DISSERTATION

# ESSAYS ON FISCAL DECENTRALIZATION AND TAXATION


Submitted by

Ardyanto Fitrady

Department of Economics


In partial fulfillment of the requirements

For the Degree of Doctor of Philosophy

Colorado State University

Fort Collins, Colorado

Summer 2014

Doctoral Committee:

Advisor: Harvey Cutler
Co-Advisor: Anita Alves Pena

David Mushinski
Stephan Kroll

Copyright by Ardyanto Fitrady 2014

All Rights Reserved

ABSTRACT

ESSAYS ON FISCAL DECENTRALIZATION AND TAXATION

This dissertation investigates two important topics in economics. First, the impacts of spillovers of public goods on the potential benefits from decentralization in an urban economy. Second, the role of tax evasion and uncertainty on the optimal taxation for two-class economy.

Theoretical model and numerical simulations are used to study the first topic in Chapter 2. The results from the numerical simulation shows that the spillover level has an impact on the potential gain of decentralization. The general results of the numerical simulations demonstrate that as the degree of spillovers increases, the potential gain of decentralization over centralization diminishes in both cases of metropolitan areas in developed and developing countries. These results celebrate Oates' decentralization theorem where decentralization is more beneficial when spillovers among jurisdictions are relatively low. However, the result also shows that the impact of the spillover level on the potential gain of decentralization varies across different levels of income vis-à-vis income inequality. It shows that a metropolitan area with lower mean income will be suffer more from the spillover than a metropolitan area with higher mean income. The numerical simulation also shows that a higher level of inequality amplifies the benefit of decentralization. It illustrates that the developed country—in this case the United States—that generally has higher income inequality, potentially gains more benefits from decentralization.

Theoretical models are used in Chapter 3 to examine the importance of tax evasion in the optimal taxation theory that are built based on previous studies. Although one can find that most of the results are intuitive, the model shows that considering tax evasion and uncertainty is important in implementing tax policies, particularly in the process of setting the tax rates for income tax and sales tax. The income tax for the high-type individual will be higher as the degree of tax evasion increases and the income tax of the low-type individual decreases as the probability of being detected for the high-type increases, ceteris paribus. The result also shows that the optimal income tax rate for the low-type individual increases as the marginal utility of mimicking the low type or the marginal utility of income for the mimicker increases, ceteris paribus. In other words, income tax for the low-type will increase if the high type has more incentive to mimic the low-type.

## ACKNOWLEDGMENTS

First of all, I would like to thank my family. Most of all, my beloved wife Dian, and our dearly loved sons Saka and Arana whose love and patience support me for all the years I spent in pursuing my doctoral degree and still continue until today. I also thank my parents for their endless support and prayers they have provided me throughout my life.

I am a very grateful to my advisor Harvey Cutler his guidance and supports in my dissertation along with his noteworthy patience. He was always available for discussion over the development of this dissertation. My deep gratitude, appreciation, and respect go to my co-advisor Anita Alves Pena for her insightful advice, constructive comments, and remarkable kindness. I cannot thank you enough. Thank you to David Mushinski and Stephan Kroll for their valuable time and constructive comments. Many elements of my dissertation were improved through discussions with them.

Many thanks to all my professors in the Economics Department from whom I learned a lot, especially Alexandra Bernasek and Robert Kling who are always willing to provide advices in my early years as a graduate student. Special thanks to Barbara Alldredge, Brooke Taylor, and Jenifer Davis for their dedication to the success of graduate students in the Department of Economics, Colorado State University.

I would also like to express my gratitude and thanks to the Faculty of Economics and Business, Universitas Gadjah Mada that has supported me in my last year in Ph.D. program.

Last but not least, I would like to offer my sincerest thanks to all my friends in the United States and Indonesia who have supported me throughout of my academic life in Colorado State University.

While all contributions helped enormously in the process of this dissertation, any errors are my own.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

# Introduction

This dissertation focuses on frameworks of two important topics in public economics, i.e. urban fiscal decentralization and optimal taxation. Discourses on both topics are important in public and urban economics, implying that there are always possibilities to improve the existing theory with different methods and assumptions. Both theoretical and numerical approaches will be used to investigate the important aspects of urban fiscal decentralization. It is not the purpose of this dissertation to change fundamental understandings of existing theory or models. Instead, it focuses more on contributing to the existing model or theory by introducing new variables into the models. Parallel results will enrich the existing theory by demonstrating some robustness of the model as new variables are introduced in the model.

The second chapter discusses fiscal decentralization in an urban area where the authority to provide local public goods, as well as the taxation to finance the provision, can be centralized in one large jurisdiction or decentralized into two small jurisdictions or more. The debate on this issue lies on two main views of fiscal decentralization. First, fiscal decentralization will benefit the economy in the aggregate. This idea is mostly based on Oates' fiscal decentralization theorem (1972). Second, fiscal decentralization is not better than centralization. This idea is based on the belief that the effect of spillovers offsets the benefits generated by fiscal decentralization as in Besley and Coate (2003). Moreover, when there is no spillover in the economy, Calabrese et al. (2012) find that centralized tax policy is better than decentralized tax policy.

Pareto optimum characterizations with spillovers[1]in urban settings are presented in Chapter 2. Following utilitarian economics, the optimal solution of the model is solved by maximizing total welfare of agents in the economy with three main constraints: housing market equilibrium, government balanced budget, and balanced-budget transfer; while considering spillovers among jurisdictions. As expected, the Pareto optimum characterization are not analytically solvable that leads to numerical simulations in the next section.

The next section in Chapter 2 conducts the empirical study using numerical simulations. By utilizing the model, assuming some parameters in the model, and employing empirical data, we calibrate the model in order to reach the equilibrium. Cases of urban areas in the United States and Indonesia (as a comparison of developed and developing economies) are used to illustrate two different empirical cases. We also use two types of taxation in the numerical simulations, i.e. efficient tax (head tax) and property tax to analyze the difference between both cases. Estimations have been made to gauge whether the benefits from fiscal decentralization dominates the costs when spillovers across jurisdictions are present.

The results from the numerical simulation shows that the spillover level has an impact on the potential gain of decentralization. The general results of the numerical simulations demonstrate that as the degree of spillovers increases, the potential gain of decentralization diminishes in both cases of metropolitan areas in developed and developing countries. These results celebrate Oates' decentralization theorem where decentralization is more beneficial when spillovers among jurisdictions are relatively low due to the optimal levels of local public good provisions.

[1]In this research, *spillover* is defined as a positive externality of local public goods in a certain jurisdiction to households' utility in other jurisdictions. We assume that there is no negative externality in the economy from local public good provision.

However, the result also shows that the impact of the spillover level on the potential gain of decentralization varies across different levels of income vis-à-vis income inequality. It shows that a metropolitan area with a relatively low mean income will be negatively affected by spillover more than a metropolitan area with a relatively high mean income. The numerical simulation also shows that a higher level of inequality amplifies the benefit of decentralization. It illustrates that developed countries—that generally have higher income inequality—potentially gain more benefits from decentralization.

Chapter 3 focuses on the optimal direct-indirect taxation in the economy, particularly income taxation and sales taxation. Related research, from Ramsey (1927) to Saez (2001), are discussed in the literature review. The discussion is divided into three main areas: (1) sales taxation; (2) income taxation; and (3) direct-indirect taxation. Each section focuses on the model used and its corresponding results. This chapter contributes to popular discussion not only to public economists but also to policy makers.

A theoretical model is built based on Nava et al. (1996) and Boadway et al. (1994) by taking into account uncertainty and tax evasion. The modification introduces intuitive results on the effects of tax evasion and uncertainty on income-tax and sales-tax rates. The chapter is closed by conclusions and policy implications. Although one can find that most of the results are intuitive, the model shows that considering tax evasion and uncertainty is important in implementing tax policies, particularly in the process of setting the tax rates for income tax and sales tax.

The model in Chapter 3 shows that in the Pareto optimum the sign of income tax for the high-type agent depends on the relation between the (taxed) private good and leisure. If the private good is a substitute for leisure, then the income tax for the high-type will be negative. Conversely, if the private good is a complement for leisure, the income tax of the high type will be positive. This result resonates the result in Nava et al. (1996). As far as the tax evasion is concerned, the income tax for the high-type individual will be higher as the degree of tax evasion increases. We also find that if the taxed private good and leisure are complementary—i.e. when the income tax is positive for the high type—and the elasticity of demand with respect to income tax is larger than the elasticity of labor supply with respect to income tax, the commodity tax rate should be higher than the income tax rate to optimize the welfare.

For the case of low-type, the model shows that the income tax of the low-type individual decreases as the probability of being detected for the high-type increases, *ceteris paribus*. The result shows that the optimal income tax rate for the low-type individual increases as the marginal utility of mimicking the low type or the marginal utility of income for the mimicker increases, *ceteris paribus*. In other words, income tax for the low-type will increase if the high type has more incentive to mimic the low-type. Conversely, the optimal income tax decreases as the marginal utility of tax revenue, the number of low-type agents, proportion of reported income, probability that the high-type will be detected when hiding the income, or the cost of hiding the income increases.

CHAPTER 2

# Fiscal Decentralization, Local Public Goods, and Welfare of Cities

## 2.1 Introduction

### 2.1.1 Background

The idea of fiscal decentralization or fiscal federalism[1] has been started from Plato's idea about government in his work *The Republic*—one of most influential works in western philosophy (Hooghe and Marks, 2007). However, a formal approach on how economists look at the fiscal decentralization—or fiscal federalism—started in Oates' seminal book in 1972 *Fiscal Federalism*. This literature shows that a decentralized system, under the assumption of no spillover effects or spillovers among jurisdictions, will generate higher social welfare than centralized system. The trade-off between those two systems depend on the heterogeneity of preference and the degree of spillovers (Besley and Coate, 2003). Although the final result is not really unexpected, Oates provides analysis to see this issue from economic point of view for the first time. Many works in the literature have been written based on Oates' fiscal decentralization theorem in the following years up to today and influence economic and political systems widely.

The influence of fiscal decentralization brings together the idea of *an ideal political system* where smaller governments should have more authority to determine the types of goods and services that should be provided in localities and decide how to finance them. The

[1] In general, fiscal decentralization is the devolution by the central government to local governments (states, regions, municipalities) of specific functions with the administrative authority and fiscal revenue to perform those functions (Kee, 2004). Slightly different definition as opposed to fiscal federalism will discuss later in the literature review.

reasons are mainly because local governments know better about their available resources and how to allocate the resources so the allocation will be more efficient. Moreover, local governments know better what their people need so the allocation is expected will be more effective.[2]

Samuelson (1954) asserts that there is no market solution for public good provision, where what he means by market solution in local economy is a decentralized policy. Based on his assertion, public goods are always under provided since there is no feasible method to charge the consumers (Tiebout, 1956). In other words, there is no reason to believe that public goods will be provided efficiently. This argument is criticized by Tiebout (1956). He conjectures that competition among publicly elected governments for mobile households may yield an efficient provision of local public goods. However, Tiebout's hypothesis is not free from critiques as many articles are written to argue against Tiebout's hypothesis. We will discuss in detail about the theory in the literature review section.

The idea of fiscal decentralization does not only have impacts on developed countries such as the United States, but also has impacts on developing countries, such as Indonesia. More than that, the issue is not only interesting in the context of high level of government system such as central government *versus* state governments, but also in the smaller jurisdictional system such as a metropolitan area (MA) *versus* cities in the corresponding MA. Should we give the power to finance and to provide local public goods in large jurisdiction (centralized system) or let smaller jurisdictions decide what they need to do (decentralized system)? Here, urban economics comes into play.

[2]One may recall Musgrave's three-function of government (1939): stabilization, income redistribution, and resource allocation, where the first two functions belong to the central government and the last function belongs to the local government.

Urban economic literature has a wide spectrum of discussions of decentralization policy and spillover effects independently.[3] However, to the best of my knowledge there is no literature that specifically focuses on the discussion of fiscal decentralization where spillover effects exist in the case of urban system. One piece of the literature focusing on fiscal decentralization in urban area is from Calabrese et al. (2012). However, their article discusses Tiebout equilibrium where spillovers do not exist.[4] Thus, it is of interest to observe theoretically how externalities or spillover effects will affect social welfare in a setting of a metropolitan area. Adding the spillovers will provide an important insight into urban economics where spillovers among jurisdictions are an essential aspect to be considered in urban economics.

Based on the discussion above, it is very appealing to do research on this area: how fiscal decentralization will impact social welfare via local public good provision with the existence of spillovers among jurisdictions in the urban area? The research will be conducted from perspectives of urban economics and public economics. A theoretical model will be built and data for urban areas in the United States and Indonesia will be calibrated to investigate the empirical aspects of the model.

### 2.1.2 Research questions

One of the most important issues in the debate of fiscal decentralization is the impact of decentralization on social welfare. Some economists argue that decentralization will be beneficial to society as in Oates' decentralization theorem. In contrast, some economists

[3] See Oates (1972), Oates (1999), Zhang and Zou (1998), Davoodi and Zou (1998), Besley and Coate (2003), Calabrese et al. (2012) for some references.

[4] One of the assumptions in Tiebout (1956) is no externalities or spillovers.

argue that centralization is better since the spillovers among jurisdictions will not be zero. Utilizing a political economy approach, Besley and Coate (2003) also provide that there is some level of externalities or spillovers where decentralization will be beneficial and vice versa.

To the best of my knowledge, none of fiscal decentralization literature points out whether fiscal decentralization leads to a gain or a loss for the society in the context of urban economy where the spillovers exist. Thus, the model in this dissertation is constructed to examine that area, particularly the impact of fiscal decentralization and spillovers on the urban welfare.

Specifically, the research questions are:

1. How can we model the impact of fiscal decentralization on social welfare when the spillovers exist in the context of an urban economy?
2. Do the results support the existing theory?
3. Does fiscal decentralization benefit cities and society when some degree of spillover is present?
4. Are the results consistent both in developed and developing countries?

### 2.1.3 Research outline

We will explore topics of fiscal decentralization, local public good provision, and their impacts on welfare in the context of urban economics. There will be two main sections in this chapter, i.e. the model and numerical simulations. The latter consists of two different results: (1) comparison between urban area in developed and developing countries; and (2) impacts of decentralization on different levels of income and inequality on the social welfare.

The chapter is divided into six sections. Section 2.2 will discuss literature review on the related topics mentioned above. It will begin with the concept of economics of the cities and equilibrium in each jurisdiction. It is then followed by the literature review in fiscal decentralization and a discussion on the evolution of public good theory, including Samuelson (1954), Tiebout (1956), and Bewley (1981). The chapter then is closed by a summary leading to the theoretical model in Section 2.3.

Section 2.3 discusses the theoretical model that will explain the description of the model and the characterization of Pareto optimum. The model is then used as bases of computer models and numerical simulations in Section 2.4 and 2.5.

Section 2.4 discusses the calibration using data for urban area in the United States and Indonesia. It is started with a description about the methodology, data, and assumptions used in the simulations. It is followed by the cases of efficient and property taxation in the United States and Indonesia that mainly focus on the welfare impacts of decentralization. The last section in this chapter will conclude the results, compare them with previous research and provide some policy implications.

## 2.2 Literature Review

This section discusses related literature that provides the underpinning theoretical framework to construct the model and conduct the analysis. We start with discussion on economics of cities and its equilibrium. Section 2.2.1 discusses spillover effects across jurisdictions that potentially affect the equilibrium while Section 2.2.2 discusses some literature on fiscal decentralization. The discussion will cover a *pure* economic theory from Oates (1972) and the political economy aspect of it. The following section then will cover a discussion of local public goods and welfare. This section will cover the positive approach from Musgrave (1959), normative approach from Samuelson (1954), Tiebout hypothesis from Tiebout (1956), the theory of clubs from Buchanan (1965), and other related literature on this topic. In the end of this section we will summarize the literature review, explaining how it relates to the model in this dissertation, and lead to the next section that will discuss the model in details.

### 2.2.1 Economics of cities and equilibrium

Urban economics is the main perspective that is used in this dissertation. Like other branches of economics, such as in Brueckner (1986), the most challenging part in urban economics is to construct rigorous economic explanation for a variety of observed regularities in the structures of real-world cities. This section will discuss an economic model for a metropolitan area (MA) when there are no externalities as constructed in Tiebout's model in his seminal work in 1956. We will discuss more on the literature of spillover effect across cities or jurisdictions and the fundamental theories of fiscal decentralization and local public finance in the next sections.

The welfare effect of decentralization is ambiguous in the literature. Calabrese et al. (2012) evaluates the welfare effect of local public good provision in metropolitan area using multiple jurisdictions and mobile households. The result from their calibration shows that centralization is more efficient in property-tax equilibrium. Since we will work closely to this article, we will summarize this paper in detail.

The main objective in Calabrese et al. (2012) is welfare comparison of a centralized equilibrium to a decentralized Tiebout equilibrium when the equilibrium exists. In doing so, they use theoretical models and calibration techniques employing data from metropolitan area in the United States, particularly American Housing Survey 1999. Some parameters are drawn from other research's results or set to fit the actual data.

Their main finding is that in property-tax equilibrium, centralization is more efficient than decentralization. They find that the welfare effects run counter to basic intuition concerning the gain from the Tiebout process. Specifically they find that the externality in choice of residence is the primary source of welfare loss based on the model.

There are three agents in the model: households, housing owners, and local governments. Households choose a jurisdiction to live by renting houses where housing owners live outside the jurisdictions (absentee landlords). To keep the model simple, Calabrese et al. (2012) makes the following assumptions:

- MA is divided into $J$ jurisdictions.
- Each jurisdiction has fixed boundaries where the central city has the largest area and the rest of communities have identical sizes.
- Each jurisdiction has a local housing market, provides (fully congested) public good, $g$, and charges property taxes, $t$.
- Local public good provision and the property tax rate are determined by majority rule in each jurisdiction.
- Households are renters and housing is owned by absentee landlords.
- There is a continuum of households that vary in incomes and preferences toward local public goods.
- Households behave as price takers and have preferences defined over a local public good, and housing services, $h$.

There are three stages to determine equilibrium in this model and backward induction is used to guarantee that the total utility are maximized in each stage. We assume that households are rational and forward looking meaning that their decisions always perfectly anticipate the results in the following stages. The stages are:

1. Households choose a jurisdiction and rent a home in a jurisdiction;
2. They vote in the corresponding jurisdiction for property tax that is used to finance the local public goods;
3. Local public goods are determined using local government budget balance principle, meaning that total revenues (from tax collection) are equal to total expenditures (i.e. local public good provision).

In characterizing Pareto allocation the planer's problem is to maximize social welfare function with respect to some constraints. There are three constraints we use in the model:

- balanced-budget transfer
- housing market clearance
- local government balanced budget

Utility function is $U = U(x, h, g; \alpha)$ where $x$ is private consumption, $h$ is housing consumption, $g$ is public good expenditure, and $\alpha$ is household preference for local public goods. Denote that $S \equiv [\alpha, \bar{\alpha}]x[y, \bar{y}] \subset R_+^2$ and the indirect utility function:

$$V^e(p_j, g_j, y + r(y, \alpha) - T_j, \alpha) \equiv \max_h U(y + r(y, \alpha) - T_j - p_j h, h, g_j; \alpha), \tag{2.1}$$

where $p_j$ is the housing price in jurisdiction $j$, $g_j$ is public good expenditure in jurisdiction $j$, $y$ is household income, $\alpha$ is a parameter for the preference toward local public goods, $r(y, \alpha)$

is transfers to households based on income and their preferences, $T_j$ is a head tax, and $h$ is housing consumption.

Thus, the maximization problem is:

$$\max_{r(y,\alpha),a_i(y,\alpha),R,T_i,t_i,p_i,g_i} \sum_{i=1}^{J} \Big[ \int_S \omega(y,\alpha) V^e(p_i, y + r(y,\alpha) - T_i, g_i, \alpha) \tag{2.2}$$

$$a_i(y,\alpha) f(y,\alpha)\, dy\, d\alpha$$

$$+ \omega_R (R/J + \int_0^{p_i/(1+t_i)} H_s^i(z)\, dz) \Big]$$

subject to:

$$R + \int_S r(y,\alpha) f(y,\alpha)\, dy\, d\alpha = 0 \tag{2.3}$$

$$\int_S h_d(p_i, y + r(y,\alpha) - T_i, g_i, \alpha) a_i(y,\alpha) f(y,\alpha)\, dy\, d\alpha = H_S^i(p_i/(1+T_i)) \tag{2.4}$$

$$T_i \int_S a_i(y,\alpha) f(y,\alpha)\, dy\, d\alpha + \frac{t_i p_i}{1+t_i} H_S^i(p_i/(1+T_i)) = g_i \int_S a_i(y,\alpha) f(y,\alpha)\, dy\, d\alpha \tag{2.5}$$

$$a_i(y,\alpha) \in [0,1]\ ,\ \sum_{i=1}^{J} a_i(y,\alpha) = 1 \quad \forall (y,\alpha), \tag{2.6}$$

where $T_i$ and $t_i$ are head tax and property tax rates, respectively; $a_i(y,\alpha) \in [0,1]$ is the proportion of households $(y,\alpha)$ assigned by the planner in community $i$; $h_d$ is housing demand; and $H_S^i$ is housing supply. Also let $\omega(y,\alpha) > 0$ denote the weight on household $(y,\alpha)$'s utility in social welfare function and $\omega_R > 0$ denote the weight on absentee landlords' utility in social welfare function.

The characterization of Pareto optimum confirms that the social optimum will have no property taxation, only head taxes. Moreover, Calabrese et al. (2012) also show that unilateral household choice of residence in the world of head taxation would achieve the Pareto optimum. Specifically, they state in a proposition:[5]

> *In an efficient differentiated allocation: (a) $t_i = \eta_i = 0$ and $T_i = g_i$, (b) $g_i$ satisfies the community Samuelson condition, and (c) households are assigned to the community where $V_i^e$ is at maximum.*

By the proposition, if the head taxes generate optimal local public good provision then the household choice of jurisdictions is socially optimal. The proposition above can also be viewed as a generalization of Oates' decentralization theorem where there are no spillovers of local public goods across jurisdictions in the model, costs of provision are the same for centralized and decentralized systems, and the provision is uniform under centralization (Oates, 1972, 1999).

