm substitutes toward its other input (labor). Less pollution means a lower marginal product of labor schedule. Since labor supply is fixed in a closed economy, the decrease of the marginal product of labor results in a lower market price of labor in (3.34).

As shown in Appendix E, I have the effects of the environmental tax on the equilibrium utility and the city limit derived from (3.31):

(3.35) $dV/dt = (\Psi/|A|)\,(\partial Y/\partial t) - (\Omega/|A|)\,(\partial E/\partial t) >< 0$

(3.36) $$\frac{db}{dt} = [-\frac{\partial R(b)/\partial Y}{\partial R(b)/\partial b} - (\frac{\Psi}{A})(\frac{\partial R(b)/\partial V}{\partial R(b)/\partial b})]\frac{\partial Y}{\partial t} + [-\frac{\partial R(b)/\partial E}{\partial R(b)/\partial b} + (\frac{\Omega}{A})(\frac{\partial R(b)/\partial V}{\partial R(b)/\partial b})]\frac{\partial E}{\partial t}$$

where $|A|$ is the determinant of the coefficient matrix of dV and db in equation (3.31), Ψ equals $-\{[\partial R(b)/\partial b][\int_0^b (1/H^2)\ (\partial H/\partial Y)\ dD] + [1/H(b)]\ [\partial R(b)/\partial Y]\}$, and Ω equals $\{[\partial R(b)/\partial b][\int_0^b (1/H^2)\ (\partial H/\partial E)\ dD] + [1/H(b)]\ [\partial R(b)/\partial E]\}$. Moreover, $|A|$, Ψ, and Ω are all negative, as shown in Appendix E. Beyond this, $\Psi/|A|$ can be considered as the marginal effect of income on the social welfare, while $\Omega/|A|$ can be considered as the marginal effect of abatement on social welfare. Thus, $(\Psi/|A|)(\partial Y/\partial t)$ is the marginal social cost (MSC) of environmental tax, while $(\Omega/|A|)(\partial E/\partial t)$ is the marginal social benefit (MSB) of the tax.

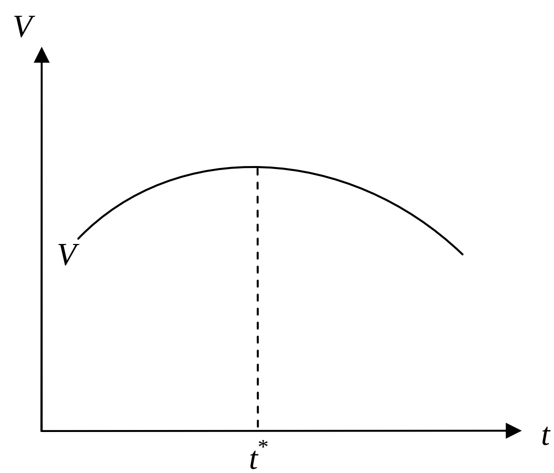
As proved by Appendix F.1, equation (3.35) implies an optimal environmental tax rate t^* exists such that the equilibrium utility (the social welfare of the city) V is maximized by the optimal rate.[33] Since I am using differentiable functions in this model, I can assume that the equilibrium utility V is strictly concave with respect to the environmental tax, which means that the V-t curve is single peaked as described by Figure 1. Given the single peaked property of the V-t curve, I may conclude:

(3.37a) $dV/dt > 0$, iff $t < t^*$

(3.37b) $dV/dt < 0$, iff $t > t^*$

which means an increase of the environmental tax improves the equilibrium utility of the closed city if it happens to be below the optimal environmental tax level, while the increase reduces the equilibrium utility if it happens to be above the optimal tax level.

Figure 3.1: The *V*-*t* Curve



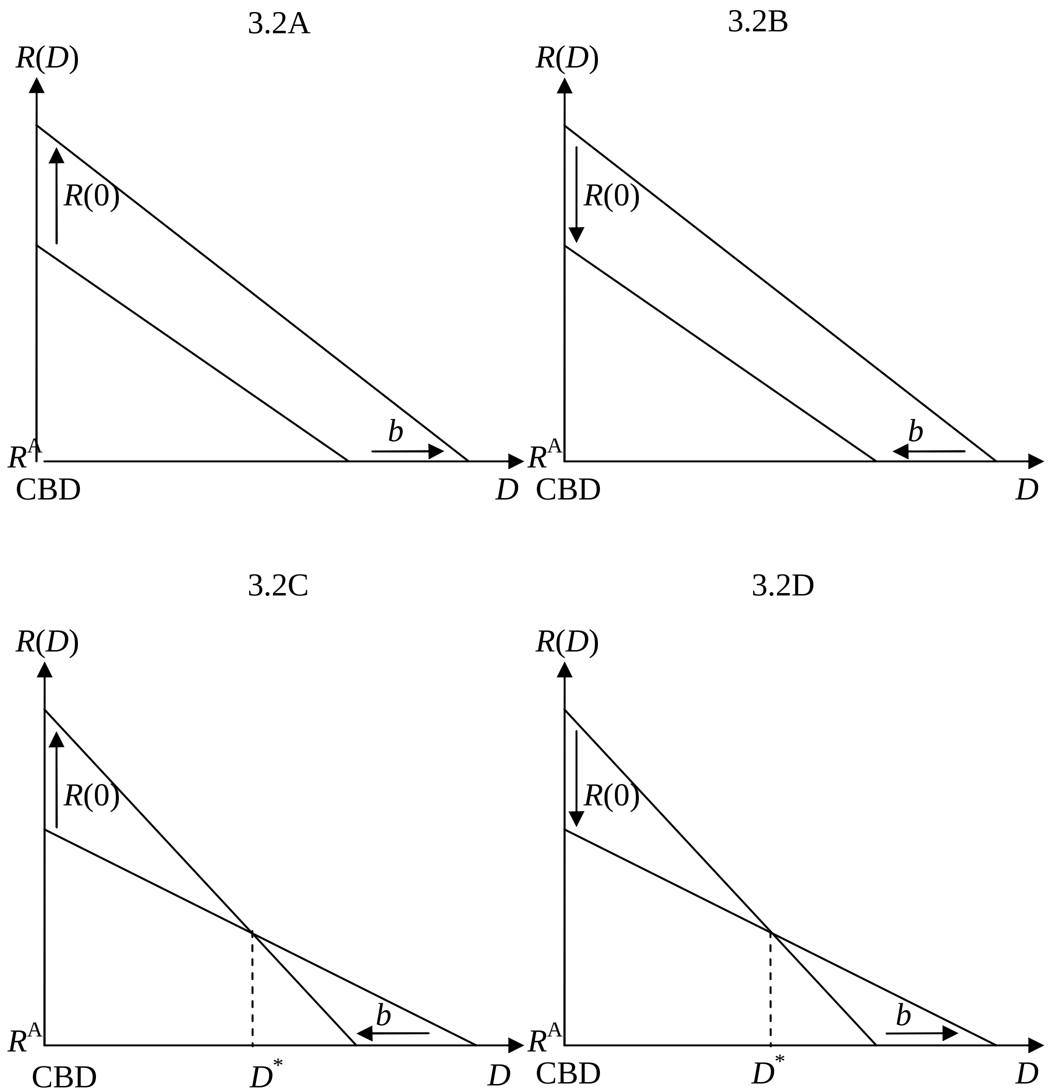
From (3.32) and (3.35), I have the effect of the environmental tax on the land rent at each point:

(3.38) $dR(D)/dt = [(\partial R/\partial Y)+(\partial R/\partial V)(\Psi/|A|)](\partial Y/\partial t)+[(\partial R/\partial E)-(\partial R/\partial V)(\Omega/|A|)](\partial E/\partial t) >< 0$

[33]Corner solutions of the optimal tax rate such as zero or infinite are not interesting and pre-excluded by my assumptions.

which means, as shown by Figure 3.2, the rent gradient $R(D)$ can shifts up (Figure 3.2A), shifts down (Figure 3.2B), pivots clockwise (Figure 3.2C), or pivots counter-clockwise (Figure 3.2D) in response to the environmental tax.[34]

Figure 3.2: The Changes of Rent Gradient[35]



[34]When utility is the only changing factor in the economy, rent gradients do not cross each other. This property, however, does not hold when factors other than the utility also change. Though the new gradient may cross the old several times, I just assume it crosses once to keep the model simple. Please note that the rent gradient is not necessary to be straight line.

[35]The rent gradient may not be straight lines, but curves.

Because all rent gradients continuously slope down from location zero to the boundary and, therefore, any pair of city limit b and rent at location zero $R(0)$ represent one particular rent gradient, the change in the rent gradient can be characterized by the changes in its two endpoints. Thus, how the city limit b and the rent at location zero $R(0)$ vary according to the environmental tax is of interest.

When the distance D is set to be zero, the rent at location zero can be written as (from equation (3.38)):

(3.39) $$\frac{dR(0)}{dt} = [\frac{\partial R(0)}{\partial Y} + \frac{\partial R(0)}{\partial V}\frac{\Psi}{|A|}]\frac{\partial Y}{\partial t} + [\frac{\partial R(0)}{\partial E} - \frac{\partial R(0)}{\partial V}\frac{\Omega}{|A|}]\frac{\partial E}{\partial t} >< 0$$

As proved by Appendix F.2 and F.3, given the strict concavity of V with respect to t, equation (3.36) and (3.39) respectively imply that particular environmental tax rates t_b^* and $t_{R(0)}^*$ exist such that the city limit b and the rent of location zero $R(0)$ are minimized respectively by the rates. Since I am using differentiable functions in this model, I can assume that both the city limit b and the rent of location zero $R(0)$ are strictly convex with respect to the environmental tax, which means that both the b-t and $R(0)$-t curves are U-shaped with single minima as described by Figure 3.3. Thus, the effects of the environmental tax on the city limit and the rent of location zero are:

(3.40a) $db/dt < 0$, iff $t < t_b^*$

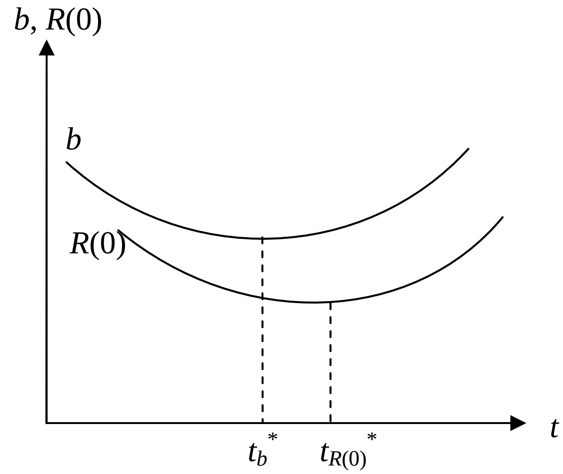
(3.40b) $db/dt > 0$, iff $t > t_b^*$

and:

(3.41a) $dR(0)/dt < 0$, iff $t < t_{R(0)}^*$

(3.41b) $dR(0)/dt > 0$, iff $t > t_{R(0)}^*$

Figure 3.3: The b-t and $R(0)$-t Curves



Intuitively, two offsetting effects on b determine the U-shape relationship of b and t. When the environmental tax increases, the falling pollution emissions encourage residents to move close to the CBD to obtain higher environmental quality with lower commuting costs. The income of residents is decreasing at the same time, however, which means residents have fewer dollars to spend on housing. This income effect means they move away from the CBD to seek lower rents for better housing. The former effect dominates before t_b^* and the latter effect dominates after t_b^*.

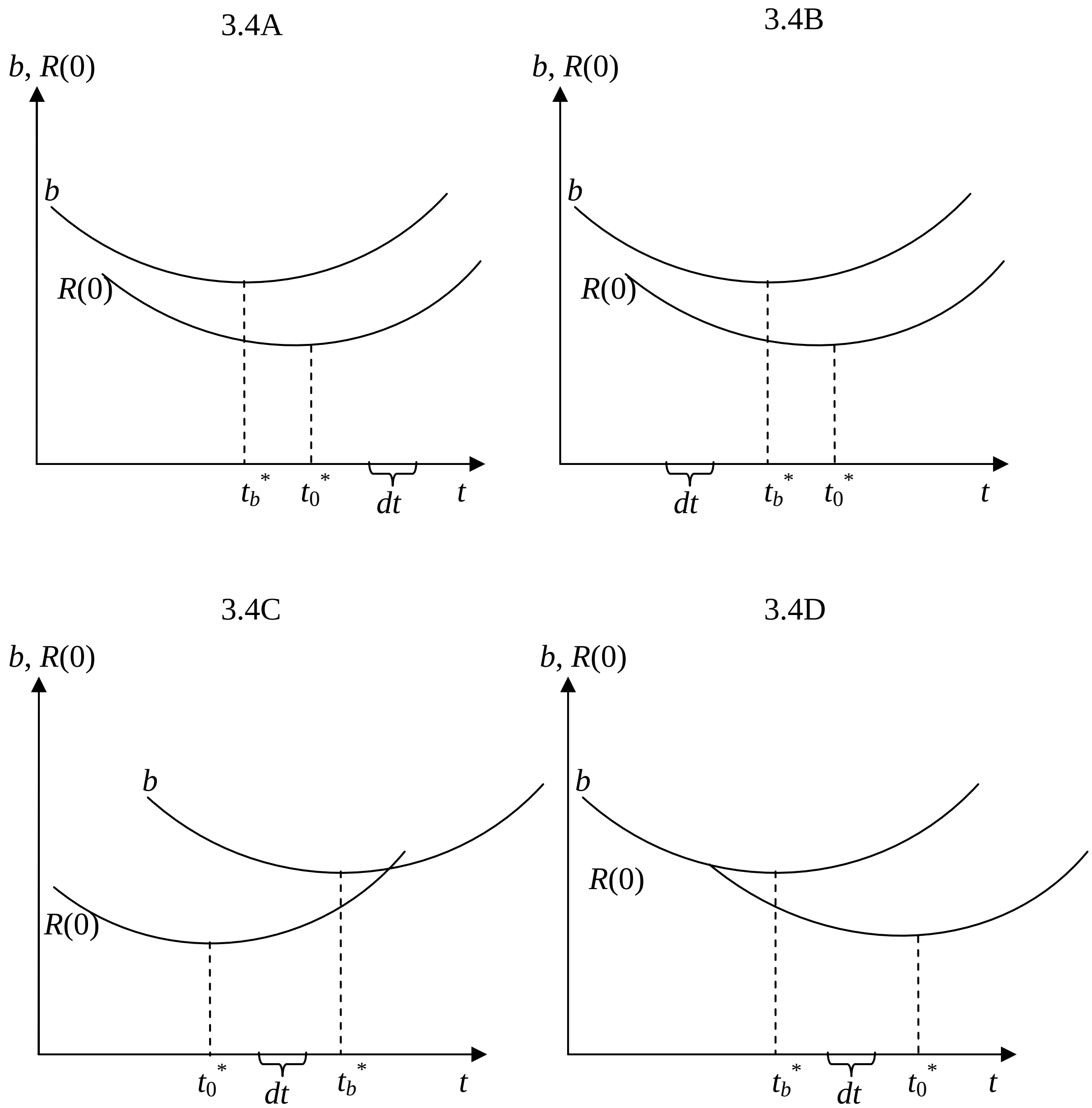
With respect to the rent of location zero, two offsetting effects also determine the U-shape relationship of $R(0)$ and t. On the one hand, when the environmental tax increases, residents bid less for rent, to compensate for their loss of income; which is an "income effect"; on the other hand, better living surroundings pull up the rent of housing at the same time, which is an "environmental effect". Finally, the former effect dominates for tax rates below $t_{R(0)}^*$, and the latter dominates for tax rates above $t_{R(0)}^*$.

Actually, as described in Appendix F.4, the effect on the rent of location zero can be extended to the rent of any location in the city (replacing location zero in (3.41) by any specific location D. That means the rent of the city is U-shaped anywhere with respect to the environmental tax rate.[36] The tax, however, does not affect rent at each location uniformly. Given a change of the tax, the changes of the rents from the "income effect" or the "environmental effect" each vary along with the distance D. Furthermore, the relative magnitudes of those two variations can change, so the minimum of the U-shaped $R–t$ curve change.

Thus, whether the rent gradient shifts up or down, or pivots clockwise or counter-clockwise depends on, as described by Figure 3.4, the relative positions of t_b^* and $t_{R(0)}^*$, and the initial level of the tax. When the initial tax rate is above both t_b^* and $t_{R(0)}^*$ (as in Figure 3.4A), a small increase of the environmental tax results in increases of both the city limit and the rent of location zero, which means the rent gradient shifts up as described in Figure 3.2A; When the initial tax rate is above both t_b^* and $t_{R(0)}^*$ (as in Figure 3.4B), an increase of the tax results in decreases of both the city limit and the rent of location zero, which means the rent gradient shifts down as described in Figure 3.2B. When the initial rate is between t_b^* and $t_{R(0)}^*$, and $t_{R(0)}^*$ happens to be smaller than t_b^* (as in Figure 3.4C), an increase of the tax results in a decrease of the city limit and an increase of the rent of location zero, which means the rent gradient pivots clockwise as described in Figure 3.2C; When the initial rate is between t_b^* and $t_{R(0)}^*$, and $t_{R(0)}^*$ happens to be larger than t_b^* (as in Figure 3.4D), an increase of the tax results in an increase of the city limit and a decrease of the rent of location zero, which means the rent gradient pivots counter-clockwise as described in Figure 3.2D.[37]

[36]Please note that the minima of the U-shapes may not be the same.
[37]Figure F.3- F.6 in Appendix F.5 combine Figure 3.2 and Figure 3.4 together in pairs, which provide more comprehensive comparisons.