However, to better approximate the real world, Calabrese et al. (2012) set head taxes, $T_i$, equal to zero and property rates, $t_i$, is positive. It generates what they call the Jurisdictional Choice Externality (JCE).[6] By definition, $\text{JCE}_i(y, \alpha)$ measures the social benefit or cost imposed on others when households $(y, \alpha)$ locates in jurisdiction $i$. Consequently, JCE becomes one of the sources of welfare loss in the model. Theoretically, they define three sources of inefficiencies in Tiebout Equilibrium:

1. property taxation generates deadweight loss.
2. majority choice of the tax rate conforms to the choice of the median-preference households in a jurisdiction, which generally differs from the choice that would maximize average welfare.
3. externalities arise in household choice of jurisdiction (JCE).

[5] Proposition 3 in Calabrese et al. (2012).

[6] Note that *externality* in Calabrese et al. (2012) is a negative externality; second, it exists within a jurisdiction, not across jurisdictions.

Using the calibration techniques, they find that JCE is the main source of welfare loss. This is counter-intuitive since the mobility that is needed to achieve Tiebout (approximately) optimal allocations of decentralization is also the main reason why one cannot gain them under property taxation.

### 2.2.2 Fiscal decentralization

Fiscal federalism and fiscal decentralization are defined with slight distinctions. Oates (1999) defines fiscal federalism as a general normative framework for assignment of functions to the different levels of government and appropriate fiscal instruments for carrying out these functions. On the other hand, Kee (2004) defines fiscal decentralization is the devolution by the central government to local governments (states, regions, municipalities) of specific functions with the administrative authority and fiscal revenue to perform those functions. Based on the definitions, one may see that the definitions of fiscal federalism and fiscal decentralization are in the same spirit of granting authority to the lower level of governments. However, fiscal federalism puts stress more on the normative aspect: the set of guidelines of how to share functions among government levels. In contrast, fiscal decentralization focuses more on the positive aspect: the implementation of distributing function among government levels.

The earliest work of decentralization can be traced back to the work of Musgrave in 1959 in his most cited work *The Theory of Public Finance*. Musgrave (1959) describes fiscal federalism as a system which permits different groups living in various states to express different preferences to public services; and this, inevitably, leads to differences in level of taxation and public services. However, it is *Oates' decentralization theorem* (Oates, 1972) that has become the central to the discussion of fiscal federalism. The theorem states that fiscal authority should be decentralized in the nonexistence of interjurisdictional externalities. Specifically, Oates (1999) points out that local government are apparently much closer to the people and posses knowledge of both local preferences and cost conditions that central agency is unlikely to have. In the same article, Oates (1999) also mentions that both in developed and developing countries, decentralization is needed to improve the performance of their public sectors.

Oates (1998) demonstrates that welfare gains from fiscal decentralization is difficult to achieve through centralized provision of all public goods. The following figures explain his idea clearly.

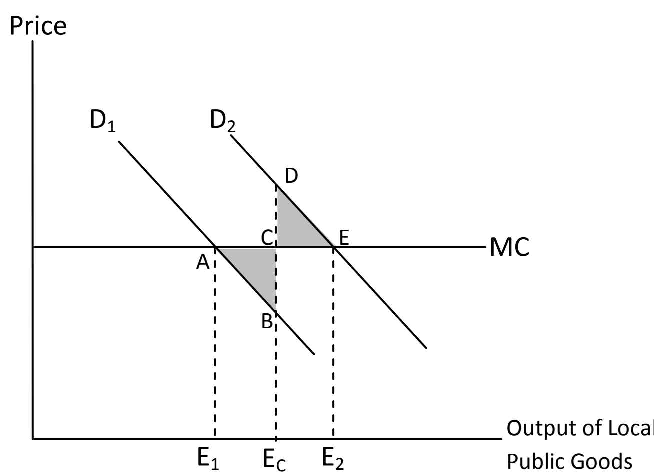



Figure 2.1: Welfare loss from centralization: variations in demand

*Sources: On the Welfare Gains from Fiscal Decentralization (Oates, 1998)*

Figure 2.1 shows demand curves for a local public goods in jurisdiction 1, $D_1$, and jurisdiction 2, $D_2$. Assuming that the marginal cost of local public good provision is constant per household, MC, then we have points A and E are the optimum level of provisions for jurisdiction 1 and 2, respectively. If the central government determines a uniform level of provisions for both jurisdictions, one can observe that it will generates welfare loss as much as the shaded area in Figure 2.1 the loss for households in jurisdiction 1 is the triangle ABC and the loss for households in jurisdiction 2 is the triangle CDE. We can also conclude that the heterogeneity of demands (see $D_1$ and $D_2$), as well as elasticity differences, will affect the magnitude of the loss. The higher the degree of diversity, the higher the loss from centralization.[7] However, in a centralized system, it is not easy politically to treat different jurisdictions based on the preferences. Thus, decentralization is the answer in this framework.

Empirically, the impacts of fiscal decentralization on welfare is not easy to conduct because of difficulties in measuring the welfare itself. Empirical work on the impacts of fiscal decentralization on economic growth have some findings where the results are inconsistent. It depends on how the researchers define fiscal decentralization (Akai and Sakata, 2002) and the characteristics of the countries. Using panel data of 46 countries in 1970-1989, Davoodi and Zou (1998) finds that there is a significant negative impact on economic growth after the implementation of fiscal decentralization in developing countries and no significant impact in the case of developed countries.[8] However, the negative relationship may be the result of unclear definition of the data on the local government expenditures since the researchers cannot distinguish between current expenditures (salary and wages) and capital spending

[7]In the same article Oates (1998) also demonstrates the case where there is interjurisdictional cost differences. However, we do not include it in this dissertation since we assume that the cost are uniform across jurisdictions.

[8]Davoodi and Zou (1998) uses per capita GDP growth as the dependent variable, and average tax rate, time-dummy variables for the implementation of decentralization, population growth, initial human capital, initial per capita GDP, and investment share of GDP as independent variables.

for most developing country cases.[9]

At the state level, the results are also ambiguous. Using panel data for China at the provincial level covering period 1978-1992, Zhang and Zou (1998) find that fiscal decentralization significantly reduces provincial economic growth.[10] A similar result can be found in Xie et al. (1999) for the United States case. Using time-series data from 1948-1994, they find that fiscal decentralization is potentially harmful to economic growth.[11] In contrast, Akai and Sakata (2002) find that the fiscal decentralization has a significant contribution to economic growth.[12] Using panel data for 50 states in the United States 1992-1996, their results sustain *Oates' decentralization theorem*. This positive relationship between fiscal decentralization and economic growth is also supported by Brueckner (2006) that incorporate human capital in his theoretical model. Using an endogenous-growth and overlapping generation models, his research demonstrates the relationship between fiscal federalism and economic growth where the analysis shows that decentralization increases the incentive to save. Consequently, it will promote a higher investment in human capital that will lead to higher economic growth.

[9] It is also worth pointing out that Davoodi and Zou (1998) mentions a statement in Musgrave (1959) that the act of local government administrators does not necessary reflect the decentralization expenditures.

[10] Zhang and Zou (1998) use annual data from 1980 to 1992 for 28 provinces in China, where the dependent variable is the real income growth rate and the independent variables are the growth rate of labor force, investment rate, the degree of openness of the provincial economy, the degree of distortion in the provincial economy, the inflation rate, the degree of fiscal decentralization (measured by ratio of consolidated provincial spending to consolidated central spending, the ratio of provincial budgetary spending to central budgetary spending, and the ratio of provincial extra-budgetary to central extra-budgetary spending).

[11] Xie et al. (1999) use per capita output growth rate as the dependent variable, and average tax rate, state government spending share, local government spending share, labor growth rate, openness index, average tariff rate, inflation rate, price of energy, and Gini index as independent variables.

[12] Akai and Sakata (2002) use the average annual growth rate of per capita gross state product($\Delta GSP$) as the dependent variable, and indicator of fiscal decentralization (measured by ratio of local government revenue and expenditure to state in 1992), population growth rate, lagged $\Delta GSP$, education, share of Democrats in the legislature in 1992, Gini index, regional dummy, share of patents, and openness index as independent variables.

Besley and Coate (2003) carry out a comparison between *Oates' decentralization theorem*—the standard approach—and a political economy approach. The standard approach reemphasizes the finding that under assumption of no spillovers, a decentralized system is superior. Centralization is only desirable if and only if the spillovers are sufficiently large.

Using a simple model, Besley and Coate (2003) illustrates the contrast between decentralization and centralization under the standard approach. Assuming two-jurisdiction world, in a decentralized system a local government chooses the amount of local public goods to maximize public good surplus in its jurisdiction.[13] The set of expenditure levels in two jurisdiction $(g_1^d, g_2^d)$ is a Nash equilibrium where:

$$g_i^d = \underset{g_i}{max}\{m_i[(1-\kappa)\ln g_i + \kappa \ln g_{-i}] - pg_i\} \text{ for } i \in \{1, 2\}. \tag{2.7}$$

$g_i^d$ is the expenditure level under decentralized system for jurisdiction $i$; $m_i$ is the mean type in jurisdiction $i$ which is equal to median type. They also assume that $m_1 \geq m_2$ and $2m_1 < \bar{\lambda}$ where $\lambda$ is a public good preference parameter with a range $[0, \bar{\lambda}]$. The parameter $\kappa \in [0, \frac{1}{2}]$ is the degree of spillovers; $\kappa = 0$ means the citizens consider that they are affected only by public goods in their own jurisdiction and $\kappa = 0.5$ means they consider they are affected equally by public goods in both jurisdictions. $p$ is the price of the local public goods that is equal to the quantity of private goods needed to produce the local public goods.[14]

[13] Besley and Coate (2003) use *district* in lieu of *jurisdiction* in the original article. This dissertation uses *jurisdiction* whenever possible for consistency.

[14] By this, Besley and Coate (2003) also assume that the local governments finance the expenditures on local public goods using head tax where each citizen in $i$ will pay a tax of $pg_i$. Each citizen is assumed to have a sufficient endowment to meet their tax obligation.

On the other hand, under a centralized system the government chooses a uniform expenditure level to maximize aggregate surplus where:

$$g^c = \underset{g}{max}\{[m_1 + m_2]\ln g - 2pg\}. \tag{2.8}$$

By taking the first derivatives for both decentralization and centralization maximization problems, one can easily observe that there is no impact of spillovers under the centralized system and it maximizes the surplus when the jurisdictions are identical.[15] Restating the conclusion above, the models show that if we have identical jurisdictions and there are spillovers among them, centralization is a better choice than decentralization. In contrast, when the jurisdictions are highly heterogeneous and the spillovers do not exist, decentralization dominates centralization. Since the surplus under decentralization is decreasing in spillovers, there is a critical level of spillovers where centralization will dominate decentralization.

However, this result relies on an unrealistic assumption that expenditures under centralization are perfectly uniform across jurisdiction. Albeit such a political pressure as discussed in *Oates' decentralization theorem*, there is a possibility for the central government to provide a different level of public goods in each jurisdiction to maximize the aggregate surplus. In that case, the centralized system may produce the surplus at least as much surplus as the decentralized system and even more when the spillovers are present.

Relaxing the assumption of uniformity, Besley and Coate (2003) utilize the idea of the citizen-candidate approach. Under decentralization, in the first stage a representative with

[15] Under centralization, if $m_1 > m_2$ then the provision in jurisdiction 1 is under-provided and the provision in jurisdiction 2 is over-provided.

preference $\lambda$ is elected in each jurisdiction. In the second stage, the level of expenditures is determined simultaneously by the elected representative in each jurisdiction.

Using a backward induction, the maximization problem of the expenditure level is as follows:

$$g_i^d(\lambda_i) = \underset{g_i}{max}\{\lambda_i[(1-\kappa)\ln g_i(\lambda_i) + \kappa \ln g_{-i}(\lambda_{-i})] - pg_i\} \text{ for } i \in \{1,2\}; \tag{2.9}$$

resulting:

$$(g_1(\lambda_1), g_2(\lambda_2)) = \left(\frac{\lambda_1(1-\kappa)}{p}, \frac{\lambda_2(1-\kappa)}{p}\right). \tag{2.10}$$

Substituting the result in the first stage (election stage), a citizen of type $\lambda$ in jurisdiction $i$ will enjoy a surplus:

$$\lambda\left[(1-\kappa)\ln\frac{\lambda_i(1-\kappa)}{p} + \kappa\ln\frac{\lambda_{-i}(1-\kappa)}{p}\right] - \lambda_i(1-\kappa). \tag{2.11}$$

Assuming citizen's preferences are a *single-peaked*, we have $(\lambda_1^*, \lambda_2^*) = (m_1, m_2)$ under decentralization. This result conveys an identical result as in the standard approach.

The result is different under centralization where the allocation of public good expenditures depend on the behavior of legislature. Using the same two-stage policy determination as in decentralization case, Besley and Coate (2003) suggest a proposition that:

*Suppose that the assumptions of the political economy approach are satisfied and that the legislature is non-cooperative. Then:*

1. *If the jurisdictions are identical, there is a critical value of spillovers, $\kappa$, where $0 < \kappa < \frac{1}{2}$ such that centralized system produces a higher level of surplus if and only if $\kappa$ exceeds this critical value.*
2. *If the jurisdictions are not identical, there is a critical value of spillovers, $\kappa$, where $0 < \kappa < \frac{1}{2}$, such that a centralized system produces a higher level of surplus if and only*

*if $\kappa$ is strictly larger than the critical value. This critical value exceeds this standard approach.*

It suggest that there is a critical value of spillovers where centralization dominates in both the identical and non-identical jurisdictions. The main difference compared to the standard approach is in part (2) where the standard approach only prescribes centralization for any spillover levels when jurisdictions are identical. Moreover, the extent of the conflict of interest among jurisdictions depends on spillovers and differences in preferences for public expenditure.

Other research on decentralization are from Wellisch (1993, 1994) where in his theoretical papers he focuses on the impact of household mobility across regions on efficiency. A characteristic of perfect mobility of households in his models support Tiebout's hypothesis and celebrate Oates' decentralization theorem as the decentralized provision of public goods is efficient. In particular, by assuming a decentralized equilibrium of the Nash-Cournot type, he finds that the equilibrium is efficient. In contrast to Oates' decentralization theorem, Wellisch (1993) finds that inter-regional benefit spillovers is efficient in his model as local governments perfectly internalize the spillovers in providing local public goods. However, Wellisch (1994) adds that central government intervention is needed to improve efficiency when some households are attached to particular regions for non-economic reasons such as cultural and nationalistic reasons.

In his papers above, Wellisch (1993, 1994) do not investigate the empirical aspects of his theoretical models. In contrast to our research, Wellisch (1993, 1994) also assume that all households have the same levels of incomes and only consider a head tax as a source for local public good provision via some production function. Moreover, both papers do not discuss fiscal decentralization[16] *per se* as the main interest of our research but more on links between household mobility and decentralized policy making (in providing local public goods).

Silva and Yamaguchi (2010) investigate decentralized enviromental policy making in the presence of imperfect household mobility. They find that as the energy supply is regulated by regional authorities and income is redistributed by central government, the equilibrium is socially optimal.[17]

Although Wellisch (1993, 1994) and Silva and Yamaguchi (2010) have different focuses and setups from our research, they agree that decentralizing some authorities to lower governments will create efficiency. Their results again favor Oates' decentralization theorem even with constrained mobility of household and spillover effects.

### 2.2.3 Local public goods and Pareto optimum

Following Musgrave (1959) a public good has two main properties: *nonrivalrous* and *nonexcludable* in consumption. The nonrivalrous property holds when consumption of a unit of goods by a consumer does not reduce the benefit of another consumer. This implies that the opportunity cost of the marginal user is zero. In other words, the marginal cost of adding another user is zero. In the other hand, the nonexcludable property holds when it is not possible to prevent others from consuming the same good.

While Musgrave (1959) discusses the positive aspect of public expenditure (for public goods), Samuelson (1954) writes the normative aspect of it in his seminal paper *A Pure Theory of Public Expenditure*. Samuelson points out the consequence—and the main problem—

[16]By fiscal decentralization, we focus on fiscal relationship between higher and lower levels of governments. See the definition of fiscal decentralization in Section 1.2.2.

[17]Silva and Yamaguchi (2010) assumes that energy supply generates both benefits (via consumption) and costs via (pollution damages).

of the non-excludability of public goods[18] i.e. a decentralized mechanism to obtain optimal public good provision; basically he argues that there is no market solution for public goods because it is difficult to reveal users' true preferences toward public goods due to free-riding problem. Consequently, there is no feasible way to charge the users of public goods.

This problem leads to the next seminal paper from Tiebout in 1956: *A Pure Theory of Local Expenditures*. In the paper, Tiebout (1956) argues that although there is no solution in national level, there is a solution at the local level for local public good provision in which a decentralized mechanism for obtaining an (approximately) optimal allocation exists.[19] The basic idea is that there is a large number of jurisdictions that provide different bundles of local public goods and taxes. People then will be sorted by *voting with their feet* to choose a jurisdiction that fit with their preferences. By doing that, they reveal their true demand for public goods and solve the preference revelation problem.

In the context of Tiebout model, a local public good is one that benefits only households in the local community. In an urban economics context, it is also worth mentioning that local public goods have a limited benefit area, meaning the benefit generated by the local public goods will only be enjoyed by a certain area close to the public goods. If households in a different jurisdiction do not benefit from the local public goods, then we can describe the situation as *no spillover effects*.

One important thing that should be understood in the Tiebout model is that in a case of local public goods, most public goods are subject to congestion. Thus, for any given

[18] Samuelson (1954) uses term *collective consumption goods* in lieu of *public goods*.

[19] Tiebout (1956) explicitly states in his paper that the word *approximate* is used to indicate the limitations of the model (see Tiebout (1956) footnote 13). Moreover, he mentions in the conclusion that "*...the problem does have a conceptual solution ... While the solution may not be perfect because of institutional rigidities, this does not invalidate its importance.*" *op.cit.*, p. 424.

level it becomes partially rivalrous. Education is one example that has been mostly used in in literature. Education is available for everyone but at some level of number of pupils, it becomes too crowded and less available to others.

There are several assumptions used in Tiebout's model whereas the key assumptions are: there are no spillover effects and the mobility of people is costless. The latter guarantees that people can move to any jurisdiction that meets households' preferences. Other assumptions, standards in many economic models today, include perfect information, the existence of a large number of jurisdictions, exogenous income, the existence of an optimal community size (related to average cost level), and that jurisdictions try to keep the population constant at an optimum level.

The fundamental question is then whether Tiebout equilibria are efficient when they exist. There is no exact answer for this question. Based on most literature, the equilibria will only be efficient under very restrictive assumptions. Bewley (1981) demonstrates a set of examples to criticize Tiebout's hypothesis on efficiency.

We show here one example to illustrate Bewey's idea, i.e. the case of pure public services (cost of provision is proportional to population). The assumptions are:

1. There are two identical regions $j = \{1, 2\}$
2. There are four consumers and four types of public services $\{A, B, C, D\}$
3. Labor ($L$) or leisure ($l$) is the only private good and each consumer is endowed with 1 unit of labor

Suppose we also have production relations: $n_j(g_{jA} + g_{jB} + g_{jC} + g_{jD}) = 2L_j$, where $n_j$ is the number of consumers in region $j$, $g_{jk}$ is the quantity of public service $k$ provided in region $j$ for $k = A, B, C, D$. Defining utility for each consumer is as follows:

$$u_A(l, g_{jA}, g_{jB}, g_{jC}, g_{jD}) = 2g_{jA} + g_{jB} \tag{2.12}$$

$$u_B(l, g_{jA}, g_{jB}, g_{jC}, g_{jD}) = g_{jA} + 2g_{jB} \tag{2.13}$$

$$u_C(l, g_{jA}, g_{jB}, g_{jC}, g_{jD}) = 2g_{jC} + g_{jD} \tag{2.14}$$

$$u_D(l, g_{jA}, g_{jB}, g_{jC}, g_{jD}) = g_{jC} + 2g_{jD}, \tag{2.15}$$

where $u_k$ is the utility function of consumer $k$, there is a hypothetical situation where we have combination of public goods in two regions are as follows:

$$\text{Region 1 has } (g_{jA}, g_{jB}, g_{jC}, g_{jD}) = (1, 0, 1, 0)$$

$$\text{Region 2 has } (g_{jA}, g_{jB}, g_{jC}, g_{jD}) = (0, 1, 0, 1)$$

If households behave following Tiebout's conjecture then {A,C} will choose to live in Region 1 and {B,D} choose to live in Region 2. Is it an optimum? The answer is no. One can easily see that society will have higher efficiency if {A,B} live in Region 1 and {C,D} live in Region 2.

Specifically Bewley (1981) identifies conditions for an efficient Tiebout's hypothesis as follows:

1. Only for a case of pure public services, not pure public goods. Pure public goods is a case where the cost is independent of population size while pure public service is a case where the cost is proportional to population size.

2. The number of regions is equal to the number of households' preferences. If one find that every households has a unique preference, then there will be one jurisdiction for each household.

According to Bewley (1981), those two conditions make Tiebout's hypothesis not an appealing result for the fundamental question of efficiency.

Nevertheless, Tiebout's hypothesis has become the center of discussions on local public good provision. In addition, Tiebout's theory has a close counterpart in term of providing a market solution for public good provision in the theory of clubs (Buchanan, 1965).[20] In the theory of clubs, as in the Tiebout model, clubs have an ability to restrict the provision of local public goods to those who are members of the clubs. Thus, the goods are now excludable and partially rivalrous. Another main difference is in the rental price for housing that is not fixed whilst the admission price to each member of a club is fixed. It means in Tiebout' setting households can reduce the consumption of housing to live in a certain jurisdiction.

The main result of the theory of the clubs is that individuals with heterogeneous preferences would be better-off by grouping in separated clubs where every member in a club has a homogeneous preference. In the context of local public goods, it implies that if there are two type of households: rich and poor, then the whole society will be better-off if each of them live in separate Tiebout jurisdiction than if they live together in a mixed jurisdiction.

However, this conclusion relies on some strong assumptions:

1. Jurisdictions only impose uniform taxation (head tax). If the poor people live in a

[20] The relationship between those two is still a subject of ongoing debate.

mixed jurisdiction, pay lower taxes than the rich people, and enjoy the same local public goods, it is not necessary that the poor people will be better-off in their own jurisdiction.

2. There is no positive externalities from having different households (with different endowments).

3. An integer number of optimal size of jurisdiction exists for each homogeneous group. If the optimal size of jurisdiction is $n$ households and there are $n+1$ households sharing the same preferences for local public goods, Tiebout's optimal equilibria will not exist.