Figure 3.4: The Relative Positions of t_b^*, $t_{R(0)}^*$, and the Initial Level of the Tax



Mathematically, if I define D^* as the point where $dR(D^*)/dt = 0$, then:

(3.42) $dR(D)/dt < 0$, iff $t < t_{R(0)}^*$ and $t < t_b^*$

(3.43a) $dR(D)/dt > 0$, iff $t_{R(0)}^* < t < t_b^*$ and $D < D^*$

(3.43b) $dR(D)/dt < 0$, iff $t_{R(0)}^* < t < t_b^*$ and $D > D^*$

(3.44a) $dR(D)/dt < 0$, iff $t_b^* < t < t_{R(0)}^*$ and $D < D^*$

(3.44b) $dR(D)/dt > 0$, iff $t_b^* < t < t_{R(0)}^*$ and $D > D^*$

(3.45) $dR(D)/dt > 0$, iff $t > t_{R(0)}^*$ and $t > t_b^*$

Moreover, D^* can be determined by substituting t, the tax level after the tax change, into (3.38).

Though the effect of the environmental tax depends on the relative positions of t_b^* and $t_{R(0)}^*$, and the initial level of the tax, only two distinct cases exist if the tax continuously changes from quite small to quite large. In the first case characterized by $t_b^* > t_{R(0)}^*$, when the tax increases from small to large, all rents first decrease, forming a downward shift of the rent gradient, then the rents close to the CBD turn back to increase while those away from the CBD still decrease, forming a clockwise pivoting, and at last, all rents increase, forming an upward shift. This case can be represented by sequentially combining Figure 3.2B, 3.2C, and 3.2A or Equation (3.42), (3.43), and (3.45). In the second case characterized by $t_b^* < t_{R(0)}^*$, all rents first decrease, forming a downward shift of the rent gradient, then the rents away from the CBD turn back to increase while those close to the CBD still decrease, forming a counter-clockwise pivoting, and at last, all rents increase, forming an upward shift. This case can be represented by sequentially combining Figure 3.2B, 3.2D, and 3.2A or Equation (3.42), (3.44), and (3.45). Specifically, if both t_b^* and $t_{R(0)}^*$ concentrate, as described by Figure 3.5, in an arbitrarily small interval of t, denoting $\varphi(t)$. Then, when the changes (increases) of the environmental tax stay below $\varphi(t)$, the rent gradient always shifts down because both the city limit and the rent of location zero are decreasing with the tax rate; while *vise versa*.

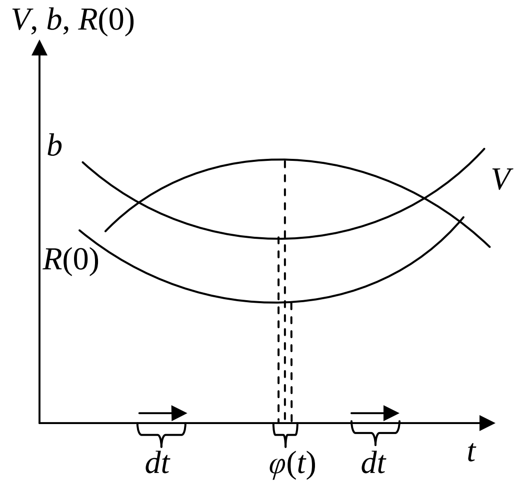
Generally, an environmental tax in a closed city can reduce the pollution emissions in the CBD of the city, and therefore increase environmental quality everywhere. It, however, decreases the wage of labor. Beyond this, when the environmental tax continuously increase from quite small to quite larger, the tax first improves the equilibrium utility of the city, while it reduces the utility after rising above the optimal level. Moreover, it first makes the city shrink but then makes the city grow. Though the tax pulls down the rent of any particular location in the city and then pulls the rent back, it has two distinct effects on the whole rent gradient.

3.4.3. Some Comprehensive Cases

(1). A quite polluted city with inadequate environmental tax as described by Figure 3.4B. In this situation, a small increase of the environmental tax dramatically decreases the pollution emissions in the CBD and, therefore, improves the environmental quality of the whole city significantly. The tax does improve the social welfare too. However, it reduces the wage level of labor. As shown by Figure 3.2B, the tax shifts down the rent gradient, which means the rent decreases everywhere in the city and the landlords bear the tax burden. At the same time, people move close to the CBD to live so that the city shrinks a bit.

(2). A very clean city with extra environmental tax as described by Figure 3.4A. In this situation, a small increase of the environmental tax can still decrease the pollution emissions in the CBD and, therefore, improves the environmental quality of the whole city although the improvement is normally not very significant. Again, it reduces the wage level of labor. As shown by Figure 3.2A, the tax, however, shifts up the rent gradient, which means the rent increases everywhere in the city and the landlords benefit. At the same time, people move away from the CBD to live so that the city expands a bit. Normally, such an increase of the tax reduces the social welfare of the city. Also, rent goes outside the city, so that is a transfer from residents to non-residents. Thus, to cut the environmental tax is a wiser choice for the local authority.

Figure 3.5: The Concentration of t^*, t_b^*, and $t_{R(0)}^*$

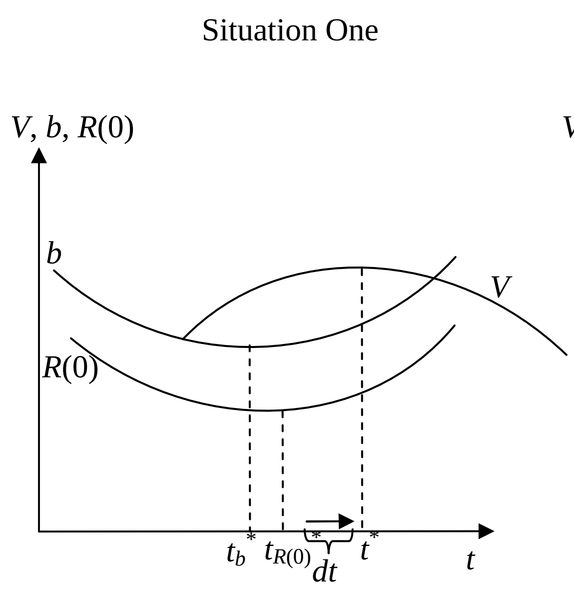


(3). A city that has distinguished environmental tax incidence. A special type of cities has the three tax rates t^*, t_b^*, and $t_{R(0)}^*$ being very close to each other (within the small interval $\varphi(t)$). In such a kind of cities, the environmental tax has distinguished influences on all sides. The environmental tax can reduce the pollution emissions, and therefore improves the environmental quality everywhere. Yet it decreases the wage of labor. Then effects of a change in tax simply depend on whether the initial t is low or high. If the initial rate is low (below $\varphi(t)$ in Figure 3.5), then it makes the city shrink, and pulls down the rent of any particular location in the city. If it is high, then it makes the city grow, and pulls the rent up.

(4). A city in which Pareto improvement environmental tax can be implemented. Typically, the burden of the environmental tax is shifted out of the city as what happens in case (1), which is definitely not Pareto efficient because landlords are bearing the tax burden. The local authority chooses to increase the tax to improve local social welfare but "export the tax burden" to non-residents (in the terminology of local public finance). This, however, may not always be the case. As described in Figure

3.6, a Pareto improvement environmental taxation may occur if t_b^* and $t_{R(0)}^*$ (suppose $t_b^* < t_{R(0)}^*$) appear at the left hand side of t^* and the initial tax rate is right between t^* and $t_{R(0)}^*$; or if t_b^* and $t_{R(0)}^*$ appear at the right hand side of t^* and the initial tax rate is right between t^* and t_b^*. In the former situation, to increase the tax can make both the social welfare of the city and the rent gradient of land rise; while in the latter situation, to decrease the tax can do so.

Figure 3.6: Pareto Optimal Taxation



3.5. CONCLUSION AND DISCUSSION

Generally, an environmental tax in a closed city targeting polluting emissions can reduce the pollution emissions in the city's central business district (CBD), and therefore improves environmental quality of the city everywhere. This is because the tax increases the price of pollution emissions (as a factor input) so that less pollution emissions are used in the production of the composite good. In addition to the reducing of pollution, the tax reduces the wage received by labor. This is because firms try to shift the burden of the environmental tax onto labor, which is inelasticly supplied. Lower pollution as an input means a lower marginal product of labor. Overall, as the environmental tax rate rises from quite low to quite high, the tax first improves the social welfare of the city until it hits the optimal level, and then it reduces welfare as the marginal benefit falls below the marginal cost of the tax. Furthermore, for a fixed number of residents, the tax first makes the city boundary shrink but then makes it grow as residents try to balance between better housing and shorter commuting. Initial increases in the tax pull down the rent for any particular location in the city and further increases pull the rent back up. The lower bids of residents that result from lower income are offset by the higher bids that result from the pursuit of better environmental quality.

The environmental tax has two possible effects on the whole rent gradient because it does not affecting the rent of every location uniformly. In the fist case, when the tax increases from small to large, all rents first decrease, forming a downward shift of the rent gradient, then the rents close to the CBD turn back to increase while those away from the CBD still decrease, forming a clockwise pivoting, and at last, all rents increase, forming an upward shift. In the second case, all rents first decrease, forming a downward shift of the rent gradient, then the rents away from the CBD turn back to increase while those close to the CBD still decrease, forming a counter-clockwise pivoting, and at last, all rents increase, forming an upward shift.

In a quite polluted city with an inadequate environmental tax, the increase of the tax shifts down the rent gradient, decreasing the rent of land everywhere in the city. In contrast, for a very clean city with a high environmental tax, an increase of the tax shifts up the rent gradient, increasing the rent of land everywhere. In one type of theoretically defined city, as the environmental tax increases from quite low to quite high, it first improves the social welfare of the city, shrinks the size of the city , and cuts the rent of any particular location in the city; and then it reduces the social welfare, expands the city, and raises the rent anywhere. In another type of theoretically defined city, furthermore, a Pareto improvement is possible, which means that the tax may increase both the social welfare of the city residents and the gradient of rent received by landlords.

The model of this research depends essentially on two facts: the inelastic labor supply of residents and the land rent decision mechanism. The former ensures all residents are fully employed, while the latter ensures that residents and landlords can bargain with each other for land rent. The land rent decision mechanism developed by Alonso (1964) reflects the natural essence of land markets. Thus, it is not flawed to use the mechanism in this research. Though the inelastic labor supply of residents is limited because leisure is not considered in the model, it is still common in theoretical urban economics models to see such an assumption. A further development of this paper can be one including labor-leisure choice in the individual's behavior, which definitely makes the research more complicated.

Chapter 4: The Effect of Corporate Income Tax on Organizational Forms

4.1 INRODUCTION

The corporate income tax has been criticized for a long time. Some economists say that they cannot see any reason to tax a corporation, because corporations are actually owned by individuals already bearing personal income tax. The coexistence of a corporate tax and a personal income tax results in double taxation only on corporation shareholders, which is inefficient because it discourages investment in the corporate sector and lowers the capital level below an optimal one. So, the corporate tax must be a distortion resulting in excess burden. However, the other economists argue that a corporate tax is necessary for other reasons such as principle-agent problems, limited liability of corporate shareholders, public mobility of corporate shares, and potential tax avoidance problems. Among those reasons, the limited liability of corporate shareholders and the public mobility of corporate shares are privileges of the organization form of corporations compared to non-corporate firms that face unlimited liability and limited mobility of ownership. In this view, the extra tax burden and the privileges of the corporate form are two offsetting factors considered seriously when investors decide to choose a particular organizational form for their investment.

Nevertheless, the literature has conflicting conclusions on the effect of the extra tax on the distribution of economic activity between the corporate and non-corporate sectors. Gravelle and Kotlikoff (1989 and 1993) theoretically find huge deadweight loss (DWL) caused by the extra corporate tax, because it distorts the choice of organizational forms. Mackie-Mason and Gordon (1997) and Goolsbee (1998), however, empirically find that the effect of the extra corporate tax on the corporate share of economic activity, especially on the corporate share of capital, is statistically significant but every small in magnitude. So, further research is needed to explore the truth.

This paper also empirically tests the effect of the extra corporate tax on the choice of organizational form between corporate and non-corporate form. It follows the works of Mackie-Mason and Gordon (1997) and Goolsbee (1998) by reexamining the effect of the extra corporate tax on the corporate share of capital, by using new data with a longer time span than used in previous investigations, by studying the role of transaction costs played in the effect of the extra corporate tax on changes of organizational forms, and by estimating the effect via new investment data instead of capital stock data used in the literature.

I find that the overall extra corporate tax rate has significantly negative effects on the corporate share of economic activity, which is consistent with findings in the literature. The effect of the extra corporate tax on the corporate share of capital stock, which is estimated by data with a longer time span than those used in the literature, stands just between two major estimates in the literature. Furthermore, the effect of the extra corporate tax on the corporate share of investment is much larger than the effect on capital. For example, a 10 percent increase in the extra corporate tax rate only results in a 0.18 percent decrease in the corporate share of capital, while it results in a one percent decrease in the corporate share of investment. These results are consistent with transaction costs in two respects: first, the corporate share of capital stock does show adjustment lags, which is different from what Mackie-Mason and Gordon (1997) find; second, the extra corporate tax shows larger effect on the corporate share of investment than on the corporate share of capital stock. Investment can react more easily than capital stock. If capital stock could be added or subtracted with less transaction costs, then the corporate share of capital would have responded to the extra corporate tax as sensitively as the corporate share of investment. For example, suppose my estimates were used to predict effects of a 10% increase in the extra corporate tax rate. If capital had been as liquid as investment, then 89 billion dollars of capital would have shifted out the corporate sector in 1997 instead of only 17 billion dollars of capital actually shifted.

Furthermore, estimates here account for the fact that the economy tends to increase the share of corporate assets during war time. Finally, the structural changes of tax system caused by the Tax Reform Act of 1986 encouraged people to increase the share of non-corporate assets.[38] These structural changes may affect the way that a change in corporate tax rate would affect the share of investment going into the corporate sector.

The paper proceeds as follows. Part 2 illustrates the background in the literature; Part 3 presents a simple theoretical model describing how firms decide whether to incorporate or not; Part 4 describes data and specifications; Part 5 shows the results; finally, Part 6 discusses conclusions and possible extensions.

4.2 LITURATURE REVIEW

The relationship between the corporate income tax and the choice of organizational form has been discussed a lot in the literature. Early works study this issue theoretically using Harberger's (1962) model, while recent works empirically estimate the effect of the extra corporate tax on the corporate share of economic activity

---

[38]Please see Auerbach (1987) and Pechman (1987) for overviews. Fullerton (1996) summarizes six consequences of the structural changes of Tax Reform Act of 1986: first, people with high income report more income from partnerships, which belongs to the non-corporate sector; second, they report more income from S-corporation, which is also belongs to the non-corporate sector; third, people who are both shareholder and manager of corporations assign higher salary to themselves; forth, economics activities are shifted from corporations to non-corporate firms caused by the repeal of the General Utilities; fifth, economics activities are shifted from corporations to non-corporate firms caused by the expansion of the corporate alternative minimum tax; sixth, people have less incentive to change labor or other income to capital gains.

such as the capital stock, annual income (loss), employment, firm sales, and even the number of firms.

In his remarkably influential paper, Arnold Harberger (1962) provides a framework to analyze the effects of the corporate income tax using a general equilibrium model. Harberger's model assumes the separation of a corporate sector and non-corporate sector in an economy, as is followed by Harberger himself (1966), Shoven (1976) and Ballard *et al* (1985). It is an issue because unlike Harberger's assumption, only few industries belong solely to the corporate sector or to the non-corporate sector. Moreover, as pointed out by many economists, Harberger's model is not easily adjusted to allow for corporate production of non-corporate goods or non-corporate production of corporate goods. Thus, any empirical work in this field has to use a mutual production model.

Gravelle and Kotlikoff (1989 and 1993) and Fullerton and Rogers (1993) present models with corporate and non-corporate production of both corporate goods and non-corporate goods within the same industry. Moreover, they assume a big elasticity of substitution between organizational forms within each industry. Gravelle and Kotlikoff investigate the organizational form together with the dead weight loss (DWL) of corporate income taxation and find that the incidence of the corporate income taxation and DWL in their mutual production model (MPM) differ very much from those in the Harberger's model. In a Harberger model, the implied DWL is about 10-20% of the corresponding tax revenue, while Gravelle and Kotlikoff find that the DWL could be greater than 100% of the corporate income tax revenue. Compared to conventional models, Gravelle and Kotlikoff's model points out that the organizational form of a firm changes when the corporate income tax changes.

According to the traditional treatment, the corporate income tax is considered to be an extra tax on the corporate income. This implies that the tax penalizes investments in the corporate sector, resulting in a smaller fraction of the total capital stock than what would be used in corporations if the same rate of tax were imposed in both the corporate and non-corporate sectors. Feldstein and Slemrod (1980) point out that such a widely accepted conception relies on an assumption of a simple personal taxation system with a single rate of tax. In their model, they replace the flat rate personal tax system with progressive personal tax rates. Then, those with high personal tax rates are willing to keep their capital gains within the corporations, since capital gains bears a relatively low effective personal tax rate. By using a more generally balanced portfolio model, they show surprisingly that a higher corporate income tax could increase, not decrease, the corporate share of total capital.

Besides the corporate tax, some non-tax factors also have striking effects on the choice of organizational form of firms. Fama and Jensen (1983) point out that non-tax factors appear to dominate in the choice of organizational form, although it is still not clear which factor is the most important. Mackie-Mason and Gordon (1997) analyze two main factors, the limited liability of corporate shareholders and the mobility of the shareholders in publicly trading their shares. Though non-corporate owners are also commonly able to get a similar right of limited liability, it is still considered one of the most important privileges of the corporate sector. The mobility of corporate shares is another big advantage of corporations. This mobility gives corporations priority access to low-cost equity capital and to an efficient solution of the principle-agent problem.

Several other empirical works are available on the relationship between corporate tax and the choice of organizational form of firms. Mackie-Mason and Gordon (1997) investigate empirically the extent to which the corporate share of total assets respond to the changes in the corporate income tax. Using data covering the period 1959-1986, their research supports the negative relationship between the corporate share of assets and the corporate tax rate. Goolsbee (1998) continues Mackie-Mason and Gordon's work by using a different data set from 1900 to 1939. In his paper, Goolsbee recalls the background of corporate taxation in the early 20th century and then represents the Mackie-Mason and Gordon theory and regression models to estimate how corporate taxation discourages incorporating. As controls in his model, he uses macroeconomic variables such as the GNP growth rate, unemployment, and the interest rate. Only the GNP growth rate, however, proves significant. Goolsbee finds that the effect of the corporate tax on incorporation is negative and significant, but small in magnitude, a similar result to that of Mackie-Mason and Gordon. Very recent empirical work is also performed by Goolsbee (2004). He uses panel data to investigate the effect of the extra corporate tax on the corporate share of employment, firm sales, and even the number of firms.

4.3 A SIMPLE THEORETICAL MODEL

The theory used in this paper is a simple decision model that indicates whether firms incorporate or not. The model is first introduced by Gordon and Mackie-Mason (1994) and Mackie-Mason and Gordon (1997) and then followed by Goolsbee (1998 and 2004).

When each firm is starting up, investors can choose between two kinds of organizational form, corporate and non-corporate. Both forms, after being established, are assumed to generate exactly the same future return Y, per unit of investment. Compared to corporations, non-corporate firms have an additional benefit G (or cost, if $G < 0$), also per unit of investment, due to various advantages or disadvantages of non-tax factors. These non-tax factors mainly include that corporations have limited liability and can access public stock market (disadvantages of non-corporate firms), while owners of non-corporate firms can deduct business losses against other personal incomes in the calculation of personal tax liability (an advantage of non-corporate firms).[39] Following what is done in the literature, I suppose G is not taxable because the government is not able to observe or measure it. Moreover, the additional benefit (or cost) G is supposed to be the same for each firm and unchangeable across industries and over time.

Corporations and non-corporate firms face different forms and levels of taxation. Non-corporate firms face only the personal income tax at rate t_p, while corporations face both the personal income tax at rate t_p and the corporate income tax at rate t_c. Though corporations bear two statutory tax rates, they may not be treated that unfairly after legal tax avoidance. Corporate shareholders can offset some of their disadvantages in income tax rates through accumulating their income at a higher real rate of return by leaving the income, partly or totally, within the corporation, which means that they bear only corporate income tax during that time. These retained earnings might increase share prices, but capital gains are taxed at a lower effective personal income tax rate than are dividends. Thus, corporate shareholders could also decrease their income tax by transforming dividend income into capital gains and then selling their shares in the stock market.

Suppose γ is the share of corporate income distributed as dividends, $1\text{-}\gamma$ is the share of corporate income kept within the corporation, and t_g is the effective capital gains tax rate. Corporations then bear an overall personal income tax rate on equity:

---

[39]More non-tax factors can be found in Mackie-Mason and Gordon (1997).

(4.1) $t_e \equiv \gamma\, t_p + (1\text{-}\gamma)\, t_g$

The effective capital gains tax rate can be expressed as:

(4.2) $t_g = \alpha\, \beta\, t_p$

where α represents the taxable share of capital gains, and β is a factor describing the effect of legal tax avoidance such as deferral advantage and step up of basis at death. Defining t_c+(1-t_c)t_e as the overall effective corporate income tax rate, the net return of a corporation is:

(4.3) $I_c = Y[1\text{-}t_c\text{-}(1\text{-}t_c)t_e]$

while that of a non-corporate firm is:

(4.4) $I_n = G + Y(1\text{-}t_p)$

A firm prefers to incorporate if and only if:

(4.5) $I_c > I_n$

After several algebra steps, that condition is the same as:

(4.6) $\text{-}G > Y[t_c\text{+}(1\text{-}t_c)t_e\text{-}t_p]$

Denoting $[t_c\text{+}(1\text{-}t_c)t_e]\text{-}t_p$ as T, the rate of overall extra tax on corporations relative to non-corporate firms, then the condition for incorporating is:

(4.7) $\text{-}G > YT$

Empirical work below uses S to denote the corporate share of activity in the economy. Given that Y and G are fixed, it is T, the overall extra corporate income tax rate, that plays a decisive role in whether or not to incorporate – that is, a decisive role in determination of S, the corporate share of economic activity. If $T < \text{-}G/Y$, a firm will choose the corporate form; and if $T > \text{-}G/Y$, a firm will choose the non-corporate form. An implicit assumption of this model, as pointed out by Mackie-Mason and Gordon (1997) themselves, is zero transaction costs such that the organizational form of a firm can be freely transferred between the corporate form and the non-corporate form according to changes of the overall extra corporate tax rate.

4.4 DATA AND SPECIFICATIONS

4.4.1 Data

A data set is collected from several reliable resources such as publications of the Internal Revenue Service (IRS) and the Bureau of Economic Analysis (BEA). These data include annual capital stock and investment data in fixed nonresidential private capital assets in the United States, published in the *Fixed Reproducible Tangible Wealth in the United States, 1925-1997* by BEA. The data include information about the distribution of capital and investment between corporate and non-corporate organizational forms, which can ideally be used to estimate the effect of the extra corporate tax on changes in organizational forms adopted by capital stock or investment. The data for capital stock are available from 1925 to 1997, while the data for investment are available from 1901 to 1997. Corporations in this publication include "all entities required to file federal corporate income tax returns (IRS Form 1120 series)" as described by BEA (1999), while non-corporate firms include sole proprietorships and partnerships.

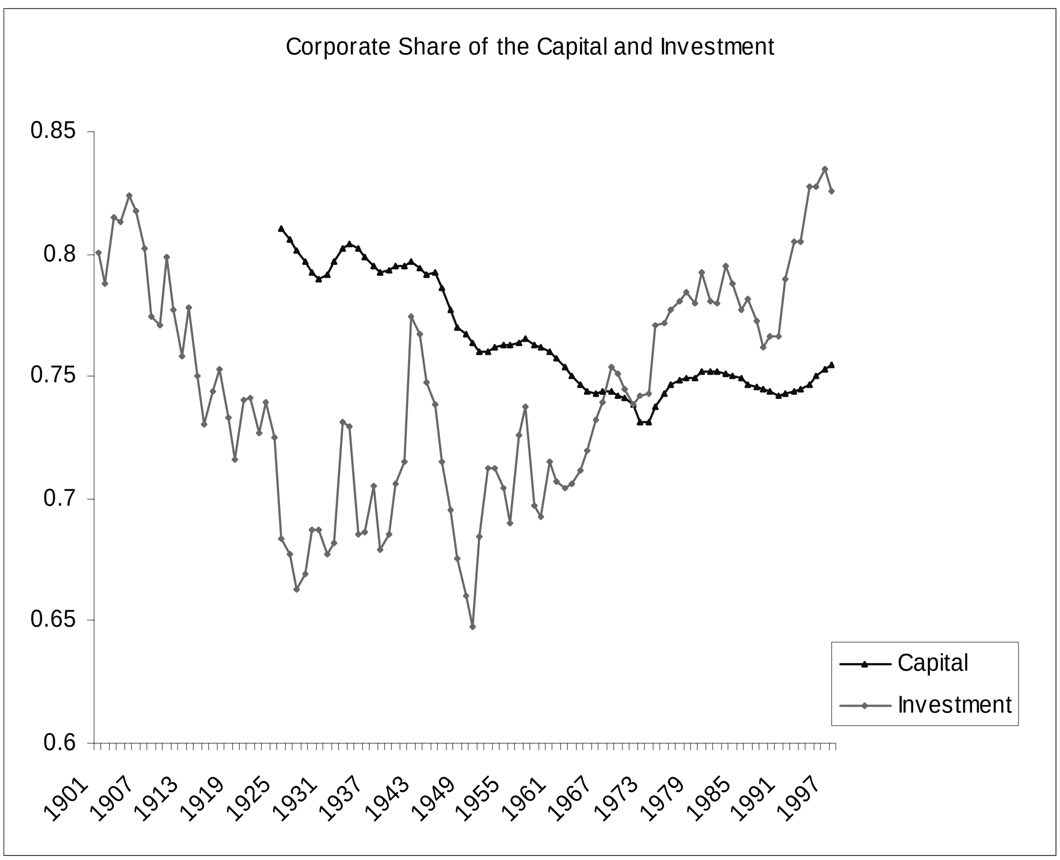
Table 4.1: Statistics of Corporate Shares and Tax Rates

	Mean	s.d.	Min	Max
Corporate share of capital stock (1925-1997)	76.48%	2.28%	73.12% (1973)	81.03% (1925)
Corporate share of investment (1901-1997)	74.33%	4.57%	64.79% (1950)	83.48% (1996)
Overall extra corporate tax rate (1901-1997)	-0.103	0.100	-0.165 (1937)	0.199 (1989)

Annual corporate shares of capital and investment are calculated using those data. Mackie-Mason and Gordon (1997) only use recent capital data from 1959 to 1986, and

Goolsbee (1998) only uses old capital data from 1900 to 1939.[40] In contrast, I have the corporate shares of capital stock from 1925 to 1997, and I also use the corporate shares of investment from 1901 to 1997. Some descriptive statistics of the corporate shares of capital and investment are presented in Table 4.1. Briefly, in the last century, corporations constitute most of the total capital and account for most of the investment activity in the United States. The corporate share of investment, however, varies more widely than does that of capital, from 0.835 in 1996 to 0.648 in 1950 (versus from 0.810 in 1925 to 0.731 in 1973).

Figure 4.1: The Corporate Share of the Capital Stock and Investment.



[40]Though Goolsbee uses data from the same source as mine, he only uses capital data from 1900 to 1939. Furthermore, the capital data from 1900 to 1924 are not available in the officially published version of the book.

Figure 4.1 shows the time trends of the corporate shares of capital stock and investment. Generally, the corporate share of investment decreases overall before the middle of last century, while it increases after that. With respect to the corporate share of capital, it decreases overall till the 1970s, but after 1973, it increases overall.[41] Moreover, the trajectory of the corporate share of investment is below that of capital before the 1970s, but they reverse after that. This indicates that the rise of the corporate share of investment from the 1950s finally results in a rise of the corporate share of capital after roughly two decades.

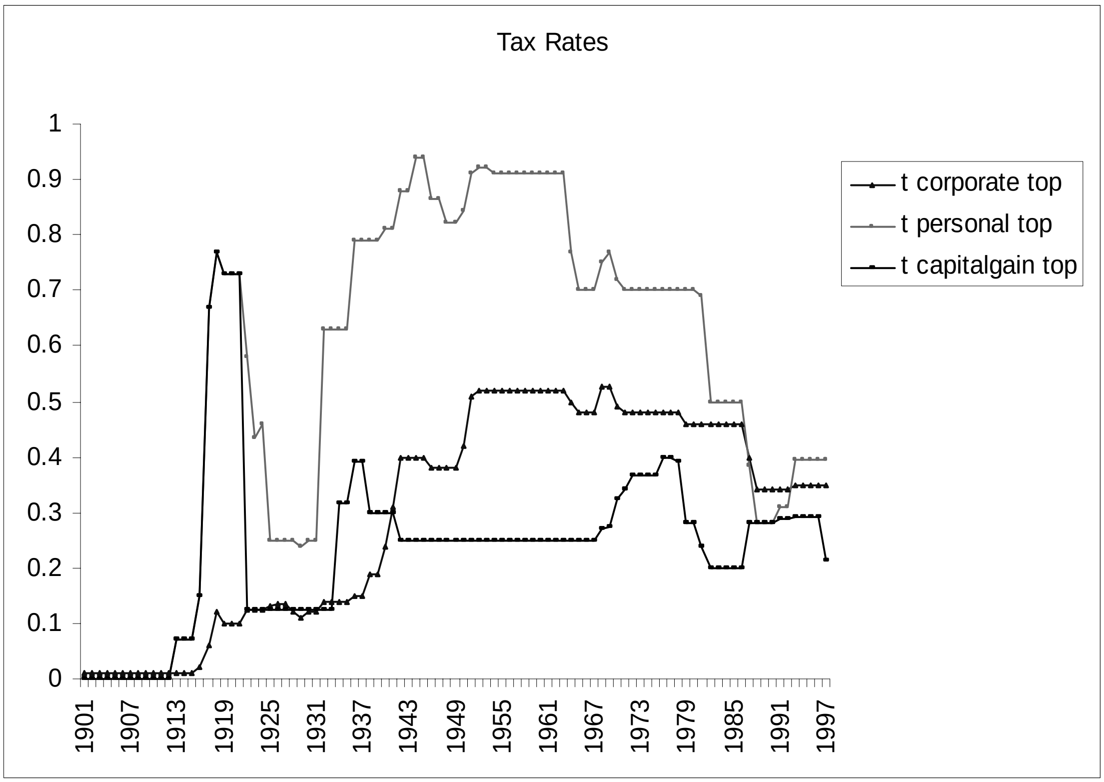
The annual personal income tax rate, corporate income tax rate, and capital gains exemptions from 1901 to 1997 are also included in the data set. Based on those rates, I calculate the effective capital gains tax rate $t_g = \alpha\beta t_p$, the effective equity income tax rate $t_e = \gamma\ t_p + (1\text{-}\gamma)\ t_g$, the overall corporate income tax rate $t_c + (1\text{-}t_c)t_e$, and the overall extra tax rate on corporations $T = [t_c+(1\text{-}t_c)t_e]$ - t_p. Particularly, I fix γ to be two-thirds, following Goolsbee (1998) who estimates that corporations distribute as dividends roughly two-thirds of total corporate incomes between 1916 and 1939. I also fix β to be 0.25, which is originally found by Feldstein *et al* (1983) and followed by Mackie-Mason and Gordon (1997) and then by Goolsbee (1998). Furthermore, if a graduated personal tax rate system exists, the rate of the top tax bracket is used in my calculation.

Some descriptive statistics of those tax rates are presented in Table 4.1. Some figures below also show the changes of those tax rates over time. Figure 4.2 shows how the personal income tax rate, corporate income tax rate and capital gains tax rate change over the last century. They all change many times, while both the personal income tax rate and corporate income tax rate vary more frequently than that of the capital gains tax.

[41]This is different from the data of Mackie-Mason and Gordon (1997), which show that the corporate share of assets is strictly decreasing from 1957 to 1989.