Adding space and multilayer government into the model critically creates inconsistency in Buchanan's concept of nonspatial (flying) clubs for describing an efficient federal system of local governments and Tiebout's concept (Hochman et al., 1995). Hochman et al. (1995) conclude that flying clubs stay in a territory that will induce an optimal financing of the local public goods based on the land rent generated in their respective market areas. Thus, if more than one layer of government serves a given territory, there must also be more than one layer of rent to finance each one of them. As result, it is very problematic to decentralize such a financial arrangement through the market mechanism, which is essential for a federal system to be consistent with Tiebout's conceptual setup. Simply put, when the spatial aspect is taken into account, a fully decentralized multilayer government is likely to be non-optimum.

### 2.2.4 Summary

The debate on the issue of small homogeneous jurisdictions (decentralized system) versus large heterogeneous jurisdictions (centralized system) is still an ongoing hot topic both in literature and in practice. In a survey of empirical literature of the Tiebout hypothesis by Dowding et al. (1994), there is mixed but marginal support for the proposition that smaller jurisdictions (i.e. larger number of jurisdictions) have satisfaction levels for some locally provided collective goods.

## 2.3 The Model

This section describes the complete model that will be explored in this dissertation. We will start with a general description of the model and followed by assumptions used in the model and its Pareto optimum characterization.

### 2.3.1 Settings

The model is constructed based on the work of Calabrese et al. (2012): *Inefficiencies from Metropolitan Political and Fiscal Decentralization: Failures of Tiebout Competition* that we discussed in the literature review. The reasons are obvious: (1) it covers case of fiscal decentralization and its impact on decentralization; (2) it covers the case of an urban economy; and (3) it has a strong theoretical framework. In short, the article provide general settings and methodology that fit this dissertation. Consequently, we will follow closely on their model and methodology. However, we will have different objectives compared to Calabrese et al. (2012) and in particularly the model here will have different settings as we will discuss in this section. The main goal in Calabrese et al. (2012) is to find the source of inefficiency of Tiebout competition. In contrast, we use their approach to find the impact of externality and decentralization on social welfare.

In constructing the model first of all we need to define who are the agents in it. There are four agents in the model:

1. households
2. absentee landlords
3. local governments, i.e. city governments
4. a higher government, i.e. a state or a province government

Households chooses cities (jurisdictions) that fit their preferences and consequently maximize their utility. Denoting $\alpha$ as a level of preference and $y$ as households income in monetary value, each household will be characterized by $(y, \alpha)$. Household utility is function of private good consumption $(x)$, housing consumption $(h)$, local public good expenditures $(g)$, and taste parameter $(\alpha)$, formally:

$$U_i = U(x_i, h_i, g_i, g_{-i}; \alpha, \kappa) \tag{2.16}$$

Denote also that $F(y, \alpha)$ is the joint distribution of household type $(y, \alpha)$ and its joint density function is shown by $f(y, \alpha)$. Let $S \equiv [\alpha, \bar{\alpha}] \times [y, \bar{y}] \subset R_+^2$. Household will consume housing in city $i$ as a function of housing price $(p_i)$, local public good expenditures $(g_i)$, non-local public good expenditures $(g_{-i})$, income $(y)$, and taste for local public goods $(\alpha)$. Formally:

$$h = h_d(p_i, g_i, \kappa g_{-i}, y, \alpha) \text{ for all } i \text{ and } (y, \alpha) \tag{2.17}$$

The absentee landlord is assumed to simplify the model and to ensure that the model will be mathematically less complicated without losing generality. However, as in Calabrese et al. (2012), we will take into account the landlords' welfare in the welfare calculation.

Landlords supply housing in jurisdiction $i$ with usual housing supply in which housing supply is a function of housing net price:[21]

$$H_s = H_s\left(\frac{p_i}{1+t_i}\right). \tag{2.18}$$

The model also assumes that the housing market clears, i.e. $h_d = H_s$.

Local government at the city level in city $i$ is assumed imposing a property tax, $t_i$, and a head tax, $T_i$. Albeit a head tax is not realistic in real world, those head taxes are needed in characterizing Pareto optimum and to observe the impacts of decentralization on welfare. The basic intuition is that the head tax is an efficient tax that does not generate deadweight loss. Thus, by imposing head tax it is assumed that there is no distortion in observing the alteration in efficiency as decentralization comes into play.

### 2.3.2 Characterization of Pareto optimum

The model assumes that the household's utility in city $i$ is maximized over housing demand, such that:

$$\begin{aligned} &V^e(p_i, g_i, \kappa g_{-i}, y + r(y,\alpha) - T_i, \alpha) \\ &\quad \equiv \max_h U(y + r(y,\alpha) - T_i - p_i h, h, g_i, \kappa g_{-i}; \alpha) \end{aligned} \tag{2.19}$$

Since the households are completely rational and correctly anticipate all equilibrium values then their housing consumption is equal to the housing demand, $h_d$. Therefore, the solution

[21] Recall that the net supply function is the net output vector that satisfy profit maximization, i.e. $\pi(\mathbf{p}) = \mathbf{p}\mathbf{y}(\mathbf{p})$ in standard notation.

for equation (2.19) is $h_d(p_i, y+r(y,\alpha)-T_i, g_i, \kappa g_{-i}, \alpha)$. If the households imperfectly perceive the benefits generated by other cities' local expenditure, denoted as $\delta$, we may have $\delta \gtreqless \kappa$ then $h_d(p_i, y + r(y,\alpha) - T_i, g_i, \delta g_{-i}, \alpha)$. $\kappa$ is related to *the true benefits* from spillover effects and $\delta$ is related to the households' perceptions on the benefits from spillover effects. In this model, we assume that households are fully rational and forward looking, where $\delta = \kappa$. The maximization problem is:

$$\max_{r(y,\alpha),a_i(y,\alpha),R,T_i,t_i,p_i,g_i,g_{-i}} \sum_{i=1}^{J} \Big[ \int_S \omega(y,\alpha) V_i^e(p_i, g_i, \kappa g_{-i}, y + r(y,\alpha) - T_i, \alpha) \\ a_i(y,\alpha) f(y,\alpha)\, dy\, d\alpha + \omega_R \Big( R/J + \int_0^{p_i/(1+t_i)} H_s^i(z)\, dz \Big) \Big] \tag{2.20}$$

subject to:

$$R + \int_S r(y,\alpha) f(y,\alpha)\, dy\, d\alpha = 0 \tag{2.21}$$

$$\int_S h_d^i(p_i, y + r(y,\alpha) - T_i, g_i, \kappa g_{-i}, \alpha) a_i(y,\alpha) f(y,\alpha)\, dy\, d\alpha \\ = H_S^i(p_i/(1+t_i)) \tag{2.22}$$

$$T_i \int_S a_i(y,\alpha) f(y,\alpha)\, dy, d\alpha + \frac{t_i p_i}{1+t_i} H_S^i(p_i/(1+t_i)) \\ = g_i \int_S a_i(y,\alpha) f(y,\alpha)\, dy\, d\alpha \tag{2.23}$$

$$a_i(y,\alpha) \in [0,1]\ ,\ \sum_{i=1}^{J} a_i(y,\alpha) = 1 \quad \forall (y,\alpha) \tag{2.24}$$

$$\kappa \in [0,1], \tag{2.25}$$

where $p_i$ is the housing price in city $i$, $g_i$ is public good expenditure in jurisdiction $i$, $y$ is household income, $r(y,\alpha)$ is transfers to households based on income and their their preferences, $R$ is transfers to homeowners, and $h$ is housing consumption.

$\kappa = 0$ implies there is no spillover effects and $\kappa = 1$ implies perfect spillover effects meaning public good provision in city $-i$ generates the same levels of benefits to the households both in city $i$ and $-i$.[22]

$N_i \equiv \int_S a_i(y,\alpha) f(y,\alpha)\, dy\, d\alpha$ where $N_i$ is number of households in jurisdiction $i$. Consequently, $N \equiv \sum_{i=1}^{J} \int_S a_i(y,\alpha) f(y,\alpha)\, dy\, d\alpha = \int_S f(y,\alpha)\, dy\, d\alpha$ where $N$ is total number of households in the metropolitan area. Also, denote $\bar{N} \equiv N/J$ and the elasticity of housing supply is $\varepsilon_s^i$, where

$$\varepsilon_s^i = \frac{H_s^{i\prime}}{H_s^i} \frac{p_i}{1+t_i}. \tag{2.26}$$

The Lagrangian function is constructed as follows:

$$\begin{aligned}
\mathcal{L} = & \sum_{i=1}^{J} \left[ \int_S \omega V_i^e(p_i, g_i, \kappa g_{-i}, y + r(y,\alpha) - T_i, \alpha) a_i f\, dy\, d\alpha + \omega_R \left( R/J + \int_0^{p_i/(1+t_i)} H_s^i\, dz \right) \right] \\
& + \Omega \left[ R + \int_S r f\, dy\, d\alpha \right] + \sum_{i=1}^{J} \eta_i \left[ \int_S h_d^i(p_i, y + r(y,\alpha) - T_i, g_i, \kappa g_{-i}, \alpha) a_i f\, dy\, d\alpha - H_S^i \right] \\
& + \sum_{i=1}^{J} \lambda_i \left[ (T_i - g_i) \int_S a_i f\, dy\, d\alpha + \frac{t_i p_i}{1+t_i} H_S^i \right]
\end{aligned} \tag{2.27}$$

[22]Note that $\kappa$ in this model is defined differently from $\kappa$ in Besley and Coate (2003) that we have discussed in the literature review.

First-order Conditions:

$$
\begin{aligned}
\frac{\partial \mathcal{L}}{\partial R} =& \Omega + \omega_R = 0 && (2.28)\\
\frac{\partial \mathcal{L}}{\partial r} =& \sum_{i=1}^{J} \omega U_1^i a_i + \sum_{i=1}^{J} \eta_i (h_d^i)_2 a_i + \Omega = 0 && (2.29)\\
\frac{\partial \mathcal{L}}{\partial a_i} =& \omega V_i^e + \eta_i h_d^i - \lambda_i (T_i - g_i) = 0 && (2.30)\\
\frac{\partial \mathcal{L}}{\partial T_i} =& -\omega U_1^i - \eta_i (h_d^i)_2 + \lambda_i = 0 && (2.31)\\
\frac{\partial \mathcal{L}}{\partial t_i} =& -\omega_R + \lambda_i (1 - t_i \varepsilon_s^i) + \eta_i \frac{(1+t_i)}{p_i} \varepsilon_s^i = 0 && (2.32)\\
\frac{\partial \mathcal{L}}{\partial g_i} =& \int_S \omega U_3^i a_i f \, dy \, d\alpha + \eta_i \int_S (h_d^i)_3 a_i f \, dy \, d\alpha - \lambda_i \int_S a_i f \, dy \, d\alpha = 0 && (2.33)\\
\frac{\partial \mathcal{L}}{\partial g_{-i}} =& \omega U_4^i \kappa + \eta_i (h_d^i)_4 \delta = 0 && (2.34)\\
\frac{\partial \mathcal{L}}{\partial p_i} =& \left[ \eta_i \int_S (h_d^i)_1 a_i f \, dy \, d\alpha - \int_S \omega U_1^i h_d^i a_i f \, dy \, d\alpha \right] \frac{1+t_i}{H_s^i} + \omega_R + \lambda_i t_i (1 + \varepsilon_s^i) \\
& - \eta_i \frac{(H_s^i)'}{H_s^i} = 0 && (2.35)
\end{aligned}
$$

These first-order conditions characterize the Pareto optimum with respect to the set of state variables above.

## 2.4 Numerical Simulations: Welfare change and equilibria

In this section, we will conduct numerical simulations/calibrations for cases of urban areas in the United States and Indonesia. The structure of this section is as follows. First we will discuss the general motivation of this chapter and how it differs compared to the previous research. Section 2.4.1 will discuss the methodology used in the numerical simulation and will be followed by data, assumptions and cases from the United States and Indonesia as archetypes of urban areas in develop and developing countries. The main motivation of this section is to observe different outcomes of decentralized taxation where spillover of public goods is present. The model treats centralized taxation as a *status quo* and calculate the impact of altering from centralization to decentralization on welfare and variables in the equilibrium such as housing price, housing demand and supply, income, population, and tax rates. The main contribution is the consideration of taking into account the spillover into the model.

### 2.4.1 Methodology

The methodology we use is numerical simulation and calibration. As shown in the appendix, we derive the equations used in the computer model based on Calabrese et al. (2012) with necessary adjustments as we have spillovers in the model. A computer model is then constructed to run the numerical simulations and calibrations.[23] The model starts with setting parameters and defining an initial guess for housing prices, local public good provision and taxes. We also set the land size where the urban economy is divided into five jurisdictions where the the central city has the largest area (40 percent) and the rest of the area belongs to four jurisdictions with equal size (15 percent for each jurisdiction). Newton-Rhapson method is then applied to find to a solution for the nonlinear programming.

Once the solution is found we calculate the change in welfare as the government tax policy shifts from centralization to decentralization. Following a standard economic theory this change is calculated using *the compensating variation* (CV). In our model, a positive sign of CV implies that households need to be compensated to reach the same utility level which

[23]This dissertation uses Matlab ver. 7.11(R2010b) to construct and run the simulations (see the appendix for the codes).

means now they have lower utility level in the decentralized regime. On the other hand, a negative sign of CV implies that households are better-off in the decentralized regime.[24] The change in landlords' welfare is also taken into account in the model by estimating the change in rents.

### 2.4.2 Data and utility function

To run the calibration we input some parameters using actual data and results from previous research. The main data used in this model is the mean income that assumed to be normally distributed. The mean income for urban areas in the United States is US$54,710 (based on 14 income classes reported by AHS estimated by Calabrese et al., 2012) and the mean income for an urban area in Indonesia, namely Central Jakarta is US$28,399 (BPS, 2012). The standard deviation of income in the US and Indonesia are 0.88623 and 0.8310 (calibrated), assuming that the United States has higher inequality than Indonesia.[25]

We assume a constant of elasticity substitution (CES) defined as:

$$U_i = [\beta_x x^\rho + \beta_h h^\rho + \beta_{g_i}(\alpha) g_i^\rho + \beta_{g_j}(\alpha)\kappa g_j^\rho]^{1/\rho} \tag{2.36}$$

where $\beta_x$ and $\beta_h$ are the parameters of goods consumption ($x$) and housing consumption ($h$), respectively. $\beta_{g_i}$ and $\beta_{g_j}$ are the parameters for the preferences toward local public

[24] Mas-Collel et al. (1995) defines compensating variation as: *the net revenue of a planner who must compensate the consumer for the price change after it occurs, bringing her back to her original utility,* $u_0$; where $CV = e(p_1, u_1) - e(p_1, u_0)$. In their definition negative CV implies that the consumer is worse-off, and vice versa. Note that we calculate CV from a consumer's perspective, i.e. $CV = e(p_1, u_0) - e(p_1, u_1)$ as reflected in the Matlab code. Thus, in our case negative CV implies that the consumer is better-off, and vice versa.

[25] Based on Gini indexes, Gini coefficients for US (2007) and Indonesia (2009) are approximately at 45 and 36.5, respectively (https://www.cia.gov/library/publications/the-world-factbook/rankorder/2172rank.html).

good in its own jurisdiction ($g_i$) and another jurisdiction ($g_j$). $\kappa$ is some measurement of spillover level. A higher values of $\kappa$ implies a higher degree of spillover. $\beta_x$ is normalized to one and other parameters are calibrated in the model, where $\beta_{g_i} = \beta_{g_j} = \alpha$. We assume that $\alpha$ is log-normally distributed. $\rho$ is the parameter for each term where the elasticity of substitution $\varepsilon_s = 1/(1-\rho)$.

The housing production function is assumed has a Cobb-Douglas functional form $h = L^{\gamma}v^{1-\gamma}$, where $L$ is land input and $v$ is another factor input. By defining a profit function and employing Hotelling's lemma we can obtain the housing supply function as $H_s^i = L_i(p_s^i)^{\frac{1-\gamma}{\gamma}}(\frac{1-\gamma}{w})^{\frac{1-\gamma}{\gamma}}$.

### 2.4.3 A case of an urban economy in the United States

The motivation of this section is to demonstrate the change in the equilibria and social welfare as we take into account spillovers in the model. We will also discuss how the results differ from those of previous literature.

#### 2.4.3.1 Calibrations and results

Table 2.1 presents the parameters used in US case for both head tax and property tax cases. We model a hypothetical metropolitan area which is divided into 5 jurisdictions where the mean income, $\mu_{lny}$, and standard deviation, $\sigma_{lny}$, are in logarithmic forms. Other parameters are set as reported in Table 2.1. The price of input, $w$, is normalized to 1 and $\gamma$ is set to 0.25 to get the supply elasticity ($\theta$) equals to 3.

Table 2.1: Parameters, US

| Parameters | |
|---|---|
| number of communities | 5 |
| $\mu_{lny}$ | 10.5171 |
| $\sigma_{lny}$ | 0.8862 |
| $\beta_h$ | 0.3559 |
| $\beta_x$ | 1 |
| $\rho$ | -0.01 |
| $\gamma$ | 0.25 |
| $\theta$ | 3 |
| $\kappa$ | 0 to 1 |

$\sigma_{lny}$ is calibrated to obtain realistic results from the simulations. $\beta_h$ is calibrated so the households will have income share for housing between 25 and 27 percent and $\beta_x$ is normalized to 1. $\rho$ is the parameter of CES utility function and it is assumed to be equal to $-0.01$.

**Efficient taxation**

We simulate the model with efficient taxation using head tax where the corresponding local government spends exactly the same amount of local revenues to provide local public goods. Table 2.2 shows the equilibrium at the degree of spillover is 0.5.

Table 2.2: Equilibrium at $\kappa = 0.5$ for efficient taxation, US

| Jurisdiction | Price | Local public goods | Head tax |
|---|---|---|---|
| 1 | 11.40 | 1414.45 | 1414.45 |
| 2 | 13.24 | 3113.92 | 3113.92 |
| 3 | 13.79 | 4559.52 | 4559.52 |
| 4 | 14.42 | 6804.09 | 6804.09 |
| 5 | 15.57 | 13385.72 | 13385.73 |

As shown in Table 2.2, jurisdiction 1 (the central city) has the lowest housing price[26], local public goods provision, and head tax at equilibrium. Those three values increase as we move from jurisdiction 1 to jurisdiction 5. In equilibrium, the central city has 53 percent of the population and mostly populated by low-income households with mean income of US$21,500. In contrast, jurisdiction 5 has the smallest proportion of population, i.e. 7 percent and populated by mostly the wealthiest households with mean income of US$203,572 (see Table 2.3).[27]

Table 2.3: Populations and mean incomes (head tax), US

| Jurisdiction | Populations | Mean incomes |
|---|---|---|
| 1 | 0.53 | 21514.20 |
| 2 | 0.16 | 47369.65 |
| 3 | 0.13 | 69464.26 |
| 4 | 0.11 | 103512.53 |
| 5 | 0.07 | 203572.87 |

Table 2.4 reports the housing equilibrium where the housing demand is equal to the housing supply and income share for housing is approximately 25 percent in for all households. As the government taxation policy shifts to the decentralization regime, 28.26 percent of the population are better-off where households gains the benefit by US$1619.08 (3 percent of the mean income).

[26]The housing price is measured in per square foot per year

[27]A comparison between Monte Carlo and quadrature results for community size, mean income, and mean housing demand is shown in appendix A.4.

Table 2.4: Housing equilibrium at $\kappa = 0.5$ for efficient taxation, US

| Jurisdiction | Housing demand | Housing supply | Income share for housing | Land share |
|---|---|---|---|---|
| 1 | 250.29 | 250.29 | 0.2515 | 0.40 |
| 2 | 146.79 | 146.79 | 0.2518 | 0.15 |
| 3 | 165.91 | 165.90 | 0.2519 | 0.15 |
| 4 | 189.89 | 189.89 | 0.2519 | 0.15 |
| 5 | 238.75 | 238.75 | 0.2521 | 0.15 |

Table 2.5: Change in welfare for efficient taxation, US

| $\Delta$Average CV | $\Delta$Rents | $\Delta$Welfare | Better-off (%) |
|---|---|---|---|
| -969.35 | 649.73 | -1619.08 | 28.26 |

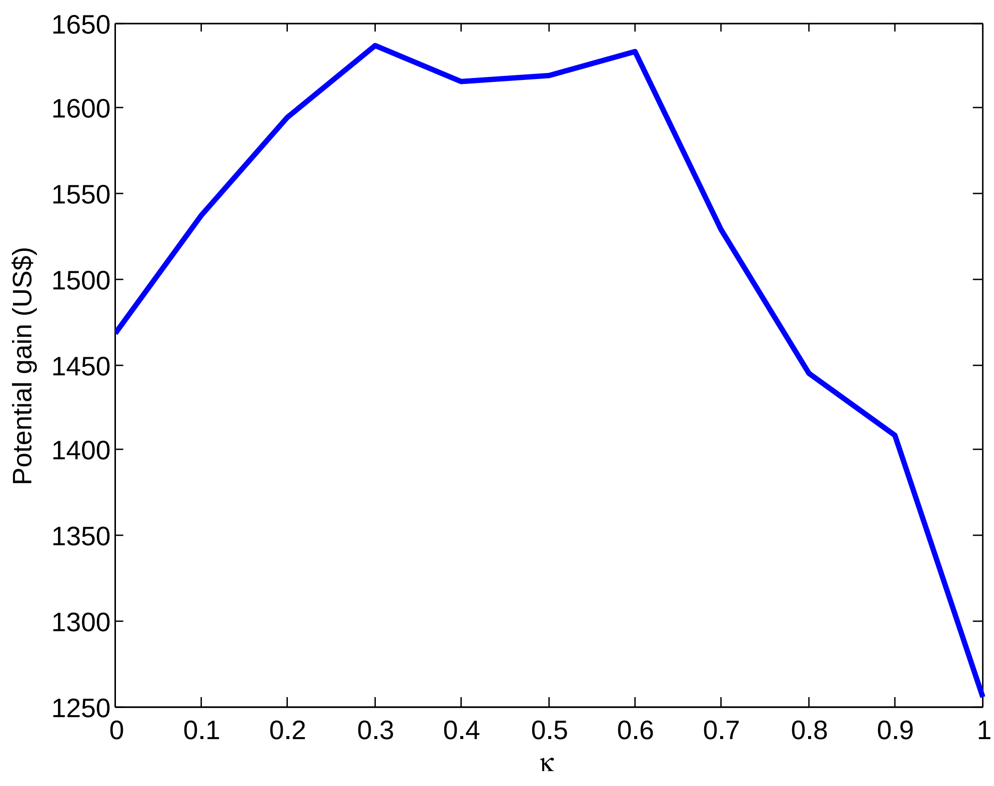



Figure 2.2: Spillover and potential gain of decentralization in head-tax world, US

As we described previously, the main interest of this research is to observe how the degree of spillover will impact the potential gain (of moving from centralized policy to decentralized policy). Figure 2.2 exemplify the impact. The curve suggests that as the spillover increases, generally the potential gain of decentralization decreases. This result support Oates' decentralization theorem (Oates, 1972): decentralization is favorable when spillover among jurisdictions is relatively small. In our case, higher spillovers among jurisdictions will impede the locals' capability to set optimal tax rates. Figure 2.2 shows that when the spillover is relatively low, the benefit from having positive spillovers (from other jurisdictions) dominates the cost of having "non-optimal" public good provisions. However, as the degree of spillover increases, the potential gain of decentralization declines.

### Property taxation

In this section, we discuss the case where governments implements property taxation, as opposed to efficient taxation. As before, we will examine the equilibrium, welfare change, and a relationship between the spillover and the potential gain from the decentralization.

Table 2.6 shows that the housing price, local public good provision, and property tax rate increase as we move from jurisdiction 1 (central city) to jurisdiction 5. The price and local public good provision at the equilibrium are somewhat higher than those in the efficient allocation case. The property tax rates in Table 2.6 are the percentage of annual implicit rents that can be converted to property tax rates with a conversion rate 7-9 percent (Poterba, 1992; Calabrese and Epple, 2010; Calabrese et al., 2012).

Table 2.6: Equilibrium at $\kappa = 0.5$ for property taxation, US

| Jurisdiction | Price | Local public goods | Property tax |
|---|---|---|---|
| 1 | 14.33 | 1916.14 | 0.3559 |
| 2 | 16.24 | 3099.78 | 0.3598 |
| 3 | 17.21 | 3882.77 | 0.3569 |
| 4 | 18.42 | 5035.59 | 0.3545 |
| 5 | 20.82 | 8046.76 | 0.3566 |

In this property-tax setting, the distribution of population is more balanced than that of the efficient allocation case. The central city has the largest population (approx. 40 percent) and the the rest of jurisdictions have fairly equal populations. It replicates the results from Calabrese et al. (2012) when they are assuming there is no spillover in the economy. It is also reported that jurisdictions have stratified mean income where the central city has the lowest mean income (US$ 26,954) and jurisdiction 5 has the highest mean income (US$113,310) in Table 2.7.

Table 2.7: Populations and mean incomes (property tax), US

| Jurisdiction | Populations | Mean incomes |
|---|---|---|
| 1 | 0.39 | 26954.37 |
| 2 | 0.15 | 43692.87 |
| 3 | 0.15 | 54753.88 |
| 4 | 0.15 | 71008.76 |
| 5 | 0.16 | 113309.55 |

Table 2.8: Housing equilibrium at $\kappa = 0.5$ for property taxation, US

| Jurisdiction | Housing demand | Housing supply | Income share for housing | Land share |
|---|---|---|---|---|
| 1 | 197.92 | 199.08 | 0.27 | 0.40 |
| 2 | 108.82 | 107.70 | 0.27 | 0.15 |
| 3 | 129.04 | 129.21 | 0.27 | 0.15 |
| 4 | 159.36 | 159.19 | 0.27 | 0.15 |
| 5 | 228.69 | 228.76 | 0.27 | 0.15 |

Table 2.9: Change in welfare for property taxation, US

| $\Delta$Average CV | $\Delta$Rents | $\Delta$Welfare | Better-off (%) |
|---|---|---|---|
| 195.32 | -59.36 | 254.67 | 0 |

The housing market equilibrium is shown in Table 2.8. We have some discrepancies in the housing equilibrium where housing demand is not exactly equal to the housing supply in each jurisdiction. However, the the discrepancies are relatively very small. Income share for housing is calculated approximately 0.27—which is a realistic number (Quigley and Raphael, 2004).

In this property-tax world society in total is worse-off when the regime is changed into a decentralized taxation. Intuitively, when the degree of spillover is relatively high ($\kappa = 0.5$), the negative impacts of non-optimal taxation and local public good provisions on welfare dominate the benefits on enjoying positive externalities from other jurisdictions. As the tax policy moves to decentralization, the average CV decreases by US\$195.32 and the rents decreases by US\$59.36 (that will reduce landlord's welfare level). Moreover, there is no household will be better-off in this case. In aggregate the society is worse-off by US\$254.67.

The potential cost of decentralization occurs in all level of spillover between 0 and 0.9 as shown in Figure 2.3.[28] The trend of the graph in the property-tax case is unequivocal, as opposed the corresponding figure in the efficient-taxation case. The graph demonstrates that as the spillover increases, the potential loss of decentralization also increases. This result firmly preserves Oates' decentralization theorem (Oates, 1972).

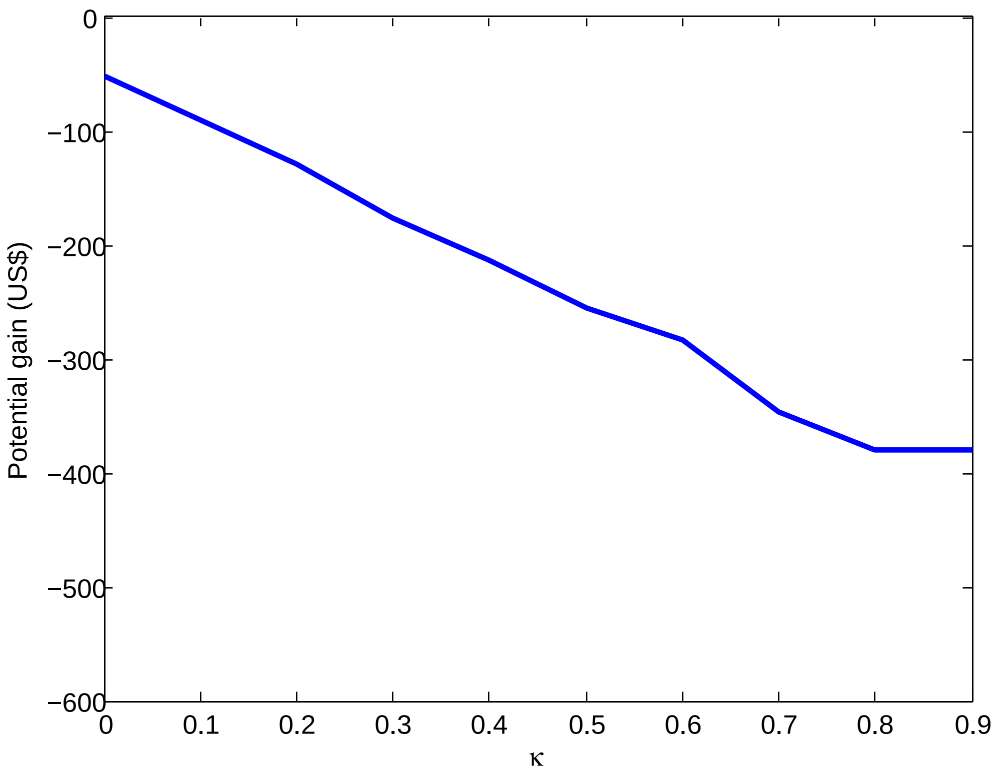



Figure 2.3: Spillover and potential gain of decentralization in property-tax world, US

[28] One of the limitations of the numerical simulation we used in this research is that it is very sensitive to the choices of parameters and initial guess, where some combination of parameters and initial guess might generate indefinite values or even complex numbers. For that reason, we can only have $0 \leq \kappa \leq 0.9$ to depict the impact of spillover on the potential gain of decentralization.

### 2.4.4 A case of an urban economy in Indonesia

#### 2.4.4.1 Calibrations and results

In this section, we will discuss how some degree of spillover ($\kappa = 0.5$) will affect the equilibrium in housing market, local public good provision, and tax rates in a developing country. We will also discuss the impact of spillover on welfare gain of decentralization in both head-tax and property-tax worlds. We use Indonesia as representative of developing countries. Indonesia is fairly new to fiscal and tax decentralization, especially for decentralized property tax.[29] We will also divide the discussion into two section: a case for head tax and a case for property tax. Table 2.10 reports the parameters use for Indonesian case.

Table 2.10: Parameters, Indonesia

| Parameters | |
|---|---|
| number of communities | 5 |
| $\mu_{lny}$ | 9.8614 |
| $\sigma_{lny}$ | 0.8310 |
| $\beta_h$ | 0.3959 |
| $\beta_x$ | 1 |
| $\rho$ | -0.01 |
| $\gamma$ | 0.37 |
| $\theta$ | 1.7 |
| $\kappa$ | 0 to 1 |

$\beta_h$ and $\gamma$ are set to 0.37 and 0.3959 to get housing elasticity approximately 1.7 and housing income share between 27% and 33%. Those values are based on the work of Monkkonen (2013) who studies housing markets in urban areas in Indonesia.

**Efficient taxation**

[29] Indonesia implemented fiscal decentralization in 2000 (Act no.22/1999) and decentralized property taxation in 2011 (Act no.28/2009).

Although head tax is not applicable in many places in the world, including Indonesia, we use this type of tax as a benchmark when the tax does not generate distortion. A head tax will not alter the household behavior and consequently will not create dead-weight loss in the economy. A head tax is an efficient tax for that reason. We start with a head tax in this section and then will be followed a discussion on the case of property tax in Indonesia.

The numerical simulation results in Table 2.11 shows that at $\kappa = 0.5$ housing price, local public good provision, and head tax are stratified in five jurisdictions where the central city has the lowest values and jurisdiction 5 has the highest values of those three variables. The housing price in the central city is US$8.5 and the housing price in the jurisdiction 5 is US$11.39. As expected, all corresponding values in equilibrium are lower than those in the US case as the consequences of lower mean income. However, the distribution of population in the metropolitan area is about fairly identical as reported in Table 2.12. The mean income is only US$12,645.79 in the central city and US$102,237.52 in jurisdiction 5. Housing market is in a equilibrium with income share for housing approximately 27 percent as shown in Table 2.13.

Table 2.11: Equilibrium at $\kappa = 0.5$ for efficient taxation, Indonesia

| Jurisdiction | Price | Local public goods | Head tax |
|---|---|---|---|
| 1 | 8.85 | 809.59 | 809.59 |
| 2 | 9.96 | 1779.07 | 1779.07 |
| 3 | 10.33 | 2447.02 | 2447.02 |
| 4 | 10.82 | 3571.59 | 3571.59 |
| 5 | 11.39 | 6545.00 | 6545.00 |

Table 2.12: Populations and mean incomes (head tax), Indonesia

| Jurisdiction | Populations | Mean incomes |
|---|---|---|
| 1 | 0.60 | 12645.79 |
| 2 | 0.14 | 27849.24 |
| 3 | 0.11 | 38244.28 |
| 4 | 0.09 | 55814.16 |
| 5 | 0.06 | 102237.52 |

Table 2.13: Housing equilibrium at $\kappa = 0.5$ for efficient taxation, Indonesia

| Jurisdiction | Housing demand | Housing supply | Income share for housing | Land share |
|---|---|---|---|---|
| 1 | 230.89 | 230.89 | 0.27 | 0.40 |
| 2 | 107.82 | 107.82 | 0.27 | 0.15 |
| 3 | 115.46 | 115.46 | 0.27 | 0.15 |
| 4 | 125.73 | 125.73 | 0.27 | 0.15 |
| 5 | 138.31 | 138.31 | 0.27 | 0.15 |

Table 2.14: Change in welfare for efficient taxation, Indonesia

| $\Delta$Average CV | $\Delta$Rents | $\Delta$Welfare | Better-off (%) |
|---|---|---|---|
| -419.2141 | -1668.59 | 1249.37 | 30 |

The change in welfare as the regime move to a decentralized tax policy is reported in Table 2.14. Although only 30 percent of households are better-off, the average CV has a positive value meaning, on average, total households is better-off. However, since the rents decreases by a larger value (i.e. landlords are worse-off), the total society (households and landlords) is worse-off. In brief, at $\kappa = 0.5$, in the head-tax world, the decentralization policy has negative impacts in the metropolitan area by US\$1249.37. This result is different from the result we have in the US, where in the head-tax world, decentralization policy potentially provide more benefits to the metropolitan area.

Compare to the US case, the impact of spillover is smoother in this Indonesian case as shown in Figure 2.4. As the curve shown, there is no significant impact on welfare gain when the spillover level is relatively low ($0 \leq \kappa \leq 0.3$). At that range, there is an indication that the costs of having non-optimal public-good provisions compensate the benefits from having positive spillover of local public good. However, the higher kappa result in lower potential gain of decentralization for the households. For instance, at $\kappa = 1$ decentralization will cost the households approximately US$1430.86. In general, we can conclude that decentralized tax policy potentially generates higher costs to social welfare in the urban area in Indonesia and the cost is higher as the degree of spillover increases.

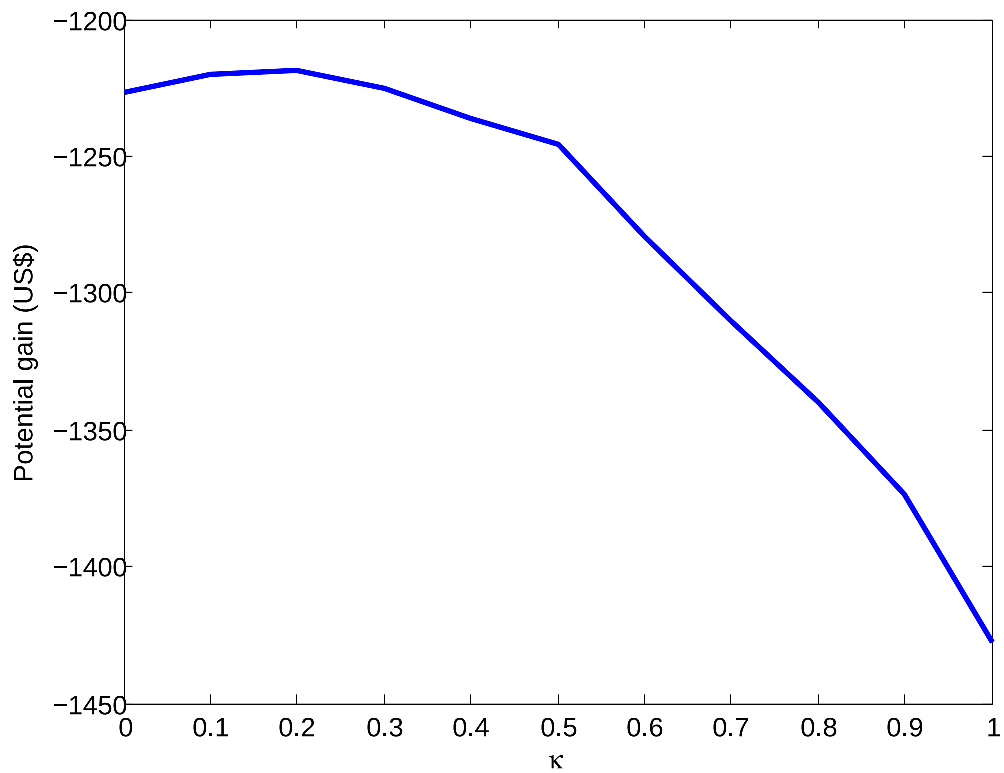



Figure 2.4: Spillover and potential gain of decentralization in head-tax world, Indonesia

### Property taxation

For the purpose of comparison, we intentionally use $\kappa = 0.5$ in Indonesian case where the governments implement property tax. A detail worth mentioning is that the equilibrium housing prices and property tax in this case have different patterns compared to our previous results as shown in Table 2.15. Both prices and property tax rates satisfy housing market condition as shown in Table 2.17. Intuitively, as the local public goods increases from jurisdiction 1 to jurisdiction 5, when households pay a high price, they have to be compensated by a lower property tax rate, and vice versa. For instance, in jurisdiction 5 the housing price (US$10.09) is lower than jurisdiction 2,3, or 4. Consequently, the property tax rate in jurisdiction 5 is relatively high to finance a high provision of local public goods. Using the same conversion as before, the *actual* property tax rates are in the range 0.4-0.9 percent, which is close to the actual property tax rate in Indonesia.[30]

In this property-tax equilibrium, households are concentrated in the central city (61 percent) with mean income US$12,831 vis-à-vis only few households choose to live in jurisdiction 5 (2 percent) with mean income US$133,526.

Table 2.15: Equilibrium at $\kappa = 0.5$ for property taxation, Indonesia

| Jurisdiction | Price | Local public goods | Property tax |
|---|---|---|---|
| 1 | 9.91 | 916.59 | 0.08 |
| 2 | 13.90 | 1579.07 | 0.08 |
| 3 | 11.50 | 2047.02 | 0.09 |
| 4 | 11.38 | 3571.59 | 0.09 |
| 5 | 10.09 | 6545.00 | 0.11 |

[30]Based on Act no.12/1994, property tax rate in Indonesia is 0.5 percent (http://www.tarif.depkeu.go.id/Bidang/?bid=pajak&cat=pbb).

Table 2.16: Populations and mean incomes (property tax), Indonesia

| Jurisdiction | Populations | Mean incomes |
|---|---|---|
| 1 | 0.61 | 12830.81 |
| 2 | 0.23 | 32101.17 |
| 3 | 0.08 | 51520.81 |
| 4 | 0.05 | 76801.40 |
| 5 | 0.02 | 133525.54 |

Table 2.17: Housing equilibrium at $\kappa = 0.5$ for property taxation, Indonesia

| Jurisdiction | Housing demand | Housing supply | Income share for housing | Land share |
|---|---|---|---|---|
| 1 | 231.67 | 231.98 | 0.29 | 0.40 |
| 2 | 156.29 | 156.70 | 0.29 | 0.15 |
| 3 | 111.31 | 111.06 | 0.29 | 0.15 |
| 4 | 109.08 | 109.09 | 0.29 | 0.15 |
| 5 | 86.096 | 86.032 | 0.29 | 0.15 |

Table 2.18: Change in welfare for property taxation, Indonesia

| $\Delta$Average CV | $\Delta$Rents | $\Delta$Welfare | Better-off (%) |
|---|---|---|---|
| -1847.46 | -1371.48 | -475.985 | 100.00 |

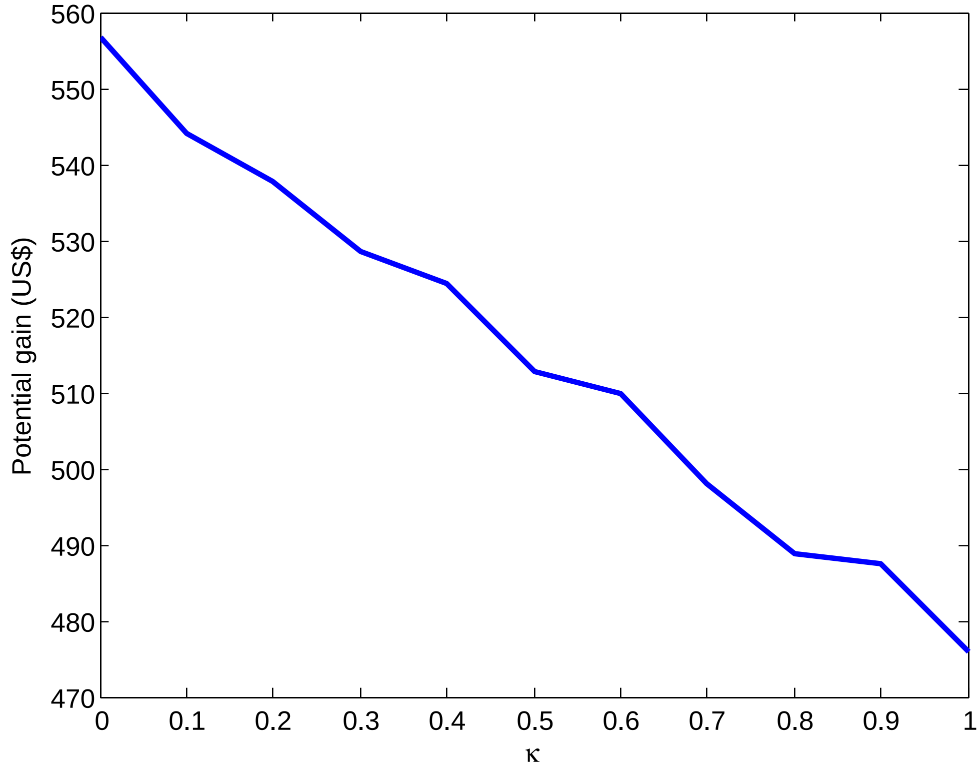



Figure 2.5: Spillover and potential gain of decentralization in property-tax world, Indonesia

In the property-tax world at $\kappa = 0.5$, the change in average CV is -US$1847 meaning the average households are better-off. In contrast, landloards are worse-off since the rents decline by US$1371. As a result, there is net potential gain from decentralized taxation as much as US$476. In this specific setting all households enjoy the benefit of decentralization.

The impact of spillover level than can be concluded from Figure 2.5. The graph demonstrates that as $\kappa$ increases from 0 to 1, the potential gain from decentralization decreases. The graph demonstrates the same pattern as what we have in US case with opposite sign of the potential gain from decentralization. However, the general conclusion remains the same: decentralization is more favorable when there is no spillover or when spillover is relatively low. Again, this result substantiates the idea of Oates' decentralization theorem.

### 2.4.5 Summary

We evaluate the potential gains and costs of decentralized tax policy in urban economies when the spillovers are present. Using degree of spillover equal to 0.5, we find that the answer depends on the spillover level, income level, and type of taxation we have in the urban economy. Using metropolitan areas in the United States as benchmark, in an efficient-tax setting, decentralization will benefit the urban area with relatively high mean income such as Boston metropolitan area in the United States. In contrast, the same setting will generate costs to a metropolitan area that has a relatively low mean income (compared to metropolitan area in the US), such as Jakarta, Indonesia. In the property-tax setting, the results reverse. Decentralization benefits the urban economy when the mean income is relatively low and generates costs to the urban economy when the mean income is relatively high.

However, our focus is more on the impact of spillover on the welfare in lieu of the magnitude of the potential benefits or costs. We find that for all cases, the potential benefits decreases as the degree of spillover increases, albeit there is an anomaly in the efficient-tax setting for low degrees of spillovers. From cost perspective, the potential cost of decentralization increases as the degree of spillover increases. Simply put, we favor lower spillover to gain more from decentralized tax policy. Our results resonate the Oates decentralization theorem where decentralization is favorable when the spillover level is low.