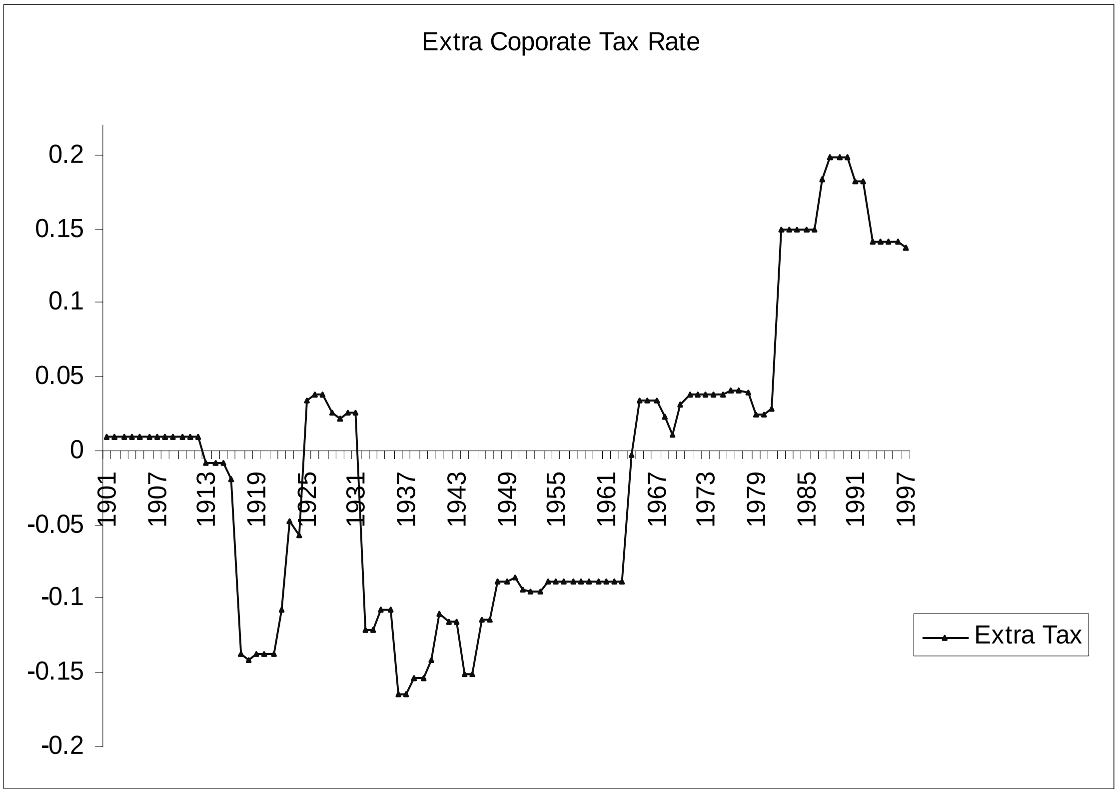
Furthermore, both the personal income tax rate and corporate income tax increase dramatically during the 1930s and the early 1940s. After that, those rates stay at a high level from the middle of the 1940s to the middle of the 1960s. Then, they decrease to much lower levels. With respect to the capital gains tax rate, it is almost unchanged from 1942 to 1967.

Figure 4.2: Various Tax Rates



As described in Figure 4.3 and Table 4.1, the overall extra corporate tax rate changes frequently and significantly around zero in the last century. The minimum is a negative 0.165 in 1936 and 1937, and the maximum is a positive 0.199 from 1988 to 1990. Generally, the overall extra corporate tax rate is negative from 1932 to 1963, while it is positive after 1965.

Figure 4.3: The Extra Corporate Tax Rates



4.4.2 Specifications

From the analysis of the theoretical model, the corporate share of economic activity S, such as that of capital or investment, is expected to be highly correlated with the overall extra corporate tax rate T. So, a best linear predictor of the corporate share of economic activity is a linear combination of T and other control variables. Two equations are estimated here, one of which has the corporate share of capital as the dependent variable, while the other has the corporate share of investment as the dependent variable. Following Mackie-Mason and Gordon (1997) and Goolsbee (1998), I include the intercept, the overall extra corporate tax rate, time, and time squared in the equations as independent variables. Time and time squared are used to control for time trend. Beyond this, I also include dummy variables to control for the WWII and the Tax Reform Act of 1986 (TRA86). Particularly, the TRA86 dummy variable helps control for the effect of changes of corporate tax rules and tax structures in 1986 other than the changes of tax rates.[42] Those two equations are:

(4.8) $SC_t = \alpha_0 + \alpha_1 T_t + \alpha_2\, time + \alpha_3\, (time)^2 + \alpha_4\, (WWII) + \alpha_5\, (TRA86) + \varepsilon_t$

(4.9) $SI_t = \beta_0 + \beta_1 T_t + \beta_2\, time + \beta_3\, (time)^2 + \beta_4\, (WWII) + \beta_5\, (TRA86) + \eta_t$

where SC_t is the corporate share of capital stock in year t, SI_t is the corporate share of investment in year t, T_t is the overall extra corporate tax rate in year t, time is the year itself divided by 100, *WWII* is the dummy variable for years from 1941 to 1945, and *TRA*86 is the dummy variable for all years after the Tax Reform Act of 1986 (year 1986 included). The error terms ε_t and η_t are assumed to have first order serial correlation (AR(1)) following Mackie-Mason and Gordon (1997) because time series data are used. Each such regression uses pure time series data with all the years available for the U. S. (*SC* for 1925 -1997 and *SI* for 1901 - 1997). Except for the new dummy variables (*WWII* and *TRA*86), the first equation is originally used by Mackie-Mason and Gordon (1997) for 1962 – 1986 and by Goolsbee (1998) for 1900 – 1939. The *SI* equation is new.[43]

In addition, the equations are estimated with additional independent variables such as the dependent variable lagged by one, two, or three years to control for transaction costs, the extra corporate tax lagged by one or two years to control for the long run effect of the tax, and the GDP growth rate to control for the overall economic

[42] Comprehensive examinations of the effect of the changes of tax structures of TRA86 on economic activities can be found in Pechman (1987) and Auerbach (1987). As pointed by them, changes of tax rules other than tax rates affect economic activities seriously, if not more seriously than the rate does. Thus, a control for TRA86 is necessary to make my estimator for the effect of the tax rate more robust.

[43] Mackie-Mason and Gordon (1997) do not have the WWII and TRA86 dummies in the equation. Goolsbee (1998) also includes a GNP growth rate in the equation to control for the effect of overall economic growth.

growth and the business circle. Particularly, when lagged dependent variables and lagged extra corporate tax rates are added to the right hand side (RHS) of the equations, the error terms (ε_t and η_t) are assumed to be without any order serial correlation because lagged variables are already included as independent variables.

When transaction costs of transfers between organizational forms are considered, equation (4.9) with SI could have advantages over (4.8) with SC. Investment is new capital added in the current period, and it normally can react sooner than the capital stock, which is accumulated investment of many previous periods. Therefore, the corporate share of investment is expected to be more sensitive to the overall extra corporate tax rate than that of capital. Mackie-Mason and Gordon (1997) point out that transaction costs can be very important in preventing firms from transferring their organizational forms between the corporate and the non-corporate form. Although those authors are aware of the importance of transaction costs, they just assume firms are able to change their forms freely. Later, Goolsbee (1998) also investigates only the corporate share of capital and ignores transaction costs. Thus, given that investment has fewer transaction costs than capital, a further investigation of the effect of the extra corporate tax on both the corporate share of investment and the corporate share of capital stock can help explain the role of transaction costs.

The coefficients of the overall extra corporate tax rate are expected to be negative in both equations (4.8) and (4.9). The implication is that a higher overall extra corporate tax reduces the corporate share of capital or investment. Furthermore, β_1 is expected to be larger than α_1 because changes of investment have fewer transaction costs or lags than those of capital stock. Thus, the effect of the extra corporate tax on the corporate share of investment is expected to be larger than that of capital. Finally, I have no expectation about the coefficients of time effects.

4.5 RESULTS

4.5.1 Effects on the Corporate Share of Capital Stock

Estimates of equation (4.8) for corporate capital stock (SC) are presented in column 1 of Table 4.2. Then column 2 shows the results when lagged dependent variables are included in equation (4.8). Column 1 is estimated based on 73 observations from 1925 to 1997 by using ordinary least squares (OLS) with first-order serial correlation corrections, and the Durbin-Watson statistic is calculated after the correction. Column 2 is estimated based on 70 observations from 1928 to 1997 by using OLS only.[44] For comparison, results of Mackie-Mason and Gordon (1997) and Goolsbee (1998) are also presented in columns 3 and 4 of Table 4.2.

First consider the (SC) estimation (equation (4.8) with result in column 1 of Table 4.2), which is similar to estimations of Mackie-Mason and Gordon (1997) and Goolsbee (1998). When I add more years of data, I find that the coefficient on the extra corporate tax rate is -0.0317, the absolute value of which is similar to the estimate of Goolsbee, 0.0304, which is supposed to have the opposite sign because he uses the non-corporate share of capital stock as the independent variable. However, it is lager than that of Mackie-Mason and Gordon, -0.00920. The -0.0317 is not reliable, however, because the corrected Durbin-Watson statistic is still too low to get rid of the serial correlation problem.[45]

[44]Limdep 7.03 is used in both estimations.

[45]When I try to estimate equation (4.8) using OLS with first-order serial correlation correction and using data from 1962 to 1986 as Mackie-Mason and Gordon (1997) do in their work, I get statistically insignificant coefficient on the extra corporate tax and a bad Durbin-Watson statistic. Please refer to column 1 of Table G.1 in Appendix G for the result. This may be due to the fact that I use data from a source different from theirs. They manually aggregate data “from numerous IRS publications and data tapes” (page 486, Mackie-Mason and Gordon 1997), while I get aggregated data directly from the BEA publications. Furthermore, as described in my Footnote 41, their corporate share of capital stock has different trends from mine during 1957 to 1989. With respect to replicating the work of Goolsbee (1998), since the capital stock data are not available for 1900 – 1924, I am not able to do so.

Table 4.2: Estimated Coefficient for the Corporate Share of Capital Stock

(standard errors in the first parentheses, and *p*-value in the second parentheses)

	Equation (4.8) 1925-1997	Equation (4.8) with lagged dep. var. 1928-1997	Mackie-Mason and Gordon 1962-1986	Goolsbee (non-corp. share of capital stock) 1901-1939
Intercept	100.1197	29.3325	0.964	0.071
	(20.5809)	(5.7876)	-	-
	(0.0000)	(0.0000)	-	-
Extra tax rate	-0.0317	-0.0181	-0.00920	0.0304
	(0.0117)	(0.0055)	-	-
	(0.0067)	(0.0017)	-	-
Time	-10.0615	-3.0056	0.000765	0.037
	(2.1008)	(0.5952)	-	-
	(0.0000)	(0.0000)	-	-
$Time^2$	0.2547	0.0774	-0.954E-4	0.0018
	(0.0536)	(0.0154)	-	-
	(0.0000)	(0.0000)	-	-
WWII	0.0034	0.0030		
	(0.0020)	(0.0010)	-	-
	(0.0855)	(0.0031)		
TRA86	-0.0029	-0.0017		
	(0.0028)	(0.0012)	-	-
	(0.3013)	(0.1489)		
Lagged dep. Var. (-1)		1.5347		
	-	(0.1170)	Not Significant	-
		(0.0000)		
Lagged dep. Var. (-2)		-0.9495		
	-	(0.1888)	-	-
		(0.0000)		
Lagged dep. Var. (-3)		0.2183		
	-	(0.1089)	-	-
		(0.0494)		
Durbin-Watson	0.9307	1.8212	1.71	-
Coefficient in AR(1)	0.9142			
	(0.0477)	-	-	-
	(0.0000)			

Note: The “dep. var.” means dependent variable, and the “non-corp.” means non-corporate.

When three lagged dependent variables (one, two, and three periods lagged) are added to the RHS of equation (4.8) (column 2 of Table 4.2), and the error term ε_t is assumed to be standard normal, the serial correlation problem is solved, so that I can get a very good Durbin-Watson statistic (1.8212).[46] Thus, I feel that the -0.0181 is a more accurate estimate of the coefficient of the overall extra corporate tax rate than -0.0317. Compared to the figure of Mackie-Mason and Gordon (1997), -0.0092, this estimate of -0.0181, estimated by adding more year to their work, represents a larger effect of the extra corporate tax on the corporate share of capital stock; compared to the figure of Goolsbee (1998), 0.0304, the -0.0181 represent a smaller effect.

The coefficient of WWII (0.0030) is positive and statistically significant, which indicates that the economy tends to increase the share of corporate assets during war time. The coefficient of TRA86 (-0.0017) is negative but not statistically significant.[47] The sign of this coefficient, however, is consistent with the prediction of Fullerton (1996) that the structural changes of both personal and corporate taxation in TRA86 encourage investors to increase the share of non-corporate assets.[48]

To control for the long run effect of the tax, two lagged extra corporate tax rates (one and two periods lagged) are added to the RHS of equation (4.8) with lagged dependent variables, but their coefficients are not statistically significant.[49] This indicates the extra corporate tax rate tends to affect the current corporate share of capital

[46]The dependent variable lagged four periods is not statistically significant. Please refer to column 2 of Table G.1 in Appendix G for the result when the forth lagged dependent variable is added to the RHS of equation (4.8).

[47]When the dummy variable TRA86 is eliminated in the regression, the coefficients of other independent variables change a bit. Please refer to column 3 of Table G.1 in Appendix G for those changes.

[48]Fullerton summarizes the Tax Reform Act's six structural changes of the personal income taxation other than changes of tax rates. Four of the six changes (the first, the second, the forth, and the fifth) encourage investors to invest in the non-corporate section, while the other two (the third and the sixth) are neutral. These considerations suggest a negative sign for the TRA86 dummy.

[49]Please refer to column 4 of Table G.1 in Appendix G for the regression result when lagged extra corporate tax rates are added to the RHS of equation (4.8) with lagged dependent variables.

stock but not that of the future. To control for the effect of the overall economic growth and the business circle, the GDP growth rate, which is adjusted by CPI, is added to the RHS of equation (4.8) with lagged dependent variables. Its coefficient, however, is not statistically significant in this case.[50]

As expected, the overall extra corporate tax rate has a significantly negative effect on the corporate share of capital stock. It is, however, small in magnitude, which is consistent with the findings in the literature. A 10 percent increase in the extra corporate tax rate of a particular year only results in a 0.181 percent decrease in the current corporate share of capital stock. This means that roughly 15.8 billion dollars of capital shift out of the corporate sector in that year in response to the 10 percent increase of the extra corporate tax rate, relative to the 8,725 billion dollars of capital stock for 1997 that is used in my calculation.

Beyond this, the significance of the coefficients of lagged dependent variables indicates that the current corporate share of capital depends on previous shares. The combined effect of the corporate shares of the previous three periods is positive (1.5347 - 0.9495 + 0.2183 = 0.8035), which indicates that transaction costs are slowing down the speed of change in organizational forms in response to the extra corporate tax. Though Mackie-Mason and Gordon (1997) find no such transaction costs effect, I do find the effect when using data including more periods.[51]

---

[50]Please refer to column 5 of Table G.1 in Appendix G for the regression result when GDP growth rate is added to the RHS of equation (4.8) with lagged dependent variables.

[51]To check the effect of transaction costs, Mackie-Mason and Gordon (1997) include lagged dependent variable (the corporate share of capital stock) as independent variable in their regression. Its coefficient, however, is not statistically significant. Thus, they conclude that "no evidence of any adjustment lag" (page 493).

4.5.2 Effects on the Corporate Share of Investment

Next, estimates of equation (4.9) are presented in column 1 of Table 4.3. Then column 2 shows the results when lagged dependent variables are included in equation (4.9). Column 1 is estimated based on 97 observations from 1901 to 1997 by using OLS with AR(1), and the Durbin-Watson statistic is calculated after the correction. Column 2 is estimated based on 95 observations from 1903 to 1997 by using OLS.[52]

When I estimate equation (4.9) (column 1 of Table 4.3), I find that the coefficient on the extra corporate tax rate is -0.1087, which means the effect of the extra corporate tax rate on the corporate share of investment is larger than its effect of the capital stock. The corrected Durbin-Watson statistics is 1.8705, which shows the estimate of -0.1087 is robust. The coefficient of WWII (0.0387) is still positive and statistically significant, which confirms that the economy tends to invest proportionally more into corporations during war time than during peace time. The coefficient of TRA86 (-0.0410) is negative now and statistically significant. The sign and the significance of this coefficient verify the prediction of Fullerton (1996). After 1986, people tend to invest proportionally more in the non-corporate sector than before 1986.

After adding two lagged dependent variables (lagged one and two) to the RHS of equation (4.9) and assuming the error term η_t are standard normal (column 2 of Table 4.3), I find their coefficients are statistically significant (0.5891 and -0.2639).[53] The Durbin-Watson statistics (2.0752) is good. Thus transaction costs are also slowing down the changing speed of investment in response to the change of the extra corporate tax rate because the combined effect of the corporate shares of the previous two periods

[52] Limdep 7.03 is used in both estimations.

[53] The dependent variable lagged three periods is not statistically significant. Please refer to column 1 of Table G.2 in Appendix G for the result when the third lagged dependent variable is added to the RHS of equation (4.9).

is positive (0.5891 - 0.2639 = 0.3252). Compared to the regression of column 2 in Table 2 (the capital stock regression), however, the number of lags in this investment regression is less than that of the capital stock regression (two significant lags in the investment regression versus three in the capital stock regression). Furthermore, the magnitudes of the coefficients of lagged dependent variables in the investment regression are also smaller than those in the capital stock regression (0.5891 on the lagged one in the investment regression versus 1.5347 in the capital stock regression, and -0.2639 on the lagged two in the investment regression versus -0.9495 in the capital stock regression). These two facts indicate that the transactions costs of investment are smaller than that of the capital stock, which is consistent with the conclusion above that the effect of the extra corporate tax rate on the corporate share of investment is larger than that of the capital stock.[54]

To control for the long run effect of the tax, two lagged extra corporate tax rates (one and two periods lagged) are added to the RHS of equation (4.9) and equation (4.9) with lagged dependent variables, but none of their coefficients are statistically significant in both cases.[55] This is consistent with the previous finding that the extra corporate tax rate tends to affect the current corporate share of investment but not that of the future. To control for the effect of overall economic growth and the business cycle, the GDP growth rate is included as an independent variable in my estimations (with or without lagged dependent variables). Its coefficient, however, is still not statistically significant in both cases.[56]

---

[54]When lagged dependent variables are added to the RHS of equation (4.9), the absolute value of the coefficient on the extra corporate tax rate is smaller (changed from 0.1087 to 0.0837), but it is still much larger than 0.0181, the coefficient of the tax rate in column 2 of Table 2.

[55]Please refer to column 2 and 4 of Table G.2 in Appendix G for the regression result when lagged extra corporate tax rates are added to the RHS of equation (4.9) and equation (4.9) with lagged dependent variables.

[56]Please refer to column 3 and 5 of Table G.2 in Appendix G for the regression result when GDP growth rate is added to the RHS of equation (4.9) and equation (4.9) with lagged dependent variables.

Table 4.3: Estimated Coefficient for the Corporate Share of Investment

(standard errors in the first parentheses, and *p*-value in the second parentheses)

	Equation (4.9) 1901-1997	Equation (4.9) with lagged dep. var. 1903-1997
Intercept	286.8810 (26.3363) (0.0000)	222.1228 (32.3280) (0.0000)
Extra tax rate	-0.1087 (0.0480) (0.0234)	-0.0837 (0.0310) (0.0083)
Time	-29.4548 (2.7133) (0.0000)	-22.8137 (3.3215) (0.0000)
Time^2	0.7579 (0.0699) (0.0000)	0.5870 (0.0854) (0.0000)
WWII	0.0387 (0.0109) (0.0004)	0.0369 (0.0082) (0.0000)
TRA86	-0.0410 (0.0127) (0.0013)	-0.0415 (0.0093) (0.0000)
Lagged dep. Var. (-1)	-	0.5891 (0.1022) (0.0000)
Lagged dep. Var. (-2)	-	-0.2639 (0.0904) (0.0044)
Durbin-Watson	1.8705	2.0752
Coefficient in AR(1)	0.4515 (0.0911) (0.0000)	-

Note: The "dep. var." means dependent variable.

As expected, the overall extra corporate tax rate has a significantly negative effect on the corporate share of investment. Moreover, the magnitude of the effect is much larger than that of the extra corporate tax on the corporate share of capital. A 10 percent increase in the extra corporate tax rate of a particular year results in a one percent decrease in the current corporate share of investment, five times as big as the decrease in the corporate share of capital (ten times, if using the -0.00920 estimate of Mackie-Mason and Gordon (1997)). That means roughly 9.0 billion dollars of investment shift out of the corporate sector in that year in response to the 10 percent increase of the extra corporate tax rate, when 830 billion dollars of investment for 1997 is used in my calculation.

The difference between the effect of the extra corporate tax on the corporate share of investment and that of capital stock can be explained by lags or transaction costs. As the new capital added in the current period, investment can choose its organizational form according to the extra corporate tax rate with less transaction costs; in contrast, capital stock is the accumulated investment of many previous years, and so it encounters extra transaction costs when changing organizational form according to the extra tax rate, because that capital stock has already chosen an organizational form in previous periods. Therefore, the overall extra corporate tax rate would have had a greater effect on the corporate share of capital if no firm faced any transaction costs of switching existing capital stock from one form to the other. Table 4.4 shows comparisons of the transferred amount of capital or investment calculated by using different marginal effects of the extra corporate tax rate from Mackie-Mason and Gordon (1997), from Goolsbee (1998), and from current results. Again, 8,725 billion dollars of capital and 830 billion dollars of investment are the figures for 1997, and a ten percent increase in the overall corporate tax rate is used in these calculations. The table shows that about 95 billion dollars of capital stock would have shifted out of the corporate sector in response to the ten percent increase in the overall corporate tax rate in 1997 if capital had been as liquid as investment, while only about 16 billion dollars of capital stock actually shifted in the year (8 billion dollars if the coefficient of Mackie-Mason and Gordon (1997) is used).

Table 4.4: Comparison of the Transferred Amount of Capital and Investment

(Calculated by a 10% increase of the overall extra corporate tax rate and figures of 1997)

Unit: Billions of dollars (1997)

	Equation (4.9)	Equation (4.8) with lagged dep. var.	Mackie-Mason and Gordon	Goolsbee
Coefficient	0.1087	0.0181	0.0092	0.0304
Capital shifted	94.8,	15.8	8.0	26.5
Investment shifted	9.0	1.5	0.8	2.5

Note: The "dep. var." means dependent variable.

4.6 CONCLUSIONS AND DISCUSSIONS

As expected, the overall extra corporate tax rate has significantly negative effects on the corporate share of economic activity, which is consistent with findings in the literature. The effect of the extra corporate tax on the corporate share of capital stock, which is estimated using data with a longer time span than those used in the previous literature, stands just between two major estimates in that literature. Furthermore, the effect of the extra corporate tax on the corporate share of investment is much larger than the effect on capital. For example, a 10 percent increase in the extra corporate tax rate only results in a 0.18 percent decrease in the corporate share of capital, while it results in a one percent decrease in the corporate share of investment.

The big difference between those two effects is caused by differential transaction costs for two reasons: first, the corporate share of capital stock does show adjustment lags here, and this result differs from Mackie-Mason and Gordon (1997); second, the extra corporate tax shows a larger effect on the corporate share of investment, which has less transaction costs than capital stock. These transaction costs slow down the response of the corporate share of capital stock to the extra corporate tax rate, while they provide less impediment to new investment. For example, given a 10% increase in the extra corporate tax rate, 89 billion dollars of capital would have shifted out the corporate sector in 1997 if capital had been as liquid as investment, while about only 17 billion dollars of capital actually shifted according to my estimates.

Furthermore, the economy tends to increase the share of corporate assets during war time. The structural changes of personal and corporate income taxation caused by the Tax Reform Act of 1986 encourage investors to increase their share of non-corporate assets.

A further extension of this paper could add the non-tax factor G into the regression models. Though G cannot be observed or measured directly, many good proxies may be used instead. One is annual bankruptcy rates, which can be used to describe the effect of limited liability of corporate shareholders, while the other is the amount of trading of corporate shares in a year, which can be used to describe the effect of the mobility of corporate shares.

Chapter 5: Conclusions

Chapter 2 shows that in the one-bracket case, both the optimal income tax rate and the negative intercept (government transfer) become larger when earning inequality becomes more serious. In the two-bracket case, I find that the marginal rate of the lower bracket is greater than that of the upper bracket when the spread of the wage is relatively small, but it is larger when the spread is relatively large. Beyond that, surprisingly, I find that with a relatively large elasticity of substitution between consumption and leisure in the consumer's utility function, the two-bracket tax structure converges to the one-bracket structure when the wage spread becomes relatively quite large. Furthermore, though the optimal lower bracket rate and income threshold do not show monotonicity, the optimal upper bracket rate and government transfer are increasing with the wage spread.

Chapter 3 shows that in a general equilibrium analysis, an environmental tax in a closed city targeting polluting emissions can reduce the pollution emissions in the city's central business district (CBD), and therefore improves environmental quality of the city everywhere. This is because the tax increases the price of pollution emissions (as a factor input) so that less pollution emissions are used in the production of the composite good. In addition to the reducing of pollution, the tax reduces the wage received by labor. This is because firms try to shift the burden of the environmental tax onto labor, which is inelasticly supplied. Lower pollution as an input means a lower marginal product of labor. Overall, as the environmental tax rate rises from quite low to quite high, the tax first improves the social welfare of the city until it hits the optimal level, and then it reduces welfare as the marginal benefit falls below the marginal cost of the tax. Furthermore, for a fixed number of residents, the tax first makes the city boundary shrink but then makes it grow as residents try to balance between better housing and shorter commuting. Initial increases in the tax pull down the rent for any particular location in the city and further increases pull the rent back up. The lower bids of residents that result from lower income are offset by the higher bids that result from the pursuit of better environmental quality.

The environmental tax has two possible effects on the whole rent gradient because it does not affecting the rent of every location uniformly. In the fist case, when the tax increases from small to large, all rents first decrease, forming a downward shift of the rent gradient, then the rents close to the CBD turn back to increase while those away from the CBD still decrease, forming a clockwise pivoting, and at last, all rents increase, forming an upward shift. In the second case, all rents first decrease, forming a downward shift of the rent gradient, then the rents away from the CBD turn back to increase while those close to the CBD still decrease, forming a counter-clockwise pivoting, and at last, all rents increase, forming an upward shift.

In a quite polluted city with an inadequate environmental tax, the increase of the tax shifts down the rent gradient, decreasing the rent of land everywhere in the city. In contrast, for a very clean city with a high environmental tax, an increase of the tax shifts up the rent gradient, increasing the rent of land everywhere. In one type of theoretically defined city, as the environmental tax increases from quite low to quite high, it first improves the social welfare of the city, shrinks the size of the city , and cuts the rent of any particular location in the city; and then it reduces the social welfare, expands the city, and raises the rent anywhere. In another type of theoretically defined city, furthermore, a Pareto improvement is possible, which means that the tax may increase both the social welfare of the city residents and the gradient of rent received by landlords.

Chapter 4 shows that the overall extra corporate tax rate has significantly negative effects on the corporate share of economic activity, which is consistent with findings in the literature. The effect of the extra corporate tax on the corporate share of capital stock, which is estimated by data with a longer time span than those used in the literature, stands just between two major estimates in the literature. Furthermore, the effect of the extra corporate tax on the corporate share of investment is much larger than the effect on capital. For example, a 10 percent increase in the extra corporate tax rate only results in a 0.18 percent decrease in the corporate share of capital, while it results in a one percent decrease in the corporate share of investment. These results are consistent with transaction costs in two respects: first, the corporate share of capital stock does show adjustment lags, which is different from what Mackie-Mason and Gordon (1997) find; second, the extra corporate tax shows larger effect on the corporate share of investment than on the corporate share of capital stock. Investment can react more easily than capital stock. If capital stock could be added or subtracted with less transaction costs, then the corporate share of capital would have responded to the extra corporate tax as sensitively as the corporate share of investment. For example, suppose my estimates were used to predict effects of a 10% increase in the extra corporate tax rate. If capital had been as liquid as investment, then 89 billion dollars of capital would have shifted out the corporate sector in 1997 instead of only 17 billion dollars of capital actually shifted.

Furthermore, estimates here account for the fact that the economy tends to increase the share of corporate assets during war time. Finally, the structural changes of tax system caused by the Tax Reform Act of 1986 encouraged people to increase the share of non-corporate assets. These structural changes may affect the way that a change in corporate tax rate would affect the share of investment going into the corporate sector.

Appendices

APPENDIX A: ADDITIONAL FIGURES FOR CHAPTER 1

Figure A.1: The Optimal Two-Bracket Tax Rates (σ = 0.3, 0.5, 0.6, and 0.7)

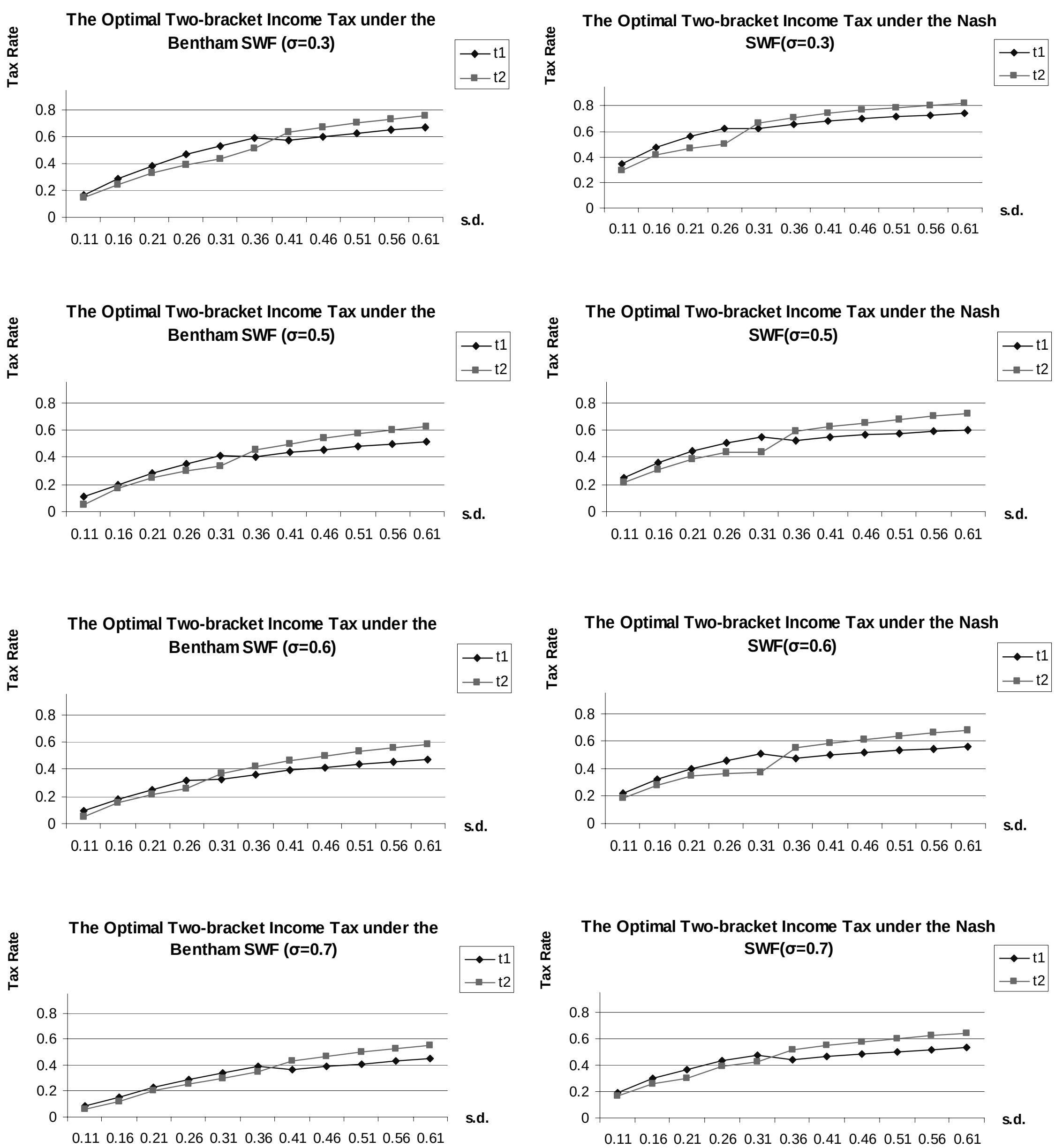
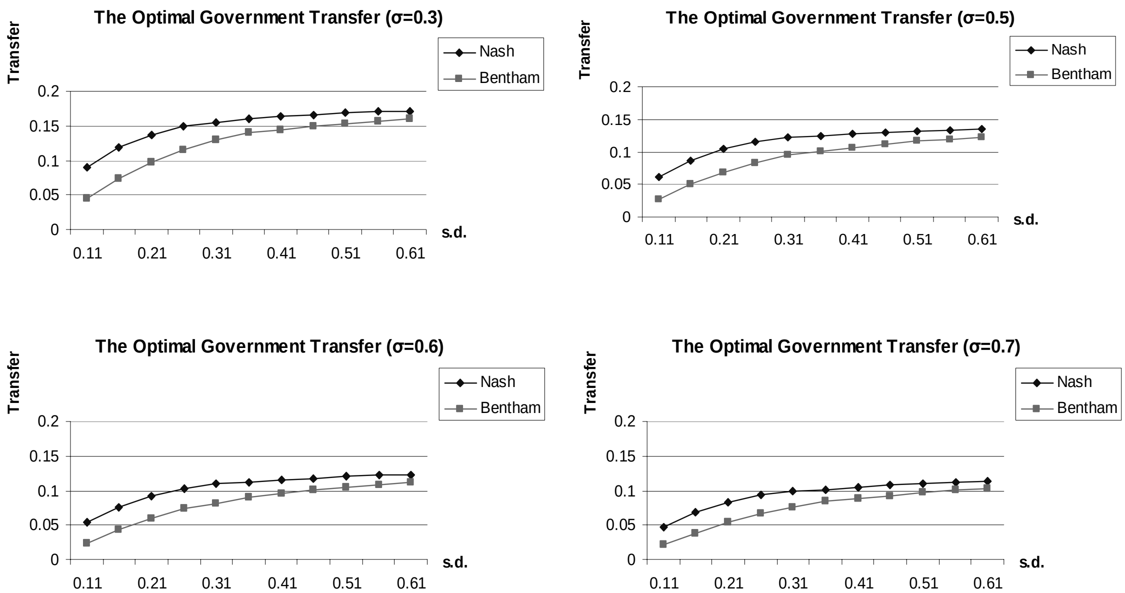


Figure A.2: The Optimal Government Transfer (σ = 0.3, 0.5, 0.6, and 0.7, two-bracket)



APPENDIX B: ADDITIONAL TABLES FOR CHAPTER 1

Table B.1: The Optimal Two-Bracket Income Tax with Larger σ under the Bentham SWF.

s.d. / σ	0.1609			0.2109			0.2609			0.3109			0.3609		
	T_1	T_2	$\hat{Y}$	t_1	t_2	$\hat{Y}$	t_1	t_2	$\hat{Y}$	t_1	t_2	$\hat{Y}$	t_1	t_2	$\hat{Y}$
0.9	0.142	0.119	0.397	0.196	0.171	0.492	0.224	0.252	0.044	0.272	0.304	0.051	0.308	0.350	0.053
1.0	0.134	0.118	0.287	0.187	0.164	0.498	0.244	0.220	0.539	0.267	0.293	0.045	0.336	0.336	0.041
1.1	0.128	0.113	0.316	0.181	0.167	0.387	0.235	0.209	0.632	0.286	0.251	0.830	0.000	0.328	0.000
s.d. / σ	0.4109			0.4609			0.5109			0.5509			0.6109		
	T_1	T_2	$\hat{Y}$	t_1	t_2	$\hat{Y}$	t_1	t_2	$\hat{Y}$	t_1	t_2	$\hat{Y}$	t_1	t_2	$\hat{Y}$
0.9	0.338	0.390	0.056	0.000	0.417	0.000	0.000	0.444	0.000	0.000	0.467	0.000	0.000	0.487	0.000
1.0	0.000	0.374	0.000	0.000	0.406	0.000	0.000	0.434	0.000	0.000	0.458	0.000	0.000	0.478	0.000
1.1	0.000	0.366	0.000	0.000	0.399	0.000	0.000	0.428	0.000	0.000	0.453	0.000	0.000	0.475	0.000

Table B.2: The Optimal Two-Bracket Income Tax with Larger σ under the Nash SWF.