## 2.5 Numerical Simulations: Income level, inequality, and welfare change

In this section, we discuss two main issues for robustness and further conclusions. First, how different income level will affect the welfare gain from decentralization at different levels of spillover. Second, how inequality will effect welfare gain from decentralization at different levels of spillover.[31]

### 2.5.1 Income and welfare change

Figure 2.6 presents a relation between three vectors: the mean income in lognormal distribution ($\mu_{ln_y}$), spillover ($\kappa$), and the potential gain of decentralization.[32] The 3-D surface shows an indication that an economy with higher mean income will gain more benefit from decentralized tax policy at various level of spillover level, *ceteris paribus.* Interestingly, when the mean income is relatively low, a higher spillover level will decrease the benefit from decentralization. In contrast, when the mean income is relatively high, a higher spillover level will increase the benefit from decentralization. Intuitively, holding the same standard deviation of income, using a higher mean income in the model will *generate* wealthier households in the same interval of income distribution that makes the proportion of the tax for each household relatively lower. As a result, the cost of having non-optimal public good provision is getting smaller while holding the benefit from enjoying the positive spillover constant.

[31] We use efficient tax to illustrate this case. Relatively small variation in parameter(s) might generate disequilibria. The calibration shows that the property-tax model is more sensitive to the change of parameters such as mean income or standard deviation of income. See footnote 28.

[32] Using Heien (1968) where $\mu_y = \exp(\mu_{ln_y} + 0.5\sigma^2_{ln_y})$ the actual range is approximately between US\$35,000 and US\$80,000.

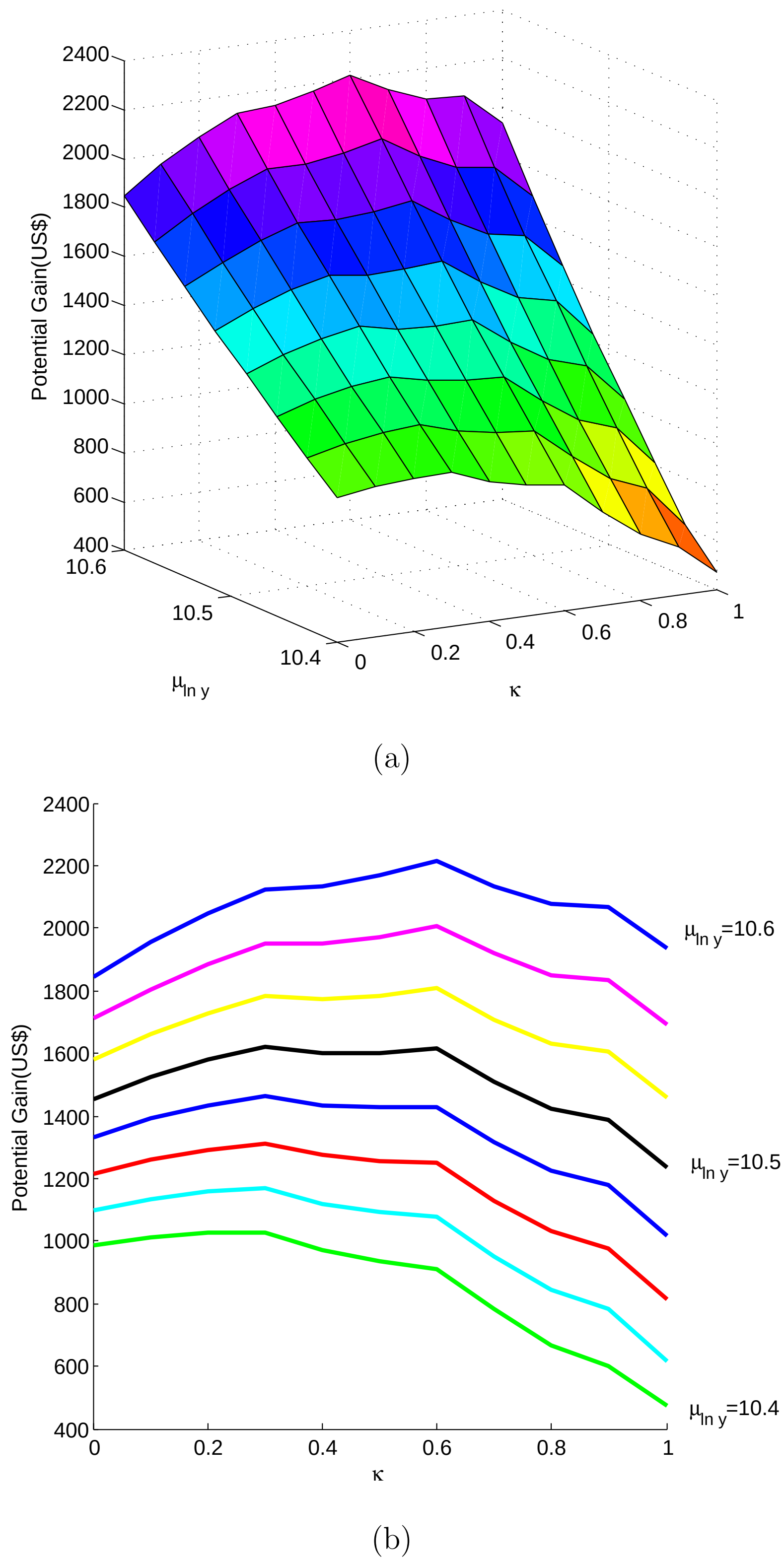


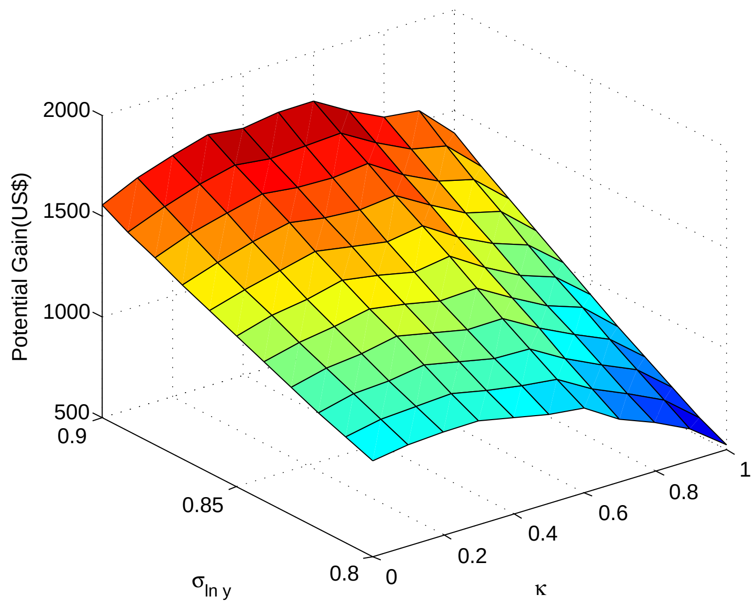

Figure 2.6: Income level and welfare change at various degrees of spillover, US

### 2.5.2 Inequality and welfare change

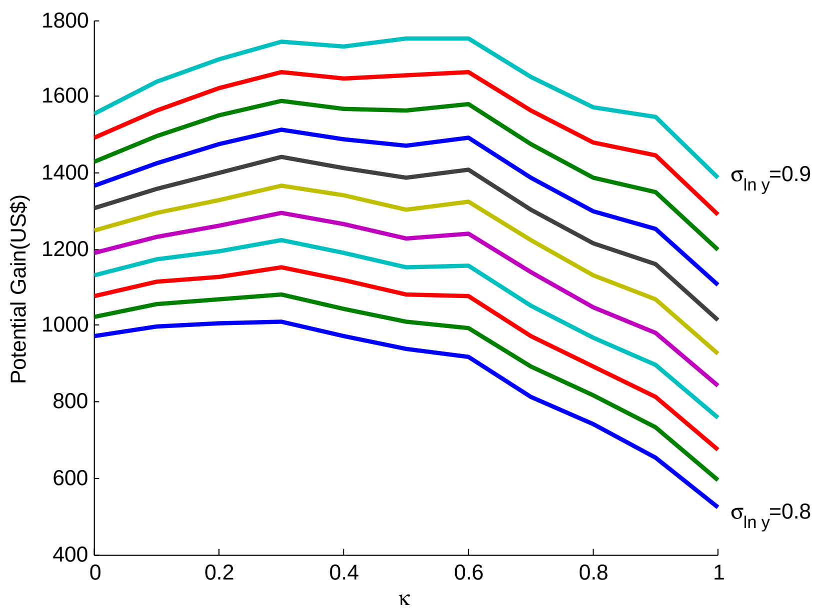
To observe the impact of inequality on welfare gain of decentralization, we use standard deviation of the the lognormal distribution as a proxy of inequality. In this case, a higher the standard deviation $\sigma_{lny}$ implies a higher inequality.[33]

Figure 2.7 demonstrates that as inequality increases, the potential gain of decentralization is linearly higher. This pattern is consistent in different levels of spillover. However, Figure 2.7 also shows that when the inequality is high, the impact of spillover level on the benefit from decentralization become more inconclusive, at least visually.

[33] We calibrate the value of $\sigma_{lny}$ by taking into account the discrepancy between inequality in US and Indonesia. However, it is worth noting that some development economists attempt to relate the Gini coefficient and lognormal distribution, e.g., Chotikapanich (1997) formulates a relation between the Gini coefficient and lognormal as $G = 2\Phi(\frac{\sigma_{lny}}{\sqrt{2}}) - 1$ where $G$ is Gini coefficient and $\Phi$ is standard normal integral up to the ordinate $\frac{\sigma_{lny}}{\sqrt{2}}$.




(a)




(b)

Figure 2.7: Inequality and welfare change at various degrees of spillover, US

### 2.5.3 A higher degree of spillovers and switching to centralization

Our previous results show that: as the degree of spillovers increases reasonably large, decentralization will be less preferable—supporting Oates decentralization theorem (1972). As an extension, the model also enables us to investigate the degree of spillovers, $\kappa$, that will turn the decision on tax policy from decentralization to centralization by increasing the value of $\kappa$ until the potential benefit from decentralization is zero. The results will not change our previous conclusions about how spillover will impact the potential benefit of decentralization but they will provide information about the degree of spillovers when centralization is more beneficial than decentralization in the context of countries at different points in the development process.

Figure 2.8 and 2.9 show cases for the head-tax case in the United States and the property-tax case in Indonesia.[34] Figure 2.8 shows that, in the US head-tax case, as the degree of spillovers higher than 1.6, centralized-tax policy will be more beneficial for the society than decentralized tax policy. This means that centralized-tax policy is better than decentralized-tax policy when the total benefit enjoyed by households from other jurisdictions is 1.6 times higher than total benefit enjoyed by household in the corresponding jurisdiction. Figure 2.9 provides information about Indonesian property-tax case where centralization will be beneficial when the degree of spillovers is very high ($\kappa > 6.8$). This means that the total benefit enjoyed by households from other jurisdiction is 6.8 times higher than the total benefit enjoyed by households in the corresponding jurisdiction.

From policy perspective, the very high degree of spillovers necessary for the benefit

[34]Both the property-tax case in the United States and the head-tax case in Indonesia favor centralization even when $\kappa = 0$ (the potential gain of decentralization is already negative when $\kappa = 0$).

of decentralization to hit zero in both cases above implies that decentralization is better than centralization for these cases since it is unlikely that spillovers would reach these levels. Naturally, when the degree of spillover is high enough, for example larger than one, any degree of rivalry and exclusiveness may induce the local government to set a price to limit the number of consumers of the good. If instead, local public goods are non-rivalrous and non-excludable, the policy option left for the government is centralizing taxation.

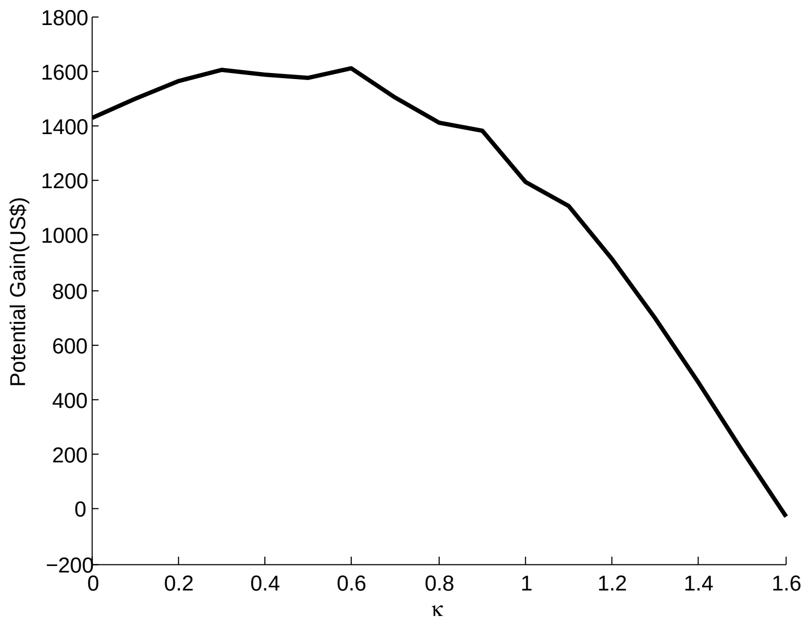



Figure 2.8: Potential gain from decentralization in the head-tax world, US

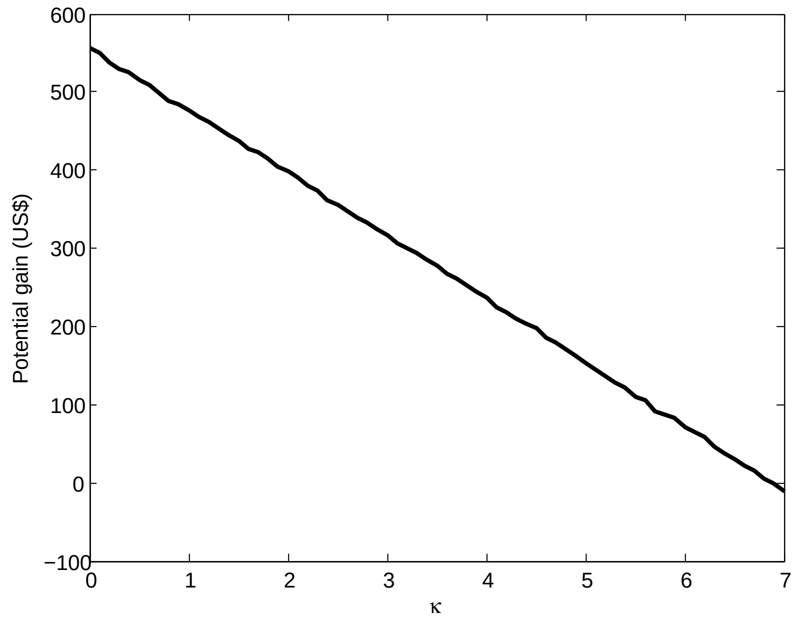



Figure 2.9: Potential gain from decentralization in the property-tax world, Indonesia

### 2.5.4 Summary

We analyze the impact of moving from centralization to decentralization in an urban economy when spillover is present for different mean income and different level of inequality. We find that the impact of decentralization at different level of spillover will depend on the income level, where a higher income level will gain more benefits as the spillover increases. In contrast, a relatively low income level will have lower benefits from spillover. This notion suggests that the impact of decentralization and spillover on welfare will depend on the income level.

In addition, we find that different levels of inequality have different impact on the benefit from decentralization. At a certain degree of spillover, an urban area with higher inequality will enjoy more benefit than an urban area with lower inequality. This result favors the idea that decentralization is needed most when the economy has heterogeneous population or jurisdictions.

Extending the degree of spillovers above the level of $\kappa = 1$ from the baseline specifications also provides us with intuition how high spillovers will switch the policy from decentralization to centralization for the US head-tax case and the Indonesia property-tax case. We find that the spillovers need to be very high (such that the benefit to those outside the community exceeds that for those within the community) to be considered as a reason to move from decentralization to centralization in both cases.

### 2.6 Conclusions and policy implications

This research has inquired a question about the impact of fiscal decentralization on welfare both in developed and developing countries, especially in tax policy where spillovers of public goods among jurisdictions are present. Two types of taxation are used in this research, i.e. efficient tax and property tax. The theoretical framework of the model is built in Section 2.3 to support the construction of numerical simulation in Section 2.4. Estimations has been made to gauge whether the benefits from fiscal decentralization (as in Tiebout equilibrium literature) dominates the costs (as in the tax-competition literature). Moreover, as the main contribution, this research takes into account the impact of spillover on the benefits of decentralization in the metropolitan area as discussed in Oates' decentralization theorem.

The results from the numerical simulation shows that spillover level does affect the potential gain of decentralization. Unexpectedly, in the case of efficient taxation, the impacts of spillover are ambiguous. Relatively low level of spillover has a positive impact on the potential gain of decentralization. In this case, the benefit from consuming local public goods in different jurisdictions dominates the cost of having non-optimal tax rate and non-optimal public good provision. However, as the degree of spillover increases, the potential gain of decentralization diminishes in both cases of developed and developing countries. Moreover, in the US head-tax case and Indonesia property-tax case, a very high degree of spillovers is necessary to switch the policy from decentralization to centralization.

On the other hand, the impact of spillover level in property-tax case is unequivocal. The numerical simulation shows that the spillover level does have negative impact on the potential gain from decentralization. As the degree of spillover increases, the potential gain of decentralization decreases in both developed and developing countries.

Overall, the results substantiate Oates' decentralization theorem where decentralization is more beneficial when spillovers among jurisdictions are relatively low. However, we have to be aware that the impact of spillover level on the potential gain of decentralization varies across different levels of income vis-à-vis income inequality. The result shows that a metropolitan area with a relatively low mean income will be negatively affected by spillover, and vice versa. The numerical simulation also shows that a higher level of inequality and income amplify the benefits of decentralization. Although one might consider that this result is politically incorrect, it illustrates that developed countries—that generally have higher income inequality—potentially gain more benefits from decentralization.

From policy perspective, the results can provide some insights of decentralizing tax policy. First, as a contrast to what we normally believe, the policy makers should be aware that decentralization may generate loss to the society. It is very important that government needs to be very careful in determining the type of taxation to decentralize. One important factor is the level of spillover among jurisdictions, where a different level of spillover will have different impacts on the welfare. This suggests that some taxes are better to remain under centralized policy especially when the spillovers are very high as we discuss in Section 2.5.3. A careful measure is important to define whether a certain tax should be decentralized or centralized. Second, as the consequence of spillover, coordination among jurisdictions is very important to minimize the negative impact of spillover in the metropolitan area. Consolidating the resources and coordination in choosing the type and level of public good provision are potentially beneficial to all jurisdictions in the metropolitan area and will minimize the negative impact of spillover and potential costs of duplication.

At least there are four potentially appealing topics for future research related to this discussion. First, a research question about how different type of taxation, such as income tax, may have impact on the equilibrium and the social welfare. Second, a question about the robustness of the model to different number of jurisdictions[35] and the assumption of land distribution, as well as the optimal number of jurisdictions in decentralized policy. This question is relevant in developing countries that are new to decentralization, such as Indonesia, where some jurisdictions have been proposing to have territorial reforms. Three, whether that higher-income households enjoy more benefits from decentralized-tax policy as indicated in Section 2.5 will be an intriguing topic for further research. It is interesting to investigate whether, within community, higher-income households are better-off than lower-income households as decentralized-tax policy is implemented by the government. Four, taking into account administrative costs in implementing centralization versus decentralization could be examined empirically and might provide additional insights.

[35] Calabrese et al. (2012) show that their model is robust to different assumptions regarding the number of jurisdictions.

CHAPTER 3

# Optimal Income-Sales Taxation, Tax Evasion, and Uncertainty

## 3.1 Introduction

### 3.1.1 Background

In the history of modern optimal tax theory from Ramsey (1927) on optimal commodity taxation to Saez (2001) on using elasticity to derive optimal income taxation, not much research incorporates tax evasion and uncertainty into optimal taxation models. Moreover, most analysis focuses on indirect taxes alongside direct taxes rather than on finding the optimal tax mix which is more practical politically in the real world (Boadway et al., 1994). This is in line with a statement in Bradford and Rosen (1976),where there are two different problems have been studied in the modern optimal tax theory: finding an optimal set of commodity taxes and finding the optimal degree of income taxation.[1]

This research examines the optimal taxation in the presence of tax evasion and uncertainty with focus on income-sales tax mix where the model is built based on Nava et al. (1996), Boadway et al. (1994), and del Mar Racionero (2001). Similar to them, the model utilizes a standard two-class economic model of optimal taxation where the government cannot observe the agents' ability but can observe income. Also, the government uses both non-linear income and linear commodity taxes. By adopting Nava et al. (1996), we deal with a two-class economy with two types of labor and two commodities and also taking into account the existence of under-reported income (tax evasion) and uncertainty. Intuitively,

[1] Bradford and Rosen (1976) assumes that the revenue system is based on income rather than commodity taxation.

by employing a balance budget constraint in the model, income tax will be higher as tax evasion in the economy increases. In contrast to Cremer and Gahvari (1995), this paper incorporates the uncertainty of being detected from hiding some income to gain benefits from lower tax payment, rather than the uncertainty of the wage. The cost function is adopted from Boadway et al. (1994)[2]. We assume that the government implements balanced-budget policy.

To answer the research questions, this research utilizes some standard assumptions (see Section 3.3.1) and we only focus on certain variables in a system that we build as as simple as possible. We believe this type of setting can provide intuitive results. Adding more complexity will only generate a model with non-intuitive and mathematically complex results and very likely only add small contributions to our understandings that are not the main purpose of this research. Simply put, the main motivation of this research is to enrich the existing model by adding more realistic variables with intuitive results. For instance, the degrees of tax evasion in developed and developing countries are very likely to be different as developed countries in general have better tax systems and institutions. It is expected that the results can explain the real world and also contribute to the theory, at least by taking into account uncertainty and tax evasion.

### 3.1.2 Research questions

Specifically we will answer four main research questions. Subject to a government revenue constraint:

1. How will uncertainty affect the optimal income and sales tax rates?

[2]Boadway et al. (1994) defines the cost of hiding one dollar of income is $G(1-\sigma)$, where $\sigma$ is the proportion of reported income. $G(.)$ is assumed strictly convex and non-negative.