The Optimal Two-bracket Income Tax with Large σ under the Nash SWF

s.d. / σ	0.1609			0.2109			0.2609			0.3109			0.3609		
	T_1	T_2	$\hat{Y}$	t_1	t_2	$\hat{Y}$	t_1	t_2	$\hat{Y}$	t_1	t_2	$\hat{Y}$	t_1	t_2	$\hat{Y}$
0.9	0.251	0.223	0.287	0.323	0.287	0.375	0.380	0.344	0.471	0.373	0.426	0.055	0.403	0.466	0.059
1.0	0.234	0.208	0.307	0.303	0.257	0.514	0.323	0.359	0.047	0.361	0.407	0.053	0.391	0.447	0.058
1.1	0.221	0.201	0.273	0.288	0.256	0.461	0.344	0.302	0.659	0.351	0.391	0.052	0.000	0.421	0.000
s.d. / σ	**0.4109**			**0.4609**			**0.5109**			**0.5609**			**0.6109**		
	T_1	t_2	$\hat{Y}$	t_1	t_2	$\hat{Y}$	t_1	t_2	$\hat{Y}$	t_1	t_2	$\hat{Y}$	t_1	t_2	$\hat{Y}$
0.9	0.429	0.500	0.062	0.451	0.528	0.065	0.470	0.552	0.069	0.487	0.573	0.072	0.503	0.591	0.075
1.0	0.417	0.480	0.063	0.000	0.489	0.000	0.000	0.508	0.000	0.000	0.524	0.000	0.000	0.538	0.000
1.1	0.000	0.450	0.000	0.000	0.473	0.000	0.000	0.492	0.000	0.000	0.508	0.000	0.000	0.522	0.000

APPENDIX C: THE MAXIMIZAITON OF RENT GRADIENT SUBJECT TO UTILITY CONSTRAINT IN CHAPTER 3

(3.9) **max** $_{X,H}$ $R(D) = [1/H(D)]\,[Y - C(D) - T(D)]$

(3.10) **s.t.** $V = U\,[C(D), H(D), Q(D, E)]$

This is equivalent to:

(3.46) **max** $_{X,H}$ $L = [1/H(D)][Y - C(D) - T(D)] + \mu\{V - U\,[C(D), H(D), Q(D, E)]\}$

where μ is the Lagrange multiplier. The first order conditions are:

(3.47a) $\partial L/\partial C = -\, 1/H - \mu\, U_C = 0$

(3.47b) $\partial L/\partial H = -\, (1/H^2)[Y - C - T(D)] - \mu\, U_H = 0$

(3.47c) $\partial L/\partial \mu = V - U[C, H, Q(D, E)] = 0$

From (47a):

(3.48) $\mu = -1/(H\, U_C) < 0$

I have three unknowns, μ, H, and X, while I have three equations (3.47a)-(3.47c). So, the maximization (3.46) is solvable. Since I assume the second order conditions are satisfied, the solution of (3.46) is the solution of the maximization of (3.9) subject to the constraint (3.10). So, I have:

(3.11a) $C = C(D, Y, E, V)$

(3.11b) $H = H(D, Y, E, V)$

(3.11c) $\mu = \mu(D, Y, E, V)$

Substituting C and H into (3.9), I get the family of the rent gradients:

(3.12) $R = R(D, Y, E, V)$

By the envelope theorem, $\partial R/\partial \sigma = \partial L/\partial \sigma$, where σ is a particular parameter in the first order conditions. So, from (3.42), I have:

(3.23a) $\partial R/\partial D = -[1/H(D)]\, T_D - \mu(D)\, U_Q\, Q_D >< 0$

(3.23b) $\partial R/\partial Y = 1/H(D) > 0$

(3.23c) $\partial R/\partial E = -\mu(D)\ U_Q\ Q_E < 0$

(3.23d) $\partial R/\partial V = \mu(D) < 0$

From (3.47a) and (3.47b), I have:

(3.49) $m - [Y - C - T(D)]/H = 0$

Where $m \equiv U_H/U_C$, which is first denoted by Wheaton (1974). As approved by Wheaton (1974), as long as the utility is strictly quasi-concave and both the composite good and the land space have positive income effect, m has properties such as $\partial m/\partial C > 0$ and $\partial m/\partial H < 0$. Totally differentiating (3.47c) and (3.49):

(3.50) $U_C\ dC + U_C\ dH = -\ (U_Q\ Q_D)\ dD - (U_Q\ Q_E)\ dE + dV$

(3.51) $$(\frac{\partial m}{\partial C} + \frac{1}{H})dC + (\frac{\partial m}{\partial C} + \frac{m}{H})dH = -[\frac{T_D}{H} + \frac{\partial m}{\partial Q}Q_D]dD + \frac{1}{H}dY - (\frac{\partial m}{\partial Q}Q_E)dE$$

Let dD, dE, and dV be zero, I have the partial derivations of C and H with respect to income Y:

(3.22a) $\partial C(D)/\partial Y = -m/\{H\ [\partial m/\partial H - (\partial m/\partial C)\ m]\} > 0$

(3.22b) $\partial H(D)/\partial Y = 1/\{H\ [\partial m/\partial H - (\partial m/\partial C)\ m]\} < 0$

Let dD, dY, and dV be zero and assume $\partial Rmu/\partial Q = 0$, I have the partial derivations of C and H with respect to income E:

(3.22c) $\partial C(D)/\partial E = -U_Q Q_E\ /U_C - m\ [\partial H(D)/\partial V] >< 0$

(3.22d) $\partial H(D)/\partial E = [-U_Q Q_E\ m\ (\partial m/\partial C+1/H)]/[U_H(\partial m/\partial C-\partial m/\partial H)] > 0$

Let dD, dY, and dE be zero, I have the partial derivations of C and H with respect to income V:

(3.22e) $\partial C(D)/\partial V = 1/U_C - m\ [\partial H(D)/\partial V] >< 0$

(3.22f) $\partial H(D)/\partial V = [m\ (\partial m/\partial C+1/H)]/[U_H(\partial m/\partial C-\partial m/\partial H)] > 0$

From the utility function (3.3):

(3.52) $\partial U_C/\partial D = U_{CC}\ (\partial C/\partial D) + U_{CH}\ (\partial H/\partial D) + U_{CQ}\ (\partial Q/\partial D)$

where U_{CQ} is assumed to be non-negative. After taking derivative of (3.48) with respect to D:

(3.53) $\partial\mu(D)/\partial D = [1/H(D)][1/(U_C)^2][\partial U_C/\partial D] + [1/U_C][1/H(D)^2][\partial H(D)/\partial D] > 0$

APPENDIX D: USEFUL PATIAL DERIVATIVES FOR CHAPTER 3

After total differentiating (3.4b), (3.6), (3.8), and (3.16), I have:

(3.54) $dY = d\omega + dg$

(3.55) $dg = (-tE/N^2)\, dN + (E/N)\, dt + (t/N)\, dE$

(3.56) $dN = dL$

(3.57) $d\omega = X_{LL}\, dL + X_{LE}\, dE$

(3.58) $dt = X_{EL}\, dL + X_{EE}\, dE$

where $X_{LL} = \partial X_L/\partial L$, $X_{EE}= \partial X_E/\partial E,$ and $X_{LE} = X_{EL} = \partial X_L/\partial E = \partial X_E/\partial L$. Substituting (3.56) into (3.58):

(3.59) $dE = -(X_{EL}/X_{EE})\, dN + (1/X_{EE})\, dt$

Substituting (3.59) into (3.57):

(3.60) $d\omega = \{[X_{LL}\, X_{EE} - (X_{LE})^2]/X_{EE}\}\, dN + (X_{LE}/X_{EE})\, dt$

Let $dt = 0$ first and then $dN = 0$, I have the partial derivatives of E and ω with respect to N and t:

(3.24a) $\partial E/\partial N = -X_{EL}/X_{EE} > 0$

(3.24b) $\partial E/\partial t = 1/X_{EE} < 0$

(3.24c) $\partial\omega/\partial N = [X_{LL}X_{EE} - (X_{LE})^2]/X_{EE} \leq 0$

(3.24d) $\partial\omega/\partial t = X_{LE}/X_{EE} < 0$

Substituting (3.59) into (3.55):

(3.61) $dg = -[(t/N)/(X_{EL}/X_{EE})+ tE/N^2]\, dN + [(t/N)/X_{EE} + E/N]\, dt$

Let $dt = 0$ first and then $dN = 0$, I have the partial derivatives of g with respect to N and t:

(3.24e) $\partial g/\partial N = -(t/N)/(X_{EL}/X_{EE}) - Et/(N^2) >< 0$

(3.24f) $\partial g/\partial t = (t/N)/X_{EE} + E/N >< 0$

Substituting (3.60) and (3.61) into (3.54), and let $dt = 0$ first and then $dN = 0$, I have the partial derivatives of Y with respect to N and t:

(3.24g) $\partial Y/\partial N = \partial \omega/\partial N + \partial g/\partial N = [X_{LL}X_{EE} - (X_{LE})^2]/X_{EE} - (t/N)/(X_{EL}/X_{EE}) - Et/(N^2) >< 0$

(3.24h) $\partial Y/\partial t = \partial \omega/\partial t + \partial g/\partial t = X_{LE}/X_{EE} + (t/N)/X_{EE} + E/N >< 0$

APPENDIX E: TOTALLY DIFFERENTIATION OF EQUATIONS (3.20) AND (3.21) FOR CHAPTER 3

(3.20) $R^A = R(b, Y(N, t), E(N, t), V)$

(3.21) $N = \int_0^b \{1/H[D, Y(N, t), E(N, t), V]\}\, d(D)$

after totally differentiating (3.20) and (3.21), I have:

(3.27a) $$-\frac{\partial R(b)}{\partial b}db - \frac{\partial R(b)}{\partial V}dV = [\frac{\partial R(b)}{\partial Y}\frac{\partial Y}{\partial N} + \frac{\partial R(b)}{\partial E}\frac{\partial E}{\partial N}]dN + [\frac{\partial R(b)}{\partial Y}\frac{\partial Y}{\partial t} + \frac{\partial R(b)}{\partial E}\frac{\partial E}{\partial t}]dt$$

(3.27b) $$\frac{1}{H(b)}db - (\int_0^b -\frac{1}{H^2}\frac{\partial H}{\partial V}dD)dV = [1 + (\int_0^b \frac{1}{H^2}\frac{\partial H}{\partial Y}dD)\frac{\partial Y}{\partial N} + (\int_0^b \frac{1}{H^2}\frac{\partial H}{\partial E}dD)\frac{\partial E}{\partial N}]dN$$
$$+ [(\int_0^b \frac{1}{H^2}\frac{\partial H}{\partial Y}dD)\frac{\partial Y}{\partial t} + (\int_0^b \frac{1}{H^2}\frac{\partial H}{\partial E}dD)\frac{\partial E}{\partial t}]dt$$

In a closed city, $dN = 0$, so:

(3.31a) $$-\frac{\partial R(b)}{\partial b}db - \frac{\partial R(b)}{\partial V}dV = [\frac{\partial R(b)}{\partial Y}\frac{\partial Y}{\partial t} + \frac{\partial R(b)}{\partial E}\frac{\partial E}{\partial t}]dt$$

(3.31b) $$\frac{1}{H(b)}db - (\int_0^b -\frac{1}{H^2}\frac{\partial H}{\partial V}dD)dV = [(\int_0^b \frac{1}{H^2}\frac{\partial H}{\partial Y}dD)\frac{\partial Y}{\partial t} + (\int_0^b \frac{1}{H^2}\frac{\partial H}{\partial E}dD)\frac{\partial E}{\partial t}]dt$$

Divided by dt, (3.31) becomes:

(3.62a) $-[\partial R(b)/\partial b]\,(db/dt) - (\partial R(b)/\partial V)\,(dV/dt) = [\partial R(b)/\partial Y](\partial Y/\partial t) + [\partial R(b)/\partial E](\partial E/\partial t)$

(3.62b) $$\frac{1}{H(b)}\frac{db}{dt}-\left(\int_0^b -\frac{1}{H^2}\frac{\partial H}{\partial V}dD\right)\frac{dV}{dt}=\left(\int_0^b \frac{1}{H^2}\frac{\partial H}{\partial Y}dD\right)\frac{\partial Y}{\partial t}+\left(\int_0^b \frac{1}{H^2}\frac{\partial H}{\partial E}dD\right)\frac{\partial E}{\partial t}$$

By the Cramer's rule:

(3.35) $dV/dt = (\Psi/|A|)\,(\partial Y/\partial t) - (\Omega/|A|)\,(\partial E/\partial t) >< 0$

where $|A| = [\partial R(b)/\partial b][\int_0^b (1/H^2)\,(\partial H/\partial V)\,dD] + [1/H(b)][\partial R(b)/\partial V]$ is the determinant of the coefficient matrix of dV and db in equation (31), Ψ equals $-\{[\partial R(b)/\partial b][\int_0^b (1/H^2)\,(\partial H/\partial Y)\,dD] + [1/H(b)]\,[\partial R(b)/\partial Y]\}$, and Ω equals $\{[\partial R(b)/\partial b][\int_0^b (1/H^2)\,(\partial H/\partial E)\,dD] + [1/H(b)]\,[\partial R(b)/\partial E]\}$. It is easy to see $|A| < 0$, $\Psi < 0$, and $\Omega < 0$ by using equation (3.22) and (3.23). From (3.62a):

(3.63) $$\frac{db}{dt}=-\left[\frac{\partial R(b)}{\partial Y}\Big/\frac{\partial R(b)}{\partial b}\right]\frac{\partial Y}{\partial t}-\left[\frac{\partial R(b)}{\partial E}\Big/\frac{\partial R(b)}{\partial b}\right]\frac{\partial E}{\partial t}-\left[\frac{\partial R(b)}{\partial V}\Big/\frac{\partial R(b)}{\partial b}\right]\frac{dV}{dt}$$

Substituting (3.35) into (3.63):

(3.36) $$\frac{db}{dt}=\left[-\frac{\partial R(b)/\partial Y}{\partial R(b)/\partial b}-\left(\frac{\Psi}{A}\right)\left(\frac{\partial R(b)/\partial V}{\partial R(b)/\partial b}\right)\right]\frac{\partial Y}{\partial t}+\left[-\frac{\partial R(b)/\partial E}{\partial R(b)/\partial b}+\left(\frac{\Omega}{A}\right)\left(\frac{\partial R(b)/\partial V}{\partial R(b)/\partial b}\right)\right]\frac{\partial E}{\partial t}$$

APPENDIX F: PROPERTIES OF THE EQUILIBRIUM UTILITY, CITY LIMIT, AND RENT AT LOCATION ZERO IN CHAPTER 3

F.1 The Equilibrium Utility

This is to prove the properties of dV/dt. When consider the properties of $V(t)$, I have:

(3.64a) $V(0) = \underline{U}$

(3.64b) $dV(0)/dt > 0$

(3.64c) $V(\infty) = \bar{U}$

(3.64d) $dV(\infty)/dt < 0$

where $\underline{U}$ and $\bar{U}$ are positive finite numbers.

When the environmental tax approaches zero, firms want to use infinite amount of emission pollutions in the production of the composite good, but the labor supply is always fixed. So the amount of the output of X converges to a positive finite number. That implies individuals' consumption of the composite good converges too. Since the environmental tax rate is zero, the total cost spent by firms on the emissions is zero, and the government transfers to individuals are therefore zero too. By the zero profit condition (3.2), the wage of labor converges to a finite positive number because X converges while tE is zero. So, the total income of an individual converges to a finite number by equation (3.4b). By the individual's budget constraint (3.4a), the expenditure on land space converges too, and therefore the individuals' consumption of land space converges to a finite number. Since the pollution emissions explosively diverge to infinite, the environmental quality of any particular location in the city either diverge to negative infinite or converge to a small finite number (because Q can be defined in either way). Since both C and H are convergent, and Q either converges to a finite number or diverges to negative infinite, $U(C, H, Q)$ is convergent. So (3.64a) is proved.

Also, when the tax rate increases from zero to a bit, the income of individuals decreases, while the pollution emissions decrease substantially, and therefore environmental quality is improved substantially. Normally, the marginal gains from the improvement of environment quality are larger than the marginal losses from the decrease of income because the utility function strictly quasi-concave. Thus, (3.64b) is believable.

When the price of the pollution emissions (the environmental tax) approaches infinite, firms are not able to use any finite amount of emission pollutions in the production. At the same time, labor supply is still fixed. So, the amount of the output of X converges to zero. That means individuals' consumption of the composite good also converges to zero. Since firms are not using any pollution emissions, the tax revenue of the government converges to zero, and therefore the government transfer also converges to zero by the balanced government budget constraint (3.6). By the zero profit condition (3.2), the wage converges to zero because both X and tE converge to zero. Thus, income of individuals converges to zero, which means their consumption of land space converges to zero. Though, pollution emissions are eliminated, the environmental quality still converges to a finite number naturally, no matter how big the number is. Since both C and H converge to zero, and Q converges to a finite number, $U(C, H, Q)$ is convergent. So, (3.64c) is proved.

Also, when the tax rate decreases from infinite to a finite number, the pollution emission increases, while the income of individuals also increases because firms are able to produce and therefore able to give wage to individuals. At the same time, the government can collect revenue and transfer to individuals. Normally, the marginal losses from the reduction of the environment quality are less than the marginal gains from the increase of income because the utility function is strictly quasi-concave. Thus, (3.64d) is also believable.

By (3.35) and (3.64), an optimal environmental tax rate t^* exists such that the equilibrium utility V is maximized by the optimal rate. Since I am using differentiable functions in this model, I can assume that the equilibrium utility V is strictly concave with respect to the environmental tax rate.

F.2 The City Limit

This is to prove the properties of db/dt. As shown in Figure F.1, given the strict concavity of V with respect to t, dV/dt is strictly decreasing with t and crosses the horizontal axis at t^*, where the derivative is zero. Thus, curve $F(t)$ = -

$[\partial R(b)/\partial V]/[\partial R(b)/\partial b]$ (dV/dt) is strictly increasing with t and also crosses the horizontal axis at t^* because $-[\partial R(b)/\partial V]/[\partial R(b)/\partial b]$ is negative definite. Since:

(3.65a) $-[\partial R(b)/\partial Y]/[\partial R(b)/\partial b]$ $(\partial Y/\partial t)$

(3.65b) $-[\partial R(b)/\partial E]/[\partial R(b)/\partial b]$ $(\partial E/\partial t)$

equation (3.63) implies that db/dt is a curve transformed from $F(t)$, which is first shifted down by $[\partial R(b)/\partial Y]/[\partial R(b)/\partial b]$ $(\partial Y/\partial t)$ and then shifted up by $-[\partial R(b)/\partial E]/[\partial R(b)/\partial b]$ $(\partial E/\partial t)$ to form db/dt. Therefore, as described in Figure F.1, db/dt crosses the horizontal axis at least once. Thus, a particular environmental tax rate t_b^* exists such that the city limit b is minimized there.

Figure F.1 and F.2: Properties of *db/dt* and *dR*(0)/*dt*.

Figure F.1

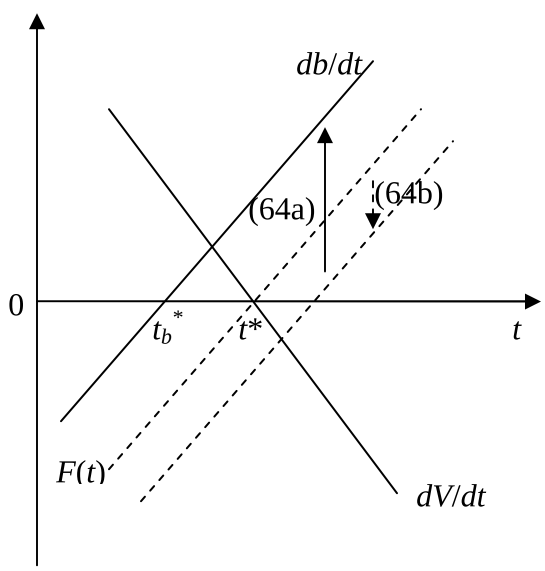
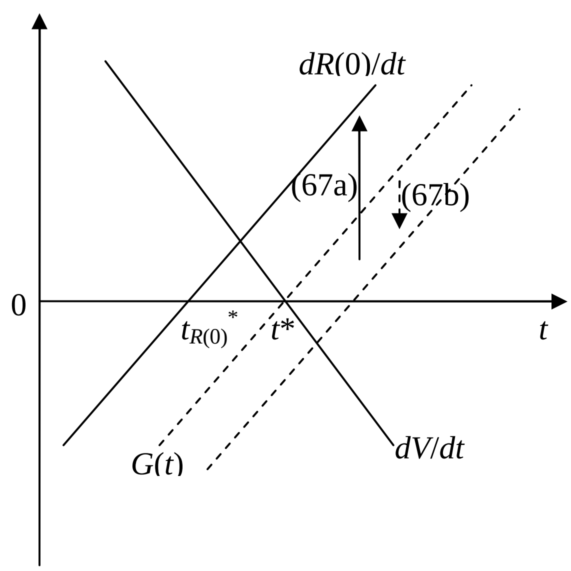


Figure F.2



F.3 The Rent at Location Zero

This is to prove the properties of $dR(0)/dt$. From (3.32):

(3.66) $dR(D)/dt = (\partial R/\partial Y)(\partial Y/\partial t) + (\partial R/\partial E)(\partial E/\partial t) + (\partial R/\partial V)(dV/dt)$

Let $D = 0$:

(3.67) $dR(D)/dt = (\partial R(0)/\partial Y)(\partial Y/\partial t) + (\partial R(0)/\partial E)(\partial E/\partial t) + (\partial R(0)/\partial V)(dV/dt)$

Given dV/dt is strictly decreasing with t and crosses the horizontal axis at t^*, curve $G(t) = [\partial R(0)/\partial V]/(dV/dt)$ is strictly increasing with t and also crosses the horizontal axis at t^* because $[\partial R(0)/\partial V]$ is negative definite. Since:

(3.68a) $(\partial R(0)/\partial Y)(\partial Y/\partial t) < 0$

(3.68b) $(\partial R(0)/\partial E)(\partial E/\partial t) > 0$

equation (3.67) implies that $dR(0)/dt$ is a curve transformed from $G(t)$, which is first shifted down by $-[\partial R(0)/\partial V]/(\partial Y/\partial t)$ and then shifted up by $[\partial R(0)/\partial E]/(\partial E/\partial t)$ to form $dR(0)/dt$. Therefore, as described in Figure F.2, $dR(0)/dt$ crosses the horizontal axis at least once. Thus, a particular environmental tax rate $t_{R(0)}{}^*$ exists such that the rent of location zero $R(0)$ is minimized there.

F.4 The Rent of Locations other than Location Zero

This is to prove the properties of $dR(D)/dt$. By replacing location 0 in section F.3 by any particular location D, $dR(D)/dt$ can be proved that a particular environmental tax rate $t_{R(D)}{}^*$ exists such that the rent of location D, $R(D)$, is minimized there.

Substituting (3.23b) – (3.23d) into (3.65):

(3.69) $\Lambda(D) = [1/H(D)]\,(\partial Y/\partial t) - \mu(D)\, U_Q\, Q_E\, (\partial E/\partial t) + \mu(D)\,(dV/dt)$

where $\Lambda(D) \equiv dR(D)/dt$. Since $\partial Y/\partial t$, $\partial E/\partial t$, and dV/dt are not functions of D, I can get the first order derivative of $\Lambda(D)$ with respect to D from (3.69):

$$(3.70)\quad \frac{d\Lambda(D)}{dD} = \{[-\frac{\partial H(D)}{\partial D}\Big/ H^2(D)] + \frac{\partial \mu(D)}{\partial D}\frac{\Psi}{|A|}\}\frac{\partial Y}{\partial t}$$
$$-[\frac{\partial \mu(D)}{\partial D}U_Q Q_E + \mu Q_E U_{QQ} Q_D + \mu U_Q Q_{ED} - \frac{\partial \mu(D)}{\partial D}\frac{\Omega}{|A|}]\frac{\partial E}{\partial t}$$

which means the sign of $d\Lambda(D)/dD$ is generally undetermined.[57] So, the effect a given change of the environmental tax on the rent can be strictly decreasing, increasing and sometimes decreasing and sometimes increasing along with the distance D.

F.5 Additional Figures for Chapter 3

Figure F.3: The Combination of Figure 3.4B and Figure 3.2B

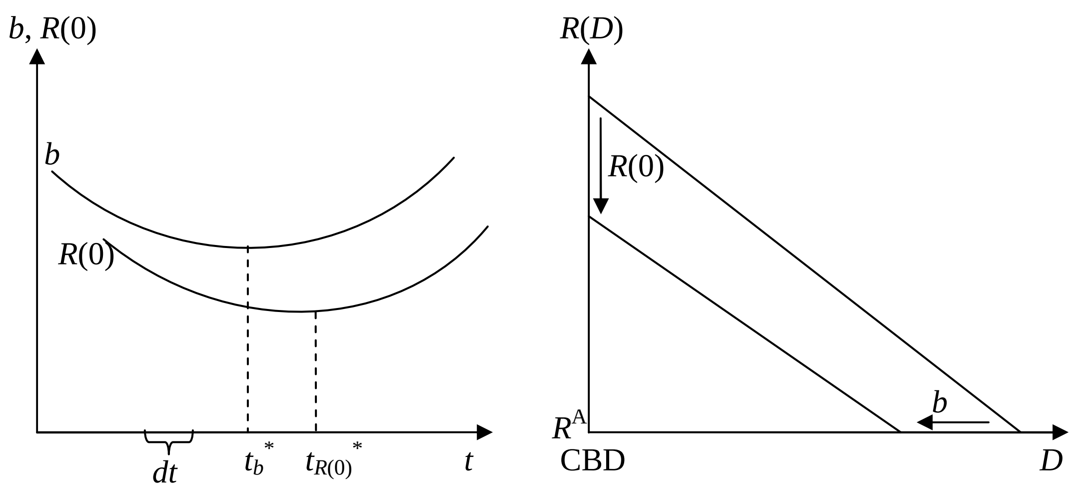
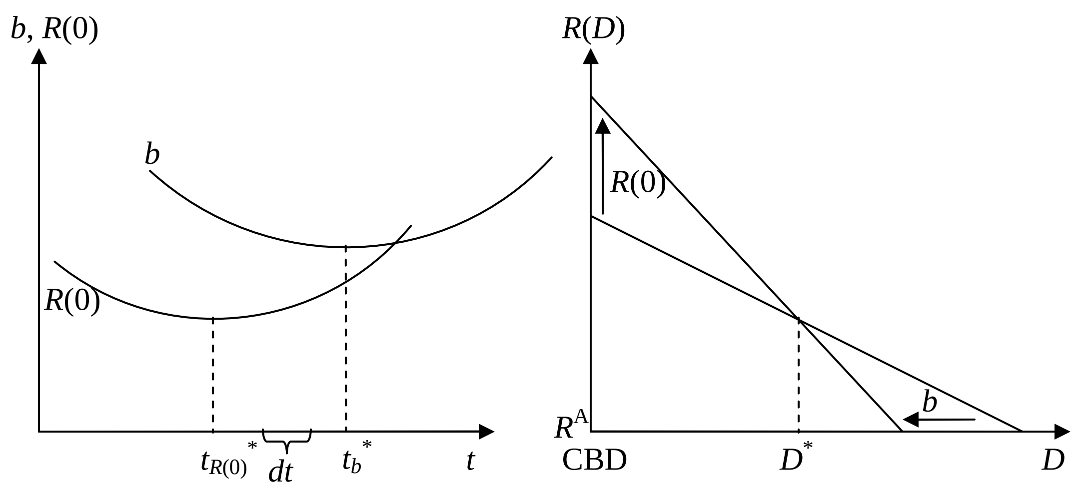


Figure F.4: The Combination of Figure 3.4C and Figure 3.2C



[57]Even $\partial\mu(D)/\partial D$ is assumed to be greater than zero (if $U_{CQ} \geq 0$) and Q_{ED} is assumed to be zero, the sign is still undetermined.

Figure F.5: The Combination of Figure 3.4D and Figure 3.2D

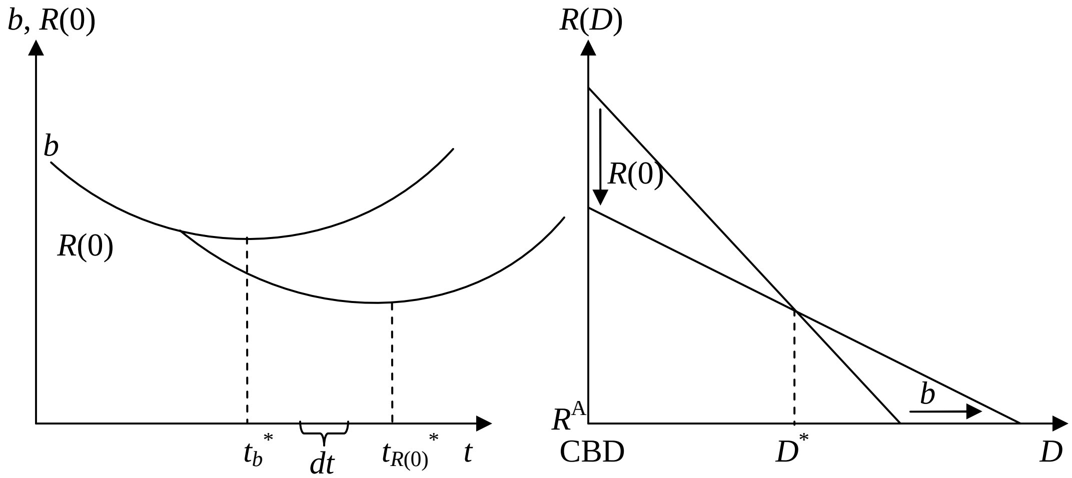
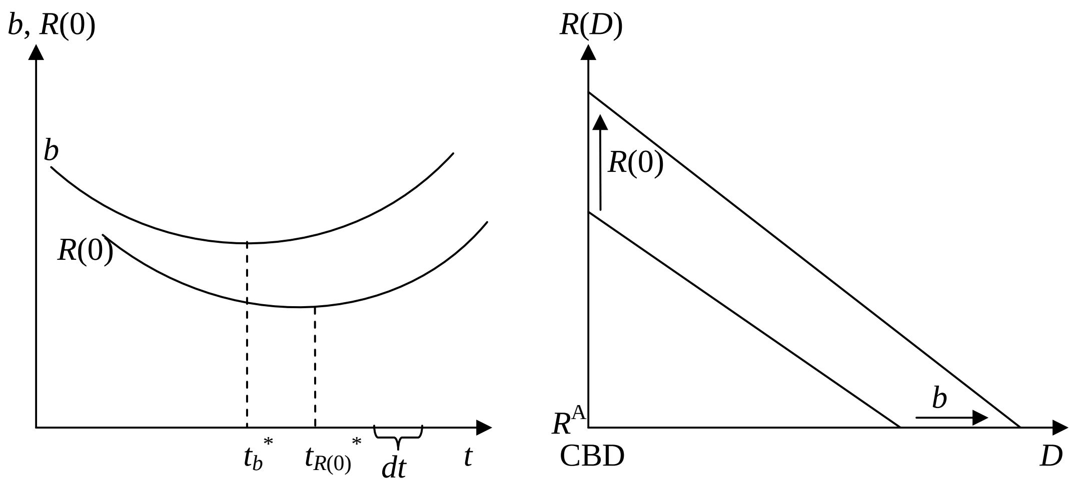


Figure F.6: The Combination of Figure 3.4A and Figure 3.2A



APPENDIX G: ADDITIONAL TABLES FOR CHAPTER 4

Table G.1: Additional Coefficients for the Corporate Share of Capital Stock

(with insignificant estimates bold and standard errors in the parentheses)

	Eq. (4.8) 1962-1986 (for FN 45)	Eq. (4.8) with four lagged dep. var. 1929-1997 (for FN 46)	Eq. (4.8) w/ lagged dep. var. but w/o TRA86 1928-1997 (for FN 47)	Eq. (4.8) w/ lagged dep. var. & tax rates 1962-1986 (for FN 49)	Eq. (4.8) w/ lagged dep. var. & GDP growth rate 1928-1997 (for FN 50)
Intercept	**205.4999** (128.8472)	29.6942 (5.7328)	25.0474 (5.0360)	29.2560 (7.2267)	28.4437 (5.8300)
Extra tax rate	**-0.0191** (0.0200)	-0.0155 (0.0056)	-0.0164 (0.0054)	-0.0189 (0.0084)	-0.0179 (0.0055)
Time	**-21.0078** (13.2142)	-3.0377 (0.5897)	-2.5602 (0.5160)	-2.9978 (0.7426)	-2.9128 (0.5997)
$Time^2$	**0.5389** (0.3388)	0.0781 (0.0152)	0.0658 (0.0133)	0.7718 (0.0191)	-0.0750 (0.0155)
GDP growth rate	-	-	-	-	**-0.0049** (0.0044)
WWII	-	0.0031 (0.0010)	0.0029 (0.0010)	0.0030 (0.0010)	0.0034 (0.0010)
TRA86	-	**-0.0014** (0.0012)	-	**-0.0017** (0.0012)	**-0.0016** (0.0012)
Lagged dep. Var. (-1)	-	1.5678 (0.1196)	1.5504 (0.1176)	1.5379 (0.1210)	1.5255 (0.1171)
Lagged dep. Var. (-2)	-	-1.1252 (0.2211)	-0.9394 (0.1904)	-0.9532 (0.1931)	-0.9078 (0.1920)
Lagged dep. Var. (-3)	-	0.5262 (0.2196)	0.1908 (0.1082)	0.2194 (0.1110)	0.1871 (0.1122)
Lagged dep. Var. (-4)	-	**-0.1827** (0.1122)	-	-	-
Lagged tax rate (-1)	-	-	-	**0.0020** (0.0112)	-
Lagged tax rate (-2)	-	-	-	**-0.0011** (0.0090)	-
Durbin-Watson	0.9293	1.8400	1.8120	1.8324	1.8064
Coefficient in AR(1)	0.8418 (0.1102)	-	-	-	-

Note: The "Eq." means equation, "dep. var." means dependent variable, and the "non-corp." means non-corporate.