2. How will tax evasion affect the optimal income and sales tax rate?
3. What factors will increase the optimal income and sales tax rate?
4. What factors will decrease the optimal income and sales tax rate?

### 3.1.3 Research outline

The essay is organized as follows. The following section will discuss literature review on taxation that will cover common approaches in optimal taxation theory. The discussion will be focused on the models, assumptions, results, and conclusions of the existing literature. The section is then closed by a summary leading to the theoretical model in section 3.3.

Section 3.3 focuses on the process of building the model. It will be started by discussing the original model and follow by the discussion how the model will change when we add tax evasion and uncertainty into the model. We put the details of the algebra in the appendices. The final results in this chapter is the optimal income tax rates for both high-type and low-type labors and also the optimal sales tax rates that directly comparable to Nava et al. (1996) which do not have tax evasion and uncertainty in their model. Related to our results, Nava et al. (1996) find that if a private good is a substitute for leisure, then it is optimal to subsidize the high-type agent's income at the margin (negative income tax).

The results in section 3.3 then will be discussed in section 3.4 along with the conclusions and the policy implications of this research.

## 3.2 Literature Review

The standard framework for optimal taxation postulates that the tax system is designed to maximize a social welfare function subject to specified constraints. The most common constraint is that the government should meet a certain (or minimum) level of revenue. The question is then what types of commodities should be taxed? How progressive or regressive should the income tax be? Those normative questions have been the subject of key debate by the most of prominent economists since Adam Smith up to today (Newbery and Stern, 1987).

For the purpose of this research, we will discuss the modern theories (utilitarian based) that are characterized by welfare maximization problem that begins in 1927 by Frank Ramsey.[3] Modern theory of public economics utilizes the fundamental theorems of economics: (1) competitive equilibrium is Pareto efficient and (2) a Pareto-efficient outcomes can be achieved as competitive equilibrium if prices are set appropriately and lump-sum incomes are allocated so that each individual can buy the consumption bundle given in the allocation at the prices that will prevail (Newbery and Stern, 1987). This section will discuss related literature review that provide underpinnings to the theoretical framework to construct the model and to support the analysis.[4]

### 3.2.1 Commodity taxation

Ramsey (1927) shows that to reach maximum welfare that meet required government revenue from commodity taxes, the proportional change in demand should be the same for all

[3]In the nineteenth century, the discussion was focused on conceptual framework such us ability to pay and tax base (Newbery and Stern, 1987). We may also find non-welfarist approaches to optimal taxation such as behavioral economic approach (see Kanbur et al., 2006) and macroeconomic or economic growth approach (see Barro and Sala-i Martin, 1992).

[4]Another factor such as institutions is not discussed in this research. Readers who are interested in this area may refer to Slemrod (1990) and Alm (1996).

commodities that give us the following famous result:

$$\sum_i T_i S_{ik} = b x_k, \qquad k = 1, ..., m \tag{3.1}$$

where $T_i$ is per unit tax of good $i$, $x_k$ is the quantity of good $k$, $S_{ik}$ is the elasticity for good $i$ with respect to the price of good $k$ and $b$ is some constant.

In the elasticity term then we have a familiar result known as *inverse elasticity rule* as follows:

$$t_r = \frac{d}{\varepsilon_{rr}} \tag{3.2}$$

where $t_r$ is the tax rate ($= T_r/P_r$), $d$ is a constant, and $\varepsilon_{rr}$ is the elasticity of uncompensated demand for good $k$. This result has an intuitive interpretation: the imposition of a tax on an inelastic good will have smaller effects on consumer behavior than that on elastic good, which consequently result in smaller deadweight loss. Equity is not an issue here since Ramsey assumes all individuals are identical.

### 3.2.2 Income taxation

Another important work on optimal tax policy is from Mirrlees (1971). Contrast to Ramsey, Mirrlees (1971) formalizes a social planner's problem with unobserved heterogeneity of individuals. In his paper, Mirrlees (1971) assumes both log normal and Pareto distribution of earning abilities and focuses on nonlinear income taxation. The question is about how to obtain an income tax schedule (tax function) that maximizes the social welfare function (sum of all individuals' utilities in the economy) subject to the revenue constraint. Mirrlees assumes that the utility function is in the form of a Cobb-Douglas utility function and the main result of his model are: (1) the marginal tax rate should be between zero and one; (2) the marginal tax rate for the person with the highest earnings should be zero; and (3) if the person with the lowest wage is working at the optimum, then the marginal tax rate that the person faces should be zero. Using a marginal tax rate, tax payers are classified into tax brackets, which determine which rate taxable income is taxed at. As income increases, it will be taxed at a higher rate than the first dollar earned.

Other than its complexity, one of Mirrlees results' weaknesses is the unweighted sum of individual utilities (Bradford and Rosen, 1976). This implies that the marginal contribution of a rich individual to the social welfare function is the same as a poor one. Moreover, Saez (2001) points out the limitation of Mirrlees' approach in application for three reasons: (1) the zero marginal rate at the top of income is a very local result; (2) based on his numerical simulations, the tax schedules are easily affected by the form of the utility function; and (3) the Mirrlees' result does not change the way the public economists think in practice about the efficiency-equity trade-off where the discussion still focuses on the efficiency side rather than the equity side. To have more applicable result, Saez (2001) using elasticities to derive the optimal income tax rates. By utilizing the elasticities, he is able to performed a numerical simulation using empirical income distributions. His simulation results suggest that the marginal tax rates labor income should be between 50%-80%.

### 3.2.3 Direct-indirect taxation

We now turn to direct-indirect taxation that will be discussed mostly based on the work of Atkinson and Stiglitz. Using one representative consumer facing a set of commodity taxes, Atkinson and Stiglitz (1972) illustrates that if goods and leisure are separable, the optimal commodity tax structure is uniform (or the optimal structure is identical to a proportional income tax). They also show that when the utility function is separable between labor and all commodities, no indirect taxes are needed. Consequently, the Ramsey result is only relevant where there are constraints to implement income taxation (Atkinson and Stiglitz, 1976). Another important theoretical result from Atkinson and Stiglitz is: *if the government only focuses on efficiency, it may use only direct taxation and it would take the form of a poll tax*

However, the result in Atkinson and Stiglitz (1972, 1976) no longer holds if the model includes the production side of the economy (Naito, 1999). In his paper, he shows that if the nonlinear income tax system is implemented for redistribution purpose, the introduction of distortions in the public sector can be a Pareto improvement (in contrast to Diamond and Mirrlees, 1971).[5]

[5] In other word, the public sector is not necessarily efficient. The intuition is that by hiring more unskilled workers, the government will generate less unskilled workers in the private sector and indirectly redistribute income from skilled to unskilled workers (Naito,1999).

The social planner's problem used in Naito (1999) is as follows:

$$\max_{L^s, L^u, x^s, x^u} \quad V^s(v(p(.)+t, x^s), L) \tag{3.3}$$

subject to (3.4)

$$V^u(v(p(.)+t, x^u), L^u) \geq \bar{U}^u \tag{3.5}$$

$$V^s(v(p(.)+t, x^s), L^s) \geq V^s(v(p(.)+t, x^u), \Omega(.)L^u) \tag{3.6}$$

$$V^u(v(p(.)+t, x^u), L^u) \geq V^u(v(p(.)+t, x^s), \frac{L^s}{\Omega(.)}) \tag{3.7}$$

$$w_s(p(.))L^s + w_u(p(.))L^u + tD(p(.)+t, x^s) + tD(p(.)+t, x^u)$$

$$\geq x^s + x^u + w_s l_q^s((1+\zeta)\Omega(.)) + w_u l_q^u((1+\zeta)\Omega(.)) \tag{3.8}$$

where $p(.) = p(L^s, L^u, x^s, x^u, t, \zeta)$ and $\frac{w_u}{w_s} \equiv \Omega(p(.))$ and $\zeta$ are given. For $i = \{u, s\}$ where $u =$ unskilled and $s =$skilled, $V^i$ is total indirect utility function for labor $i$, $v$ is individual indirect utility function, $p$ is the price, $t$ is the commodity tax on good 2 (no tax on good 1), $x^i$ is income for labor $i$, $L^i$ is the labor supply of labor $i$, $\bar{U}^u$ is the minimum utility of unskilled labor, $w_i$ is the wage of labor $i$, $D$ is the conditional demand function of good 2, $l_q^i$ is the labor used to produce public goods, and $\zeta$ is the degree of distortion in public sector.[6]

Naito (1999) also demonstrates using the model that introducing some distortion in the public sector can be a Pareto-improvement through relaxing the incentive problem in income redistribution that emerges under asymmetric information.

To summarize the debate on the optimal taxation and direct-indirect taxes, Atkinson (1977) is a good place to start. In his paper, Atkinson contrasts and compares previous

[6]If $\zeta = 0$ then the public sector is efficient, meaning the marginal rate of transformation between two types of labor is equal to that in the private sectors.

research on optimal taxation (both direct and indirect taxation) by comparing two main views: (1) desirable balance which appears to find favors with policy makers where direct taxation is used for equity purpose while indirect taxation is used for efficiency purposes; and (2) direct taxation is superior on both equity and efficiency purposes.

Comparing four models from previous research, Atkinson summarizes that if all consumers are identical then the poll tax is efficient. However, if the consumers are not identical (by varying wages) the optimal tax depends on goods consumed which will depend on marginal value of income. By considering distribution, redistribution with indirect taxes supports heavier taxes on luxury goods, albeit this is not always the case where he shows that in linear expenditure system the optimum indirect tax would be the same tax rates for all commodities. Intuitively, the progressiveness of commodity tax exactly offsets the regressiveness of the lump-sum tax.

Stiglitz (1982) works on the Pareto efficient taxation by assuming two-class economy (high and low abilities) and incorporating self-selection constraints—also known as the incentive compatibility constraints—in the planner's maximization problem. The formal representation of the model is as follows:

$$\max_{C_1,C_2,Y_1,Y_2} V^2(C^2, Y^2) \tag{3.9}$$

subject to: (3.10)

$$V^1(C^1, Y^1) \geq \bar{U}^1 \tag{3.11}$$

$$V^2(C_2, Y_2) \geq V^2(C_1, Y_1) \tag{3.12}$$

$$V^1(C_1, Y_1) \geq V^1(C_2, Y_2) \tag{3.13}$$

$$R = (Y_1 - C_1)N_1 + (Y_2 - C_2)N_2 \geq \bar{R} \tag{3.14}$$

where $C_i$ and $Y_i$ are the *i-th* individual's consumption and income, respectively. $V^i$ is *i-th* individual's indirect utility, $N_i$ is the number of individuals of type i, $R$ is government revenue, $\bar{R}$ is the revenue requirement. The result for this planner's problem will depend on the assumption on whether the self-selection constraint is binding or not. In the normal case where the self-selection constraint for high type is binding and the self-selection constraint for low type is not binding, he found that the marginal tax rate for the low type is positive and the marginal tax rate for the high type is zero as demonstrated in Mirrlees (1971).

In the existences of both income and commodity taxation, by modifying the government budget constraint, Stiglitz (1982) concludes that if the leisure and goods are separable, there should be no commodity taxation (only a income tax is needed). This notion supports the result in Atkinson and Stiglitz (1976) as we discussed previously. The intuition is that with the separability, if those two types have the same indifference curve between two commodities, one cannot use differential taxation as a basis of separation.

Boadway et al. (1994) argues that in any case the existing theory does not tell anything about the mix of direct and indirect taxes. He points out that the existing theory only has two type of results when the government wants to impose both direct and indirect taxes: (1) income taxation is enough or (2) the tax mix is indeterminate since the optimal tax structure can be achieved by an arbitrary combination of direct-indirect taxation. In order to capture this issue, Boadway et al. (1994) uses an analysis based on the concept that different taxes will have different tax evasion features where the incentive to evade will depend on the marginal tax rates. In the world of two types of individuals, two consumption goods, a direct-indirect tax mix, and the possibility of hiding some proportion of income along with the concealment costs, Boadway et al. (1994) shows that in the existence of tax evasion the incompatibility constraint play an important rule in the design of tax system. This notion bring us to the model that we will construct in Section 3.3.

As far as this research is concerned, Nava et al. (1996) shows that the good complementary with leisure should be taxed higher. It is also implicitly assumed in their paper that all households' income are fully reported which is not always true in the real world. It brings us to the notion that including the possibility of tax evasion along with its cost will change the results and provide a more *realistic* model. Unlike them, we do not focus on public expenditure policy.[7] We simply focus on the optimal taxation for both income and sales taxes.

[7] To deal with public expenditure policy, Nava et al. (1996) modify the standard Samuelson rule by adding terms related to the self-selection constraint and to the revenue of indirect taxes.

### 3.2.4 Summary

The literature review section discusses some modern approaches that has been used to define an optimal taxation theory. From the inverse elasticity rule from Ramsey (1927) to using elasticity to derive optimal income taxation from Saez (2001), there are various approaches that provide a broad range of understanding on optimal taxation.

The optimal taxation theory focuses on three cases: (1) linear commodity taxation as we see in Ramsey (1927); (2) non-linear income tax as in Mirrlees (1971); and (3) direct-indirect taxes as in Atkinson and Stiglitz (1976) and Stiglitz (1982). It is the direct-indirect taxation that becomes main interest in this research. Many papers have been written based on Atkinson and Stiglitz (1976), including Nava et al. (1996) and Boadway et al. (1994).

The optimal balance of direct and indirect taxation maybe is the most practical question in real world. The optimal taxation theory has provided such useful insights economically based on utilitarian paradigm. However, not all are acceptable politically. The notions of uniform indirect taxation and only having income taxation in the economy only make sense to economists but not for policy makers or average voters.

As expected, all type of economic models may need to include some conceptions that usually ignored in the model. The discussion on the best model is always debatable since every model will have its unique insights in its specific settings. As in Mankiw et al. (2009), how to take into account new ideas into the theory of optimal taxation remains an open question.

Using the model developed by Nava et al. (1996), this research tries to provide new insights to the existing model, i.e. a model that can provide a new guidance on how the government should define the optimal income and sales tax rates.

### 3.3 The Model

In this section, we will discuss the assumptions used in the model, the household's maximization problem, the mimicker's maximization problem, and the social planner's problem. We will discuss the main results and leave the details of algebra in the appendices.

#### 3.3.1 Assumptions

As in standard literature, we assume there are two types of households in the economy: low type and high type. The government is not able to observe the ability of the individuals (or households) but can observe the individual's income. However, the individuals are able to conceal their income and concealing is a costly behavior where the cost is linear. Government imposes an income tax rate ($t_i$) where $i$ refers to household $i$.

Both high type and low type pay tax and only possible to conceal income, not expenditure. It means there is no uncertainty in sales tax. There are two types of consumption goods (private goods): good $a$ ($x_a$) and good $b$ ($x_b$) where government only imposes sales tax on good $a$ ($\tau_a$)and leave good $b$ untaxed ($\tau_b = 0$) therefore good $b$ is treated as a numéraire. In this setting, we assume that the income taxes are non-linear and commodity taxes are linear.

The final objective of the model—as well as the main objective of this research—is to solve a social planner's problem. However, to achieve it, we need to solve for the household's maximization problem and the mimicker's maximization problem. The following sections will discuss those three problems.

### 3.3.2 The household's maximization problem

We define the subscript or superscript $i = \{1, 2\}$ as the type of the household where $i = 1$ and $i = 2$ refer to the low type and the high type, respectively. $\pi_u$ is the probability of being undetected when the household is hiding some of its income and $\pi_d$ is the corresponding probability of being detected, where $\pi_d^i = 1 - \pi_u^i$. We assume that $U_u^i > U_d^i$, where the cost of being detected is a function $G(1-\sigma_i)w_i l_i(1-\sigma_i)$ and equals to $(U_u^i - U_d^i)$. $q_a$ and $q_b$ are the after-tax prices for good $a$ and $b$ respectively, $\sigma_i$ is the the proportion of reported income, and $T_i$ is a lump-sum component. The utility, $U(.)$, is a monotone and strictly concave function and it is a function of consumption $(x_a^i, x_b^i)$ and labor supply $(l_i)$.

Here we maximizes the expected utility (see equation (3.15)) subjects to a budget constraint.

$$\max_{x_a^i, x_b^i, l_i} \pi_u^i [U_u^i(x_a^i, x_b^i, l_i)] + \pi_d^i [U_d^i(x_a^i, x_b^i, l_i)] \tag{3.15}$$

subject to

$$q_a x_a^i + q_b x_b^i = w_i l_i (1 - t_i)\sigma_i + w_i l_i (1 - \sigma_i) - T_i$$

Since $q_a$ is equal to $p_a(1 + \tau_a)$ then we can rewrite the constraint as:

$$p_a(1 + \tau_a) + q_b x_b^i = w_i l_i (1 - \sigma_i t_i) - T_i$$

Define $g(1-\sigma_i) \equiv G(1-\sigma_i)(1-\sigma_i)$, we have the maximization problem as follows:

$$\max_{x_a^i, x_b^i, l_i} U_u^i(.) - \pi_d^i g(1-\sigma_i) w_i l_i \text{ s.t. } p_a(1+\tau_a)x_a^i + q_b x_b^i = w_i l_i (1 - t_i \sigma_i) - T_i \tag{3.16}$$

By Roy's identity we obtain:

$$V_{\tau_a}^i = \frac{dU}{d\tau_a} = \left.\frac{\partial \mathcal{L}}{\partial \tau_a}\right|_* = -\alpha_i x_a^i \tag{3.17}$$

$$V_{t_i}^i = \frac{dU}{dt_i} = \left.\frac{\partial \mathcal{L}}{\partial t_i}\right|_* = -\alpha_i w_i l_i \sigma_i \tag{3.18}$$

$$V_{T_i}^i = \frac{dU}{dT_i} = \left.\frac{\partial \mathcal{L}}{\partial T_i}\right|_* = -\alpha_i \tag{3.19}$$

$$V_{\sigma_i}^i = \frac{dU}{d\sigma_i} = \left.\frac{\partial \mathcal{L}}{\partial \sigma_i}\right|_* = -w_i l_i (\pi_d^i g'(1-\sigma_i) + \alpha_i t_i) \tag{3.20}$$

where $\alpha_i$ is the marginal utility of income.

### 3.3.3 The mimicker's maximization problem

Now, let us investigate the maximization problem faced by a mimicker. A mimicker—denoted by a horizontal bar above the variables—is a high-type individual that imitate the behavior of a low-type individual.

Define that $\bar{l_2} \equiv \frac{w_1 l_1}{w_2}$, $\bar{\sigma_2} \equiv \frac{\sigma_i w_i l_i}{w_2 \bar{l_2}}$, $\bar{x}_a^2(\tau_a, t_1, T_1; \bar{l_2}, \bar{\sigma_2})$, $\bar{x}_b^2(\tau_a, t_1, T_1; \bar{l_2}, \bar{\sigma_2})$, $\bar{V}^2(\tau_a, t_1, T_1) \equiv U^2(\bar{x}_a^2(.), \bar{x}_b^2(.), \frac{w_1 l_1(.)}{w_2}; \bar{\sigma}_2)$, then the maximization problem for a mimicker is as follows:

$$\max_{\bar{x}_a^2, \bar{x}_b^2, \bar{l}_2} U(\bar{x}_a^2, \bar{x}_b^2, \bar{l}_2) - \pi_d^2 g(1-\bar{\sigma}_2) w_2 \bar{l}_2 \text{ s.t. } w_2 \bar{l}_2 (1 - t_1 \bar{\sigma}_2) - T_1 - \bar{x}_a^2(1+\tau_a) - \bar{x}_b^2 \tag{3.21}$$

resulting:

$$\bar{V}^2_{\tau_a} = -\bar{\alpha}_2 \bar{x}^2_a + \bar{\alpha}^2 \frac{\partial w_1 l_1}{\partial \tau_a} \xi \tag{3.22}$$

$$\bar{V}^2_{t_1} = -\bar{\alpha}_2 w_1 l_1 \bar{\sigma}_2 + \bar{\alpha}_2 \frac{\partial w_1 l_1}{\partial t_1} \xi \tag{3.23}$$

$$\bar{V}^2_{T_1} = -\bar{\alpha}_2 + \bar{\alpha}^2 \frac{\partial w_1 l_1}{\partial T_1} \xi \tag{3.24}$$

$$\bar{V}^2_{\bar{\tau}_2} = -w_1 l_1 (\bar{\alpha}_2 t_1 + \pi^2_d g'(1 - \bar{\tau}_2)) \tag{3.25}$$

where[8]

$$\xi \equiv \frac{U_{\bar{l}_2}}{w_2 \bar{\alpha}_2} + (1 - t_1 \bar{\sigma}_2) - \frac{\pi^2_d g(1 - \bar{\sigma}_2)}{\bar{\alpha}_2}$$

### 3.3.4 The social planner's problem

Now we turn to the social planner's problem. The planner chooses a set of $\{\tau_a, t_i, T_i\}$ that will maximizes the utility of the high-type individual $V^2(.)$ (see equation (3.26)) whilst in the same time it guarantees that the low type will have at least a certain level of utility, $U^1_s$, as it is shown in the first constraint. The second constraint is an incompatibility constraint or a self-selection constraint that make sure that the high-type individual does not masquerade as a low-type individual. Since the government guarantee that the low type will reach a certain level of utility, the low type will not try to mimic the high type.[9] The third constraint is the government's budget constraint where $n_1$ and $n_2$ are the number of low-type and high-type individuals, respectively.

[8]If the individual reports all of her income, $\bar{\sigma}_2 = 1$, we have $\xi$ as in Nava et al. (1996).

[9]This is known as the normal case in Stiglitz (1982).

The left-hand side of the inequality is the total government revenue and,—as it is shown in the third constraint—, it must be higher than or at least equal to some level of revenues, $R_0$. The social planner's problem is then constructed as follows:

$$\max_{\tau_a, t_i, T_i} V^2(\tau_a, t_2, T_2) \tag{3.26}$$

subject to:

$$MUC: \quad V^1(\tau_a, t_1, T_1) \geq U_s^1$$

$$ICH: \quad V^2(\tau_a, t_2, T_2) \geq \bar{V}^2(\tau_a, t_1, T_1; \frac{w_1 l_1(.)}{w_2})$$

$$GBC: \quad n_1[\tau_a x_a^1(.) + \sigma_1 t_1 w_1 l_1(.) + T_1] + n_2[\tau_a x_a^2(.) + \sigma_2 t_2 w_2 l_2(.) + T_2] \geq R_0,$$

where MUC is the the minimum utility constraint for the low type, ICH is the incentive compatibility constraint for the high-type, and GBC is the government budget constraint. By utilizing the results from previous section (see the appendices for details), we are able to derive an optimal income and sales taxes as follows:[10]

$$t_1 = \frac{\partial \tilde{x}_a^1 / \partial t_1}{-\partial w_1 \tilde{l}_1 / \partial t_1} \frac{\tau_a}{\sigma_1} + \frac{\gamma \bar{\alpha}_2}{\lambda n_1 \sigma_1} \xi \tag{3.27}$$

$$t_2 = \frac{\partial \tilde{x}_a^2 / \partial t_2}{-\partial w_2 \tilde{l}_2 / \partial t_2} . \frac{\tau_a}{\sigma_2} \tag{3.28}$$

$$\tau_a = \frac{\gamma}{\lambda} \bar{\alpha}_2 (x_a^1 - \bar{x}_a^2) \left[ n_1 \frac{\partial \tilde{x}_a^1}{\partial \tau_a}\bigg|_{l_1} + n_2 \frac{\partial \tilde{x}_a^2}{\partial \tau_a}\bigg|_{l_2} \right]^{-1}, \tag{3.29}$$

[10] $\mu$ and $\lambda$ are positive due to the monotonicity of preferences, and $\gamma$ is strictly positive when $U_s^1$ is sufficiently high as shown in Nava et al. (1996). We also assume that the result is not a corner solution as private good can increase to infinity and leisure is capped at 24 hours per day.

where $\tilde{x}_a^i$ and $\tilde{l}_i$ denotes the *compensated demand* for $x_a$ and *compensated labor supply*, respectively.

We begin the analysis from the tax rate for the high-type, $t_2$. Without losing the generality, we assume that $\tau_a > 0$ regardless its optimality. Intuitively, we have the denominator $-\partial w_2 \tilde{l}_2 / \partial t_2 > 0$: the higher the income tax rate, the lower the labor supply and consequently the income $w_2 \tilde{l}_2$ (note that we have a negative sign in the denominator). We also have the ratio $\frac{\tau_a}{\sigma_2}$ is positive. Therefore, the sign of $t_2$ will depend on the sign of $\partial \tilde{x}_a^2 / \partial t_2$. Moreover, $t_2$ is non zero even at the top of income distribution since there is inefficiency generated by the sales taxes. A distortionary income tax is still needed to compensate the inefficiency (Nava et al., 1996).[11]

The sign of $\partial \tilde{x}_a^2 / \partial t_2$ depends on the the propinquity between good $a$ and leisure. If good $a$ and leisure are substitutes in the Hicks perception, then $\partial \tilde{x}_a^2 / \partial t_2 < 0$ and $t_2 < 0$. It implies that it will be optimal if we subsidize the high-type individuals. In contrast, if good $a$ and leisure are complementary, then $\partial \tilde{x}_a^2 / \partial t_2 > 0$ and consequently $t_2 > 0$.

Equation (3.28) also demonstrates an intuitive result that the income tax rate is inversely related to the proportion of reported income, $\sigma_2$. The higher the degree of tax evasion (i.e. the lower reported income), the higher the income tax rate. Specifically, this result aligns with Nava et al. (1996) as it is assumed implicitly that $\sigma_2 = 1$. Assuming that good $a$ and leisure are complementary and $\tau_a > 0$, it is easy to see that $t_2 > \tau_a$ if the elasticity of demand for good $a$ with respect to the high-type income tax is higher than the elasticity of labor supply with respect to the high-type income tax, and vice versa.

[11] Nava et al. (1996) points out this in the context of comparing his results to that of Stiglitz (1982) where no marginal income tax on high-type individuals is needed when both direct and indirect taxes are non-linear.

These results then come to four proportions:

*Proposition 1. In a Pareto optimum, the income tax of the high-type individual will be negative ($t_2 < 0$) if good a is a substitute for leisure.*

*Proposition 2. In a Pareto optimum, the income tax of the high-type individual will be positive ($t_2 > 0$) if good a is a complement for leisure.*

*Proposition 3. In a Pareto optimum where tax evasion exists, the income tax rate for the high-type individual will be higher as the degree of tax evasion increases.*

*Proposition 4. In a Pareto optimum where good a and leisure are complementary, $t_2 > \tau_a$ if the elasticity of demand for good a w.r.t. the high-type income tax is larger than the elasticity of labor supply w.r.t. the high-type income tax.*

Now we turn to the result in equation (3.27). To make it comparable directly with the corresponding result in Nava et al. (1996), we will modify equation (3.27). Defining $\hat{\xi} \equiv \frac{U_{\bar{l}_2}}{w_2 \bar{\alpha}_2} + (1 - t_1 \bar{\sigma}_2)$, then:

$$t_1 = \frac{\partial \tilde{x}_a^1 / \partial t_1}{-\partial w_1 \tilde{l}_1 / \partial t_1} \frac{\tau_a}{\sigma_1} + \frac{\gamma \bar{\alpha}_2}{\lambda n_1 \sigma_1} \left[ \hat{\xi} - \frac{\pi_d^2 g(1 - \bar{\sigma}_2)}{\bar{\alpha}_2} \right] \tag{3.30}$$

One can observe that $\hat{\xi}$ is equal to $\xi$ defined in Nava et al. (1996) when $\sigma_2 = 1$, and consequently equation (3.30) aligns with the result for $t_1$ in Nava et al. (1996) when $\sigma_i = 1$ and the probability of being detected for high-type, $\pi_d^2$, is zero.[12] Both results are also identical when $\sigma_1 = 1$ and the cost of hiding the income, $g(.)$ is zero.[13] In essence, the case in Nava et al. (1996) is a special case of our analysis. It is interesting to see that in this setting the income tax rate for the low-type individual is also affected by the probability of being detected for the high type.

Assuming that the first term in (3.30) is positive as for $t_2$, then it is intuitive that $t_1$ will decrease as the probability of being detected for the high type increase, *ceteris paribus.*

[12]In this case $\hat{\xi} > 0$ as it has been proved in Nava et al. (1996).
[13]It is naturally assumed that $g(0) = 0$.

A higher $\pi_d^2$ implies higher government revenues and accordingly lower the income tax rate of the low type in this case.

The analysis of the effects of the first term in the right-hand side on $t_1$ is analog to the previous discussion on $t_2$. The second term provide more information about the optimal income tax rate for the low type where the tax rate will be positively affected by $\{\gamma, \bar{\alpha}_2\}$ and will be negatively affected by $\{\lambda, n_1, \sigma_1, \pi_d^2, g(.)\}$. These results then provide two propositions:

*Proposition 5. In a Pareto optimum, the income tax of the low-type individual will decrease as the probability of being detected for the high type increase, ceteris paribus .*

*Proposition 6. In a Pareto optimum, the optimal income tax rate for the low type will be positively affected by $\{\gamma, \bar{\alpha}_2\}$ and will be negatively affected by $\{\lambda, n_1, \sigma_1, \pi_d^2, g(.)\}$*

### 3.4 Conclusions and policy implications

In this research we have introduced uncertainty and tax evasion into an optimal tax mix problem in two-class economy with two consumption goods and labor. We assume that the government cannot observe the agent's ability and only able to observe unconcealed income, where concealing is a costly behavior for the agent. We also assume that the income taxes are non-linear and the commodity taxes are linear. The model has beed constructed based on the approach used by Boadway et al. (1994), Nava et al. (1996) and del Mar Racionero (2001).

After taking into account uncertainty and tax evasion, as in Nava et al. (1996) we also find that in the Pareto optimum the sign of income tax for the high-type agent depends on the relation between the (taxed) private good and leisure. If the private good is a substitute for leisure, then the income tax for the high-type will be negative. However, if the private good is a complement for leisure, the income tax of the high type will be positive.

As far as the tax evasion is concerned, the income tax for the high-type individual will be higher as the degree of tax evasion increases. Intuitively, this result is the consequences of the government's budget constraint: to maintain the balanced budget, income tax should be high when tax evasion is high. We also find that if the taxed private good and leisure are complementary—i.e. when the income tax is positive for the high type—and the elasticity of demand with respect to income tax is larger than the elasticity of labor supply with respect to income tax, the commodity tax rate is higher than the income tax rate.

For the case of low-type, the model shows that the income tax of the low-type individual decreases as the probability of being detected for the high-type increases, *ceteris paribus.* The result shows that the optimal income tax rate for the low-type individual increases as the marginal utility of mimicking the low type or the marginal utility of income for the mimicker increases, *ceteris paribus.* In other words, income tax for the low-type will increase if the high type has more incentive to mimic the low-type. Conversely, the optimal income tax decreases as the marginal utility of tax revenue, the number of low-type agents, proportion of reported income, probability that the high-type will be detected when hiding the income, or the cost of hiding the income increases.

In spite of the fact that the model is abstract, it has policy relevance. First, as shown above, income tax is not necessary higher for the high type (high income), or even negative to increase the total welfare. Although it might not sound politically correct, our result shows that it will depend on the agent's behavior toward leisure. Second, in order to reduce the income tax rate for the low type, the policy maker needs to build a reliable tax system that can reduce the possibility of tax evasion and consequently will increase the tax collection rate.

As policy makers and most economists believe that the income tax can be utilized as a tool to increase equality, there is a trade-off between equality and efficiency. A higher income tax rate for the high types may reduce inequality and in the same time it will also reduce efficiency as there will be less incentive for the high type to work more. One can argue that it depends on the labor supply elasticities of low and high types as their after-tax wages changes. As in Hausman (1983) and Nava et al. (1996), our results show that, by taking into account the level of tax evasion, the higher elasticity will result in a lower tax rate to maintain an optimum welfare level. Regarding equity, if the high type more responsive to the income tax rate, the income tax rate should be less progressive and equity will be more difficult to achieve. An optimum policy to deal with the trade-off remains an open question that will be an interesting for a further empirical research topic. Institutional factors, such as tax and political systems, would be important to be considered in theoretical and empirical research.

Taking into account tax evasion in the model is crucial since the case will be considerably different for developed and developing countries. Developed countries, in general, have better tax systems than developing countries do; and consequently the degrees of tax evasions will be different. Our model captures this notion. As the optimal income and sales tax rates will be higher for a higher degree of tax evasion, we can expect that developing countries will need to impose higher tax rates to meet some level of tax revenue, ceteris paribus. One way to reduce the tax rates—that will be beneficial for poor households via a lower income tax rate—is by implementing a more reliable tax system to reduce tax evasion. Further research on different impacts of tax evasions in developed and developing countries is warranted.

CHAPTER 4

# Concluding Remarks

This dissertation examines two topics in economics. First, the impacts of spillovers of public goods on the potential benefits from decentralized policy in urban economy. Do spillovers of local public goods across jurisdictions have impacts on the potential gain of decentralization in an urban economy? Are the impacts different for urban economies in developed and developing countries? Those two questions are the main research questions that we investigate in the first topic. Second, the role of tax evasion and uncertainty on the optimal taxation. Specifically, how tax evasion and uncertainty will affect the optimal taxation for both income and sales taxes in a two-class economy.

In investigating the first topic in Chapter 2, it is impossible not to mention Oates' fiscal decentralization theorem (Oates, 1972) that shed light on the formal economic intuition on centralization versus decentralization. Oates (1972) points out that decentralization will be optimal if there is no spillover. Specifically, decentralized local public good provisions based on local preferences will generate a higher social welfare than centralized local public good provision with uniform levels of public good for all jurisdictions. Enhancing the simulation model used by Calabrese et al. (2012), this dissertation illustrate the direction of the impact of spillover on the potential benefits from decentralization. The findings in Chapter 2 support the idea of Oates decentralization theorem whereas the degree of the spillover is sufficiently high, the potential benefit from having decentralization decreases. The results are consistent in both urban economies in developed and developing countries.

From policy perspective there are two potential questions in two different situations that might arise. First question: do we need to decentralize the tax policy? The answer will then depend on the levels of spillovers across a jurisdictions. An appropriate method to measure the spillover level will be another challenging task that is needed to be carefully approached. One may expect that the task will not be simple since it will involve agents' behaviors and preferences that are always difficult, if not impossible, to reveal. Second question: assuming the decentralized tax policy has been implemented, how can governments formulate a certain policy that can limit spillovers? In practice—regardless its advantages and disadvantages—the implementation of school neighborhood zoning policy or other zoning regulations is one possible answer. However, it might not be easy to find answers for another type of local public local goods, such as recreation areas or other kinds of city public amenities due to their characteristics as public goods which are non-rivalrous and non-excludable at certain levels.

The second topic is discussed in Chapter 3 where the model is built based on Nava et al. (1996) and Boadway et al. (1994). As in Nava et al. (1996), the results show that the sign of the income tax for the high type will depend on the behavior of the high type toward leisure. If the (taxed) consumption good is a complement for leisure then the income tax rate will be positive. In contrary, if the consumption good is a substitute for leisure then the optimal income tax will be negative. By taking into account tax evasion and uncertainty, the findings shows us that higher tax evasion will increase the income tax rates. In addition, the income tax rates for the low type will increase if the high type has more incentive to mimic the low-type and the probability of getting caught (when hiding some portion of income) is low.

From policy perspective, when the high type is mimicking the low type, the effectiveness of income tax as a redistribution tool will be lower. Lower tax evasion and higher tax compliance will lower the income tax rate for the low-type and consequently increase the equity. The importance of a reliable tax system and simple tax codes are imperative to reduce the potency of tax evasion. This is relevant for economies globally where degree of tax evasion may vary as tax systems and institutions in general are different.

## BIBLIOGRAPHY


Akai, N. and Sakata, M. (2002). Fiscal Decentralization Contributes to Economic Growth: Evidence from state-level cross-section data for the United States. *Journal of Urban Economics*, 52(1):93–108.

Alesina, A. and Spolaore, E. (1997). On the Number and Size of Nations. *The Quarterly Journal of Economics*, 112(4):1027–56.

Alm, J. (1996). What is an Optimal Tax System. *National Tax Journal*, 49(1):117–133.

Arnott, R. J. and Stiglitz, J. E. (1979). Aggregate Land Rents, Expenditure on Public Goods, and Optimal City Size. *The Quarterly Journal of Economics*, 93(4):471–500.

Atkinson, A. B. (1977). Optimal Taxation and the Direct versus Indirect Tax Controversy. *Canadian Journal of Economics*, pages 590–606.

Atkinson, A. B. and Stiglitz, J. E. (1972). The Structure of Indirect Taxation and Economic Efficiency. *Journal of Public Economics*, 1(1):97–119.

Atkinson, A. B. and Stiglitz, J. E. (1976). The Design of Tax Structure: Direct versus Indirect Taxation. *Journal of public economics*, 6(1-2):55–75.

Atkinson, A. B., Stiglitz, J. E., and Atkinson, A. (1980). *Lectures on Public Economics*. McGraw-Hill New York.

Ballard, C. L., Shoven, J. B., and Whalley, J. (1985). General Equilibrium Computations of the Marginal Welfare Costs of Taxes in the United States. *The American Economic Review*, 75(1):128–138.

Barr, J. L. and Davis, O. A. (1966). An Elementary Political and Economic Theory of the Expenditures of Local Governments. *Southern Economic Journal*, 33(2):149–165.

Barro, R. J. (1990). Government Spending in a Simple Model of Endogeneous Growth. *Journal of Political Economy*, 98(5):103–125.

Barro, R. J. and Sala-i Martin, X. (1992). Public Finance in Models of Economic Growth. *The Review of Economic Studies*, 59(4):645–661.

Benabou, R. (1991). Workings of a City: Location, Education, and Production. NBER Technical Working Papers 0113, National Bureau of Economic Research, Inc.

Benabou, R. (1993). Workings of a City: Location, Education, and Production. *The Quarterly Journal of Economics*, 108(3):619–52.

Benabou, R. (1996). Equity and Efficiency in Human Capital Investment: The Local Connection. *Review of Economic Studies*, 63(2):237–64.

Bergstrom, T. C. (1979). When Does Majority Rule Supply Public Goods Efficiently? *Scandinavian Journal of Economics*, 81(2):216–26.

Besley, T. and Coate, S. (2003). Centralized versus Decentralized Provision of Local Public Goods: A political economy approach. *Journal of Public Economics*, 87(12):2611 – 2637.

Bewley, T. (1981). A Critique of Tiebout's Theory of Local Public Expenditures. *Econometrica*, 49:713–740.

Blackley, D. M. (1999). The Long-Run Elasticity of New Housing Supply in the United States: Empirical evidence for 1950 to 1994. *The Journal of Real Estate Finance and Economics*, 18(1):25–42.

Boadway, R. (1982). On the Method of Taxation and the Provision of Local Public Goods: Comment. *American Economic Review*, 72(4):846–51.

Boadway, R., Marchand, M., and Pestieau, P. (1994). Towards a Theory of the Direct-indirect Tax Mix. *Journal of Public Economics*, 55(1):71–88.

Bohn, H. (1990). Tax Smoothing with Financial Instruments. *The American Economic Review*, pages 1217–1230.

BPS (2010). Statistics Indonesia (BPS). accessed March 6, 2013.

BPS (2012). *Trends of Selected Socio-economic Indicators of Indonesia*. BPS-Statistics Indonesia.

Bradford, D. F. and Rosen, H. S. (1976). The Optimal Taxation of Commodities and Income. *The American Economic Review*, 66(2):94–101.

Browning, E. K. (1987). On the Marginal Welfare Cost of Taxation. *The American Economic Review*, pages 11–23.

Brueckner, J. (1986). The Structure of Urban Equilibria: A unified treatment of the Muth-Mills model. *Handbook of regional and urban economics: Urban economics*, 2:821.

Brueckner, J. K. (1983). Property Value Maximization and Public Sector Efficiency. *Journal of Urban Economics*, 14(1):1–15.

Brueckner, J. K. (1999). Fiscal Federalism and Capital Accumulation. *Journal of Public Economic Theory*, 1(2):205–24.

Brueckner, J. K. (2004). Fiscal Decentralization with Distortionary Taxation: Tiebout vs. Tax Competition. *International Tax and Public Finance*, 11(2):133–153.

Brueckner, J. K. (2006). Fiscal Federalism and Economic Growth. *Journal of Public Economics*, 90(1011):2107 – 2120.

Brueckner, J. K., Mills, E., and Kremer, M. (2001). Urban Sprawl: Lessons from Urban Economics [with Comments]. *Brookings-Wharton papers on urban affairs*, pages 65–97.

Buchanan, J. M. (1965). An Economic Theory of Clubs. *Economica*, pages 1–14.

Calabrese, S. and Epple, D. (2010). On the political economy of tax limits.

Calabrese, S., Epple, D., and Romano, R. (2007). On the Political Economy of Zoning. *Journal of Public Economics*, 91(1-2):25–49.

Calabrese, S., Epple, D., and Romano, R. (2012). Inefficiencies from Metropolitan Political and Fiscal Decentralization: Failures of Tiebout competition. *Review of Economic Studies*, 79:1081–1111.

Calabrese, S., Epple, D., Romer, T., and Sieg, H. (2006). Local Public Good Provision: Voting, peer effects, and mobility. *Journal of Public Economics*, 90(6):959–981.

Chotikapanich, Valenzuela, R. (1997). Global and Regional Inequality in the Distribution of Income: Estimation with limited and incomplete data. *Empirical Economics*, 22:533–546.

Cook, P. J. (1972). A 'One Line' Proof of the Slutsky Equation. *American Economic Review*, 62(1).

Cremer, H. and Gahvari, F. (1995). Uncertainty, Optimal Taxation and the Direct versus Indirect Tax Controversy. *The Economic Journal*, pages 1165–1179.

Davidoff, T. (2005). Income Sorting: Measurement and decomposition. *Journal of Urban Economics*, 58(2):289–303.

Davoodi, H. and Zou, H.-F. (1998). Fiscal Decentralization and Economic Growth: A cross-country study. *Journal of Urban Economics*, 43(2):244–257.

de Bartolome, C. A. M. (1990). Equilibrium and Inefficiency in a Community Model with Peer Group Effects. *Journal of Political Economy*, 98(1):110–33.

del Mar Racionero, M. (2001). Optimal Tax Mix with Merit Goods. *Oxford Economic Papers*, 53(4):628–641.

Diamond, P. A. and Mirrlees, J. A. (1971). Optimal Taxation and Public Production I: Production efficiency. *The American Economic Review*, 61(1):8–27.

DiPasquale, D. and Wheaton, W. C. (1992). The Cost of Capital, Tax Reform, and the Future of the Rental Housing Market. *Journal of Urban Economics*, 31(3):337–359.

Dowding, K., John, P., and Biggs, S. (1994). Tiebout: A survey of the empirical literature. *Urban studies*, 31(4-5):767–797.

Edwards, J., Keen, M., and Tuomala, M. (1994). Income Tax, Commodity Taxes and Public Good Provision: A brief guide. *FinanzArchiv/Public Finance Analysis*, pages 472–487.

Epple, D., Filimon, R., and Romer, T. (1984). Equilibrium among Local Jurisdictions: Toward an integrated treatment of voting and residential choice. *Journal of Public Economics*, 24(3):281–308.

Epple, D., Gordon, B., and Sieg, H. (2010). A New Approach to Estimating the Production Function for Housing. *American Economic Review*, 100(3):905–24.

Epple, D. and Nechyba, T. (2004). Fiscal Decentralization. 4:2423–2480.

Epple, D. and Platt, G. J. (1998). Equilibrium and Local Redistribution in an Urban Economy when Households Differ in Both Preferences and Incomes. *Journal of Urban Economics*, 43(1):23–51.

Epple, D. and Romer, T. (1991). Mobility and Redistribution. *Journal of Political Economy*, 99(4):828–58.

Epple, D., Romer, T., and Sieg, H. (2001). Interjurisdictional Sorting and Majority Rule: An empirical analysis. *Econometrica*, 69(6):pp. 1437–1465.

Epple, D. and Sieg, H. (1999). Estimating Equilibrium Models of Local Jurisdictions. *Journal of Political Economy*, 107(4):645–681.

Feldstein, M. (1999). Tax Avoidance and the Deadweight Loss of the Income Tax. *Review of Economics and Statistics*, 81(4):674–680.

Fernandez, R. and Rogerson, R. (1996). Income Distribution, Communities, and the Quality of Public Education. *The Quarterly Journal of Economics*, 111(1):135–64.

Fisman, R. and Gatti, R. (2002). Decentralization and Corruption: Evidence across countries. *Journal of Public Economics*, 83(3):325–345.

Fullerton, D., Shoven, J. B., and Whalley, J. (1983). Replacing the US Income Tax with a Progressive Consumption Tax: A sequenced general equilibrium approach. *Journal of Public Economics*, 20(1):3–23.