Table G.2: Additional Coefficients for the Corporate Share of Investment

(with insignificant estimates bold and standard errors in the parentheses)

	Eq. (4.9) with three lagged dep. var. 1904-1997 (for FN 53)	Eq. (4.9) w/ lagged tax rates 1904-1997 (for FN 55)	Eq. (4.9) w/ GDP growth rate 1904-1997 (for FN 56)	Eq. (4.9) w/ lagged dep. var. & tax rates 1904-1997 (for FN 55)	Eq.(4.9) w/ lagged dep. var. & GDP growth rate 1904-1997 (for FN 56)
Intercept	201.0040 (34.6936)	312.0064 (29.6007)	307.8137 (26.7249)	209.7462 (36.3237)	212.7302 (33.6565)
Extra tax rate	-0.0786 (0.0310)	**-0.0854** (0.0592)	-0.1182 (0.0463)	**-0.0830** (0.0538)	-0.0810 (0.0312)
Time	-20.6489 (3.5636)	-32.0368 (3.0480)	-31.6050 (2.7518)	-21.5427 (3.7318)	-21.8499 (3.4577)
$Time^2$	0.5314 (0.0917)	0.8242 (0.0785)	0.8131 (0.0708)	0.5543 (0.0960)	0.5622 (0.0889)
GDP growth rate	-	-	**0.0069** (0.0280)	-	**0.0296** (0.0251)
WWII	0.0366 (0.0082)	0.0412 (0.0106)	0.0417 (0.0108)	0.0361 (0.0084)	0.0331 (0.0086)
TRA86	-0.0379 (0.0095)	0.0467 (0.0125)	-0.0478 (0.0122)	-0.0400 (0.0096)	-0.0392 (0.0095)
Lagged dep. Var. (-1)	0.6361 (0.1054)	-	-	0.6000 (0.1059)	0.6121 (0.1039)
Lagged dep. Var. (-2)	-0.3835 (0.1196)	-	-	-0.2532 (0.0933)	-0.2677 (0.0905)
Lagged dep. Var. (-3)	**0.1404** (0.0923)	-	-	-	-
Lagged tax rate (-1)	-	**-0.0647** (0.0667)	-	**-0.0198** (0.0750)	-
Lagged tax rate (-2)	-	**0.0167** (0.0599)	-	**0.0308** (0.0556)	-
Durbin-Watson	2.0521	1.8610	1.8658	2.0955	2.1127
Coefficient in AR(1)	-	0.4140 (0.000)	0.3977 (0.0000)	-	-

Note: The "Eq." means equation, "dep. var." means dependent variable, and the "non-corp." means non-corporate.

References

Alonso, W. (1964), *Location and Land Use*, Harvard University Press, Cambridge, MA.

Auerbach, A. (1987), "Tax Reform as Political Choice", *The Journal of Economic Perspectives* 1:29-35.

Auten, G. (1999), "Capital Gains Taxation: From *the Encyclopedia of Taxation and Tax Policy*", http://www.urban.org/GeraldAuten, Urban Institute, Washington D.C.

Backmann, M. J. (1957), "On the Distribution of Rent and Residential Density in Cities", Paper Presented at the Inter-Departmental Seminar on Mathematical Applications in the Social Science, Yale University.

Ballard, C., J. Shoven, and J. Whalley (1985), "The Total Welfare Cost of the United States Tax System: A General Equilibrium Approach", *National Tax Journal* 38: 125-140.

Baumol, W. J. and W. F. Oates (1988), *The Theory of Environmental Policy*, Cambridge University Press, London.

Brown, B. (1985), "Location and Housing Demand", *Journal of Urban Economics* 17: 30-41.

Bull, N., K. A. Hassett and G. E. Metcalf (1994), "Who Pays Broad-Based Energy Taxes? Computing Lifetime and Regional Incidence", *Energy Journal* 15: 145-164.

Citizens for Tax Justice (2004), "Top Federal Income Tax Rates on Regular Income and Capital Gains since 1916", http://www.ctj.org/pdf/regcg.pdf

Cooter, R. and E. Helpman (1974), "Optimal Income Taxation for Transfer Payments under Different Social Welfare Criteria", *Quarterly Journal of Economics* 88: 656-670.

Fama, E. and M. Jensen (1983), "Agency Problems and Residual Claims", *Journal of Law & Economics* 26:327-349.

Feldstein, M., and J. Slemrod (1980), "Personal Taxation, Portfolio Choice, and the Effect of the Corporate Income Tax", *Journal of Political Economy* 21: 129-158.

Feldstein, M., L. Dicks-Mireaux, and J. Poterba (1983), "The Effective Tax Rate and the Pretax Rate of Return ", *Journal of Public Economics* 21: 129-158.

Foley, D. (1967), "Resource Allocation and the Public Sector", *Yale Economic Essays* 7: 45-98.

Frankel, M. (1985), “Amenity Changes, Property Values, and Hedonic Prices in a Closed City”, *Journal of Environmental Economics and Management* 12: 117-131.

Frankel, M. (1987), “Taxes, Pollution, and Optimal Abatement in an Urban Economy”, *Journal of Urban Economics* 22: 117-135.

Fullerton, D. and D. Rogers (1993), *Who Bears the Lifetime Tax Burden*, Brookings, Washington D.C.

Fullerton, D. (1996), “Comment on “High-Income Families and the Tax Changes of the 1980s: The Anatomy of Behavioral Response (by Slemrod, J.)””, in Feldestein, M. and J. M. Poterba, ed, *Empirical Foundations of Household Taxation*: 189-192, The University of Chicago Press, Chicago, IL.

Fullerton, D. and G. Heutel (2004), “The General Equilibrium Incidence of Environmental Taxes”, NBER working paper series #11311, Cambridge, MA, 2005.

Fullerton, D. and G. Heutel (2004), “The General Equilibrium Incidence of Environmental Mandates”, working paper, the University of Texas at Austin.

Glaeser, E. (2005), “Inequality”, NBER working paper series #11511.

Goolsbee, A. (1998), “Taxes, Organizational Form, and the Deadweight Loss of the Corporate Income Tax”, *Journal of Public Economics* 69: 143–152.

Goolsbee, A. (2004), “The Impact and Inefficiency of the Corporate Income Tax: Evidence from State Organizational Data”, *Journal of Public Economics* 88: 2283-2299.

Gordon, R. and J. Mackie-Mason (1994), “Tax Distortions to the Choice of Organizational Form”, *Journal of Public Economics* 55: 279-306.

Gravelle, J. and L. Kotlikoff (1989), “The Incidence and Efficiency Costs of Corporate Taxation when Corporate and Non-corporate Firms Produce the Same Good”, *Journal of Political Economy* 97: 749-780.

Gravelle, J. and L. Kotlikoff (1993), “Corporate Tax Incidence and Inefficiency when Corporate and Non-corporate Goods are Close Substitutes”, *Economic Inquiry* 31: 501-516.

Harberger, A. (1962), “The Incidence of the Corporate Income Tax”, *Journal of Political Economy* 70: 215-240.

Harberger, A. (1966), "Efficiency effects of taxes on income from capital", in M. Krzyzaniak, ed., *Effects of the Corporation Income Tax*: 107-117, Wayne State University Press, Detroit, MI.

Henderson, J. V. (1977), "Externalities in a Spatial Context, the Case of Air Pollution", *Journal of Public Economics* 7: 89-110.

Helpman, E. and E. Sadka (1978), "The Optimal Income Tax", *Journal of Public Economics* 9: 383-393.

Hockman, O. (1978), "A Two Sector Model with Several Externalities and Their Effects on an Urban Setting", *Journal of Urban Economics* 5: 198-218.

Instituto Nacional De Estadistica, "Encuesta Nacional de Empleo Urbano (National Urban Employment Survey) 1995-1999", Geografica e Informatica (National Statistics Institute), Mexico.

Isard, W. (1956), *Location and Space Economy*, MIT Press, Cambridge, MA.

Kuznets, S. (1955), "Economic Growth and Income Inequality", *American Economic Review* 65: 1-28.

Lydall, H. F. (1968), *The Structure of Earnings*, Oxford, U.K.

Mackie-Mason, J. and R. Gordon (1997), "How Much Do Taxes Discourage Incorporation", *The Journal of Finance* 52: 477-505.

Mieszkowski, P. M. (1972), "The Property Tax, An Excise Tax or a Profits Tax?", *Journal of Public Economics* 1:73-96.

Mirrless, J. (1971), "An Exploration in the Theory of Optimum Income Taxation", *Review of Economic Studies* 38: 175-208.

Muth, R. F. (1961), "The Spatial Structure of the Housing Market", *Papers and Proceedings of the Regional Science Association* 7: 207-220.

National Bureau of Economic Research, Inc. (NBER), "CPS Merged Outgoing Rotation Groups (CPS MORG) 1979 – 2004".

Pechman, J. A. and B. A. Okner (1974), *Who Bears the Tax Burden*?, Brookings Institution Press, Washington, DC.

Pechman, J. A. (1987), "Tax Reform: Theory and Practice", *The Journal of Economic Perspectives* 1:11-28.

Phelps, E. (1973), "The Taxation of Wage Income for Economic Justice", *Quarterly Journal of Economics* 3: 331-354.

Pigou, A. C. (1932), *The Economics of Welfare,* Macmillan, London.

Porter, M. E. and Class van der Linde (1995), "Toward a New Conception of the Environment-competitiveness Relationship", *Journal of Economic Perspectives* 9: 97-118.

Rapanos, V. T. (1992), "A Note on Externalities and Taxation", *Canadian Journal of Economics* 25: 226-232.

Rapanos, V. T. (1995), "The Effects of Environmental Taxes on Income Distribution", *European Journal of Political Economy* 11: 487-501.

Sadka, E. (1976), "On Income Distribution, Incentive Effects, and Optimal Income Taxation", *Review of Economic Studies* 9: 261-268.

Samuelson, P. A. (1954), "The Pure Theory of Public Expenditure", *Review of Economics and Statistics* 36: 387-389.

Sasaki, K. (1987), "A Comparative Static Analysis of Urban Structure in the Setting of Endogenous Income", *Journal of Urban Economics* 22: 53-72.

Seade, J. (1977), "On the Shape of Optimal Tax Schedules", *Journal of Public Economics* 7: 203-236.

Sheshinski, E. (1989), "Note on the Shape of the Optimum Income Tax Schedule", *Journal of Public Economics* 40: 201-215.

Shoven, J. (1976), "The Incidence and Efficiency Effects of Taxes on Income from Capital", *Journal of Political Economy* 84: 1241-1283.

Stiglitz, J. (1982), "Self-selection and Pareto Efficient Taxation", *Journal of Public Economics* 17: 213-240.

Stiglitz, J. (1987), "Pareto Efficient and Optimal Taxation and the New Welfare Economics", *Handbook of Public Economics* V2, North-Holland, Amsterdam.

Slemrod, J., Yitzhaki S. and Mayshar J., and Lundholm, M. (1994), "The Optimal Two-Bracket Linear Income Tax", *Journal of Public Economics* 53:269-290.

Slemrod, J. (1996), "High-Income Families and the Tax Changes of the 1980s: The Anatomy of Behavioral Response", in Feldestein, M. and J. M. Poterba, ed, *Empirical Foundations of Household Taxation,* The University of Chicago Press, Chicago, IL.

Slemrod, J. and Bakija, J. (2000), “Does Growing Inequality Reduce Tax Progressivity? Should It?”, NBER working paper series #7576.

Stern, N. H. (1976), “On the Specification of Models of Optimum Income Taxation”, *Journal of Public Economics* 6: 123-162.

Stull, W. J. (1973), “A Note on Residential Bid Price Curves”, *Journal of Regional Science* 13: 107-113.

Stull, W. J. (1974), “Land Use and Zoning in an Urban Economy”, *the American Economic Review* 64: 337-347.

Tax Policy Center, “Historical Corporate Top Tax Rate and Bracket”, http://www.taxpolicycenter.org/TaxFacts/corporate/..%5Ctables%5CCorporate%5CPDF%5Chistorical.pdf, Urban Institute and Brookings Institution, Washington D. C.

Tiebout, C. M. (1956), “A Pure Theory of Local Expenditures”, *Journal of Political Economy* 64: 416-424.

U.S. Department of Commerce, Economics and Statistics Administration, Bureau of Economic Analysis (1987), *Fixed Reproducible Tangible Wealth of the United States*, *1925-1985*, Washington D.C.

U.S. Department of Commerce, Economics and Statistics Administration, Bureau of Economic Analysis (1993), *Fixed Reproducible Tangible Wealth of the United States*, *1925-1989*, Washington D.C.

U.S. Department of Commerce, Economics and Statistics Administration, Bureau of Economic Analysis (1999), *Fixed Reproducible Tangible Wealth of the United States*, *1925-1994*, Washington D.C.

U.S. Department of Commerce, Economics and Statistics Administration, Bureau of Economic Analysis (1999), *Fixed Reproducible Tangible Wealth of the United States*, *1925-1997*, Washington D.C.

U.S. Department of the Treasury, Office of Tax Analysis (2002), *Capital Gains and Taxes Paid on Capital Gains for Returns with Positive Net Capital Gains*, 1954-2000, Washington D.C.

U.S. Department of the Treasury (2002), *Internal Revenue Service, Personal Exemptions and Individual Income Tax Rates*, 1913-2002, Washington D.C.

Varian, H. (1974), “Equity, Envy, and Efficiency”, *Journal of Economic Theory* 9: 63-91.

Wheaton, W. C. (1974), “A Comparative Static Analysis of Urban Spatial Structure”, *Journal of Economic Theory* 9: 223-237.

Vita

Jijun Tan was born in Yangcheng, Shanxi, China on November 20, 1973, the son of Kuannian Tan and Xiaojin Wang. After graduation from the First High School of Yangcheng in Yangcheng, Shanxi, China in 1991, he entered Shanxi University in Taiyuan, Shanxi, China. He received the degree of Bachelor of Economics from Shanxi University in July 1995. In September 1996, he entered the Guanghua School of Management at Beijing University in Beijing, China. He was married to Xianmei Yang on July 20, 1997 in their hometown, Yangcheng, China. In July 1998, he received the degree of Master of Economics from Beijing University. Then he worked in the Guoson Security Co. in Beijing for one year as an investment banker. In August 2000, he entered the Graduate School in Economics at the University of Virginia. He received a degree of Master of Arts from the University of Virginia in May 2001. Then he transferred to the Graduate School in Economics at the University of Taxes at Austin in August 2001. The son of Xianmei Yang and his, Ruiyang Tan, was born on December 9, 2003 in Austin, Texas. The daughter of theirs, Xinyang Tan, was born on November 17, 2005 in the same city.

Permanent Address: RM3-102, BLDG 0, FENGXIANG XIAOQU, JINCHENG, SHANXI, CHINA 048000

This dissertation was typed by the author.