Glaeser, E. (2008). *Cities, Agglomeration, and Spatial Equilibrium*. The Lindahl lectures. Oxford University Press.

Goodspeed, T. J. (1995). Local Income Taxation: An externality, Pigouvian solution, and public policies. *Regional Science and Urban Economics*, 25(3):279–296.

Gordon, R. H. (1983). An Optimal Taxation Approach to Fiscal Federalism. *The Quarterly Journal of Economics*, 98(4):567–86.

Groves, H. M. and Kahn, C. H. (1952). The Stability of State and Local Tax Yields. *The American Economic Review*, 42(1):87–102.

Gyourko, J. and Tracy, J. (1991). The Structure of Local Public Finance and the Quality of Life. *Journal of Political Economy*, pages 774–806.

Hanushek, E. A. and Quigley, J. M. (1980). What Is the Price Elasticity of Housing Demand? *The Review of Economics and Statistics*, 62(3):449–54.

Hardman, A. and Ioannides, Y. M. (2004). Neighbors' Income Distribution: Economic segregation and mixing in US urban neighborhoods. *Journal of Housing Economics*, 13(4):368–382.

Haurin, D. R. (1981). Local Income Taxation in an Urban Area. *Journal of Urban Economics*, 10(3):323–337.

Hausman, J. A. (1983). Taxes and Labor Supply.

Heien, D. M. (1968). A Note on Log-linear Regression. *Journal of the American Statistical Association*, pages 1034–1038.

Henderson, J. V. (1974). The Sizes and Types of Cities. *The American Economic Review*, 64(4):640–656.

Hindriks, J. and Myles, G. D. (2006). *Intermediate Public Economics*. MIT press Cambridge.

Hochman, O., Pines, D., and Thisse, J.-F. (1995). On the Optimal Structure of Local Governments. *The American Economic Review*, pages 1224–1240.

Hooghe, L. and Marks, G. (2007). Union: Virtues and vices of scale flexibility in jurisdictional design. In *West European Politics 30th Anniversary Conference, European University Institute*.

Hoyt, W. H. and Rosenthal, S. S. (1989). Capital Gains Taxation and the Demand for Owner-occupied Housing. Working Paper Series / Economic Activity Section 92, Board of Governors of the Federal Reserve System (U.S.).

Hoyt, W. H. and Rosenthal, S. S. (1990). Capital Gains Taxation and the Demand for Owner-occupied Housing. *The Review of Economics and Statistics*, 72(1):45–54.

Iimi, A. (2005). Decentralization and Economic Growth Revisited: An empirical note. *Journal of Urban Economics*, 57(3):449–461.

Ivanov, I. and Dobreva, J. (2010). Modeling Labor Supply through Duality and the Slutsky Equation. *International Journal of Economic Sciences and Applied Research*, (2):111–122.

Jackman, R. (1973). Fiscal Federalism: W.E. Oates, (Harcourt, Brace, Jovanovich, Inc. 1972 xvi + 256 pp.). *Journal of Public Economics*, 2(2):188–191.

Jacobs, B. and Boadway, R. (2013). Optimal Linear Commodity Taxation under Optimal Non-linear Income Taxation. Technical report, CESifo Working Paper: Public Finance.

Kanbur, R., Pirttilä, J., and Tuomala, M. (2006). Non-welfarist Optimal Taxation and Behavioural Public Economics. *Journal of Economic Surveys*, 20(5):849–868.

Kee, J. E. (2004). Fiscal Decentralization: Theory as reform. *Financial Management Theory in the Public Sector*, pages 165–185.

Leach, J. (2004). *A Course in Public Economics*. CUP.

Mankiw, N. G., Weinzierl, M., and Yagan, D. (2009). Optimal Taxation in Theory and Practice. Technical report, National Bureau of Economic Research.

Mas-Collel, A., Whinston, M. D., and Green, J. (1995). *Microeconomic Theory*. Oxford university press Oxford.

Mauro, P. (1995). Corruption and Growth. *The Quarterly Journal of Economics*, 110(3):681–712.

Mayer, C. J. and Somerville, C. T. (2000). Residential Construction: Using the urban growth model to estimate housing supply. *Journal of Urban Economics*, 48(1):85–109.

Menezes, C. F. and Wang, X. H. (2005). Duality and the Slutsky Income and Substitution Effects of Increases in Wage Rate Uncertainty. *Oxford Economic Papers*, 57(3):545–557.

Mills, E. S. (1967). An Aggregative Model of Resource Allocation in a Metropolitan Area. *The American Economic Review*, 57(2):197–210.

Mirrlees, J. A. (1971). An Exploration in the Theory of Optimum Income Taxation. *The review of economic studies*, 38(2):175–208.

Mirrlees, J. A. (1976). Optimal Tax Theory: A synthesis. *Journal of Public Economics*, 6(4):327–358.

Monkkonen, P. (2013). Urban Land-use Regulations and Housing Markets in Developing Countries: Evidence from Indonesia on the Importance of Enforcement. *Land Use Policy*, 34:255–264.

Murray, S. E., Evans, W. N., and Schwab, R. M. (1998). Education-Finance Reform and the Distribution of Education Resources. *American Economic Review*, 88(4):789–812.

Musgrave, R. A. (1939). The Voluntary Exchange Theory of Public Economy. *Quarterly Journal of Economics*, 53(2):213 – 237.

Musgrave, R. A. (1959). *The Theory of Public Finance: A study in public economy*. McGraw-Hill New York.

Naito, H. (1999). Re-examination of Uniform Commodity Taxes under a Non-linear Income Tax System and Its Implication for Production Efficiency. *Journal of Public Economics*, 71(2):165–188.

Nava, M., Schroyen, F., and Marchand, M. (1996). Optimal Fiscal and Public Expenditure Policy in a Two-class Economy. *Journal of Public Economics*, 61(1):119–137.

Neary, J. P. and Roberts, K. W. (1980). The Theory of Household Behaviour under Rationing. *European economic review*, 13(1):25–42.

Nechyba, T. J. (1999). School Finance Induced Migration and Stratification Patterns: The impact of private school vouchers. *Journal of Public Economic Theory*, 1(1):5–50.

Newbery, D. and Stern, N. (1987). *The Theory of Taxation for Developing Countries*. Oxford University Press.

Oates, W. E. (1972). *Fiscal Federalism*. Harcourt College Publishers.

Oates, W. E. (1998). *On the Welfare Gains from Fiscal Decentralization*. University of Maryland, Department of Economics.

Oates, W. E. (1999). An Essay on Fiscal Federalism. *Journal of Economic Literature*, 37(3):1120–1149.

Panzar, J. C. and Willig, R. D. (1976). Vindication of a" Common Mistake" in Welfare Economics. *The Journal of Political Economy*, 84(6):1361–1363.

Poterba, J. M. (1992). Taxation and Housing: Old questions, new answers. *American Economic Review*, 82(2):237–42.

Quigley, J. M. and Raphael, S. (2004). Is housing unaffordable? why isn't it more affordable? *The Journal of Economic Perspectives*, 18(1):191–214.

Ramsey, F. P. (1927). A Contribution to the Theory of Taxation. *The Economic Journal*, 37(145):47–61.

Roback, J. (1982). Wages, Rents, and the Quality of Life. *The Journal of Political Economy*, pages 1257–1278.

Rosen, H. S. (1979). Housing Decisions and the U.S. Income Tax : An econometric analysis. *Journal of Public Economics*, 11(1):1–23.

Rosenthal, S. S., Duca, J. V., and Gabriel, S. A. (1991). Credit Rationing and the Demand for Owner-occupied Housing. *Journal of Urban Economics*, 30(1):48–63.

Ross, S. and Yinger, J. (1999). Sorting and Voting: A review of the literature on urban public finance. 3:2001–2060.

Saez, E. (2001). Using Elasticities to Derive Optimal Income Tax Rates. *The Review of Economic Studies*, 68(1):205–229.

Samuelson, P. A. (1954). The Pure Theory of Public Expenditure. *The Review of Economics and Statistics*, 36(4):pp. 387–389.

Schmidheiny, K. (2006). Income Segregation from Local Income Taxation when Households Differ in Both Preferences and Incomes. *Regional Science and Urban Economics*, 36(2):270–299.

Seade, J. K. (1977). On the Shape of Optimal Tax Schedules. *Journal of public Economics*, 7(2):203–235.

Seegert, N. (2012a). Optimal Tax Portfolios: An estimation of government tax revenue minimum-variance frontiers. *Department of Economics, University of Michigan, Ann Arbor MI.*

Seegert, N. (2012b). Optimal Taxation with Volatility: A theoretical and empirical decomposition. *Job Market Paper, Department of Economics, University of Michigan, Ann Arbor MI.*

Seegert, N. (2012c). Welfare Consequences of Volatile Tax Revenue. *Department of Economics, University of Michigan, Ann Arbor MI.*

Sieg, H., Smith, V. K., Banzhaf, H. S., and Walsh, R. (2002). Interjurisdictional Housing Prices in Locational Equilibrium. *Journal of Urban Economics*, 52(1):131–153.

Silva, E. C. and Yamaguchi, C. (2010). Interregional Competition, Spillovers and Attachment in a Federation. *Journal of Urban Economics*, 67(2):219–225.

Slemrod, J. (1990). Optimal Taxation and Optimal Tax Systems. *The Journal of Economic Perspectives*, 4(1):157–178.

Stansel, D. (2005). Local Decentralization and Local Economic Growth: A cross-sectional examination of US metropolitan areas. *Journal of Urban Economics*, 57(1):55–72.

Stiglitz, J. E. (1982). Self-selection and Pareto Efficient Taxation. *Journal of Public Economics*, 17(2):213–240.

Stiglitz, J. E. (1988). Pareto Efficient and Optimal Taxation and the New New Welfare Economics.

Stuart, C. (1984). Welfare Costs per Dollar of Additional Tax Revenue in the United States. *The American Economic Review*, 74(3):352–362.

Tiebout, C. M. (1956). A Pure Theory of Local Expenditures. *Journal of Political Economy*, 64(5):pp. 416–424.

Turnovsky, S. J. (1996). Optimal Tax, Debt, and Expenditure Policies in a Growing Economy. *Journal of Public Economics*, 60(1):21–44.

Wellisch, D. (1993). On the Decentralized Provision of Public Goods with Spillovers in the Presence of Household Mobility. *Regional Science and Urban Economics*, 23(5):667–679.

Wellisch, D. (1994). Interregional Spillovers in the Presence of Perfect and Imperfect Household Mobility. *Journal of Public Economics*, 55(2):167–184.

Wildasin, D. E. (1980). Locational Efficiency in a Federal System. *Regional Science and Urban Economics*, 10(4):453–471.

Wildasin, D. E. (1985). Income Taxes and Urban Spatial Structure. *Journal of Urban Economics*, 18(3):313–333.

Wilson, J. D. (1997). Property Taxation, Congestion, and Local Public Goods. *Journal of Public Economics*, 64(2):207–217.

World Bank Group (2013). World Development Indicators (WDI). accessed March 6, 2013.

Xie, D., Zou, H.-F., and Davoodi, H. (1999). Fiscal Decentralization and Economic Growth in the United States. *Journal of Urban Economics*, 45(2):228–239.

Yakita, A. (2003). Taxation and Growth with Overlapping Generations. *Journal of Public Economics*, 87(3-4):467–487.

Zhang, T. and Zou, H.-F. (1998). Fiscal Decentralization, Public Spending, and Economic Growth in China. *Journal of Public Economics*, 67(2):221–240.

APPENDICES

## A.1 Derivations of first-order conditions

First-order Conditions:

$$\frac{\partial \mathcal{L}}{\partial R} = \sum_{i=1}^{J} \frac{\omega_R}{J} + \Omega = 0$$

$$\Rightarrow \omega_R + \Omega = 0$$

$$\Rightarrow -\Omega = \omega_R \tag{A-1}$$

$$\frac{\partial \mathcal{L}}{\partial r(y,\alpha)} = \sum_{i=1}^{J} \int_S \frac{\partial}{\partial r}[\omega V_i^e a_i f]\, dy\, d\alpha + \Omega \int_S \frac{\partial}{\partial r}[rf]\, dy\, d\alpha + \sum_{i=1}^{J} \eta_i \int_S \frac{\partial}{\partial r}[h_d^i a_i f] dy d\alpha = 0$$

$$\Rightarrow \sum_{i=1}^{J} \int_S \omega \frac{\partial V_i^e}{\partial r} a_i f\, dy\, d\alpha + \Omega \int_S f\, dy\, d\alpha + \sum_{i=1}^{J} \eta_i \int_S \frac{\partial h_d^i}{\partial r} a_i f dy d\alpha = 0$$

$$\Rightarrow \sum_{i=1}^{J} \omega U_1^i a_i \int_S f\, dy\, d\alpha + \Omega \int_S f\, dy\, d\alpha + \sum_{i=1}^{J} \eta_i (h_d^i)_2 a_i \int_S f dy d\alpha = 0$$

$$\Rightarrow \sum_{i=1}^{J} \omega U_1^i a_i + \Omega + \sum_{i=1}^{J} \eta_i (h_d^i)_2 a_i = 0$$

$$\Rightarrow -\Omega = \sum_{i=1}^{J} \omega U_1^i a_i + \sum_{i=1}^{J} \eta_i (h_d^i)_2 a_i \tag{A-2}$$

$$\frac{\partial \mathcal{L}}{\partial a_i(y,\alpha)} = \int_S \omega V_i^e f\, dy\, d\alpha + \eta_i \int_s h_d^i f\, dy\, d\alpha - \lambda_i (T_i - g_i) \int_S f\, dy\, d\alpha$$

$$\Rightarrow \omega V_i^e + \eta_i h_d^i - \lambda_i (T_i - g_i) = 0 \tag{A-3}$$

$$
\begin{aligned}
\frac{\partial \mathcal{L}}{\partial T_i} =& -\omega U_1^i \int_S a_i f \, dy \, d\alpha - \eta_i (h_d^i)_2 \int_S a_i f \, dy \, d\alpha + \lambda_i \int_S a_i f \, dy \, d\alpha = 0 \\
\Rightarrow& -\omega U_1^i - \eta_i (h_d^i)_2 + \lambda_i = 0 \\
\Rightarrow& \lambda_i = \omega U_1^i + \eta_i (h_d^i)_2 && \text{(A-4)} \\
\frac{\partial \mathcal{L}}{\partial t_i} =& \omega_R \left( -\frac{p_i H_s^i}{(1+t_i)^2} \right) + \left[ -\lambda_i \frac{H_s^i p_i t_i}{(1+t)^2} + \lambda_i \frac{H_s^i p_i}{1+t_i} - \lambda_i \frac{p_i^2 t_i H_s^{i'}}{(1+t_i)^3} \right] + \eta_i \frac{p_i H_s^{i'}}{(1+t_i)^2} = 0 \\
\Rightarrow& -\omega_R \frac{p_i}{1+t_i} + \left[ -\lambda_i \frac{p_i t_i}{1+t} + \lambda_i p_i - \lambda_i \frac{p_i^2 t_i}{(1+t_i)^2} \frac{H_s^{i'}}{H_s^i} \right] + \eta_i \frac{p_i}{(1+t_i)} \frac{H_s^{i'}}{H_s^i} = 0 \\
\Rightarrow& -\omega_R + \left[ -\lambda_i t_i + \lambda_i (1+t_i) - \lambda_i \frac{p_i t_i}{1+t_i} \frac{H_s^{i'}}{H_s^i} \right] + \eta_i \frac{(1+t_i)}{p_i} \frac{p_i}{1+t_i} \frac{H_s^{i'}}{H_s^i} = 0 \\
\Rightarrow& -\omega_R + \left[ \lambda_i - \lambda_i t_i \varepsilon_s^i \right] + \eta_i \frac{(1+t_i)}{p_i} \varepsilon_s^i = 0 \\
\Rightarrow& -\omega_R + \lambda_i (1 - t_i \varepsilon_s^i) + \eta_i \frac{(H_s^i)'}{H_s^i} = 0 && \text{(A-5)} \\
\frac{\partial \mathcal{L}}{\partial g_i} =& \int_S \omega U_3^i a_i f \, dy \, d\alpha + \eta_i \int_S (h_d^i)_3 a_i f \, dy \, d\alpha - \lambda_i \int_S a_i f \, dy \, d\alpha = 0 \\
\Rightarrow& \omega U_3^i N_i + \eta_i (h_d^i)_3 N_i - \lambda_i N_i = 0 && \text{(A-6)} \\
\frac{\partial \mathcal{L}}{\partial g_{-i}} =& \int_S \omega U_4^i \kappa a_i f \, dy \, d\alpha + \eta_i \int_S (h_d^i)_4 \delta a_i f \, dy \, d\alpha = 0 \\
\Rightarrow& \omega U_4^i \kappa N_i + \eta_i (h_d^i)_4 \delta N_i = 0 \\
\Rightarrow& \omega U_4^i \kappa + \eta_i (h_d^i)_4 \delta = 0 && \text{(A-7)}
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial \mathcal{L}}{\partial p_i} &= \int_S -\omega U_1^i h_d^i a_i f \, dy \, d\alpha + \omega_R \frac{H_s^i}{1+t_i} \\
&\quad + \lambda_i \left[ \frac{t_i}{1+t_i} H_s^i + \frac{t_i p_i}{1+t_i} \frac{H_s^{i'}}{1+t_i} \right] + \eta_i \int_S (h_d^i)_1 a_i f \, dy \, d\alpha - \eta_i \frac{H_s^{i'}}{1+t_i} = 0 \\
&\Rightarrow - \int_S \omega U_1^i h_d^i a_i f \, dy \, d\alpha \frac{1+t_i}{H_s^i} + \omega_R + \lambda_i \left[ t_i + \frac{t_i p_i}{1+t_i} \frac{H_s^{i'}}{H_s^i} \right] \\
&\quad + \eta_i \int_S (h_d^i)_1 a_i f \, dy \, d\alpha \frac{1+t_i}{H_s^i} - \eta_i \frac{H_s^{i'}}{1+t_i} \frac{1+t_i}{H_s^i} = 0 \\
&\Rightarrow - \int_S \omega U_1^i h_d^i a_i f \, dy \, d\alpha \frac{1+t_i}{H_s^i} + \omega_R + \lambda_i t_i (1+\varepsilon_s^i) \\
&\quad + \eta_i \int_S (h_d^i)_1 a_i f \, dy \, d\alpha \frac{1+t_i}{H_s^i} - \eta_i \frac{H_s^{i'}}{H_s^i} = 0 \\
&\Rightarrow \left[ \eta_i \int_S (h_d^i)_1 a_i f \, dy \, d\alpha - \int_S \omega U_1^i h_d^i a_i f \, dy \, d\alpha \right] \frac{1+t_i}{H_s^i} + \omega_R + \lambda_i t_i (1+\varepsilon_s^i) \\
&\quad - \eta_i \frac{H_s^{i'}}{H_s^i} = 0 \qquad \text{(A-8)}
\end{aligned}
$$

### A.2 Deriving equations for MATLAB functions[1]

The mean of income and the mean of preferences toward public goods are assumed having lognormal distribution. Following (Heien, 1968), the mean, $\mu$, of a variable X (in this case X is income or preference) whereas X has a lognormal distribution $(\mu_x, \sigma_x)$ is:

$$\mu = \exp(\mu_x + \frac{\sigma_x^2}{2}) \tag{A-9}$$

**Efficient Allocation**

Assuming a CES utility function:

$$U_i = [\beta_x x^\rho + \beta_h h^\rho + \beta_{g_i}(\alpha) g_i^\rho + \beta_{g_j}(\alpha)\kappa g_j^\rho]^{1/\rho}, \tag{A-10}$$

where $\rho < 0$ and $\beta'_{g_i}, \beta_{g_j} > 0$. Let $\beta_g(\alpha) \equiv \alpha$ and $\beta_{g_j} \equiv \alpha_j$. Note that we suppress subscript $i$ on $x$ and $h$ for simplicity. Assuming the household preferences on public goods are the same for $i$ and $j$ then we have:

$$U_i = [\beta_x x^\rho + \beta_h h^\rho + \alpha g_i^\rho + \alpha\kappa g_j^\rho]^{1/\rho} \tag{A-11}$$

*Deriving indirect utility function*

$$\max_h \, [\beta_x x^\rho + \beta_h h^\rho + \alpha g_i^\rho + \alpha\kappa g_j^\rho]^{1/\rho} \tag{A-12}$$

[1] The equations are derived based on Epple and Sieg (1999) and a numerical simulation technique used by Calabrese et al. (2012).

As an exponentiation is a monotone transformation, then we can change the problem into:

$$\max_{h} \beta_x x^\rho + \beta_h h^\rho + \alpha g_i^\rho + \alpha\kappa g_j^\rho. \tag{A-13}$$

First-order conditions:

$$-p\beta_x\rho(y - r - ph)^{\rho-1} + \rho\beta_h h^{\rho-1} = 0$$

$$-p\beta_x(y - r - ph)^{\rho-1} + \beta_h h^{\rho-1} = 0 \tag{A-14}$$

Solving for $h_d$ (i.e. Marshallian demand for housing)— from A-14:

$$\beta_h h^{\rho-1} = p\beta_x(y - r - ph)^{\rho-1}$$

$$\beta_h^{\frac{1}{\rho-1}} h = p^{\frac{1}{\rho-1}}\beta_x^{\frac{1}{\rho-1}}(y - r - ph)$$

$$\beta_h^{\frac{1}{\rho-1}} h + php^{\frac{1}{\rho-1}}\beta_x^{\frac{1}{\rho-1}} = (p\beta_x)^{\frac{1}{\rho-1}}(y - r)$$

$$h(\beta_h^{\frac{1}{\rho-1}} + p^{\frac{\rho}{\rho-1}}\beta_x^{\frac{1}{\rho-1}}) = (p\beta_x)^{\frac{1}{\rho-1}}(y - r)$$

$$h_d = \frac{(p\beta_x)^{\frac{1}{\rho-1}}(y - r)}{\beta_h^{\frac{1}{\rho-1}} + p^{\frac{\rho}{\rho-1}}\beta_x^{\frac{1}{\rho-1}}} \times \left[\frac{(1/\beta_h)^{\frac{1}{\rho-1}}}{(1/\beta_h)^{\frac{1}{\rho-1}}}\right]$$

$$h_d = \frac{(p\frac{\beta_x}{\beta_h})^{\frac{1}{\rho-1}}(y - r)}{1 + (p^\rho\frac{\beta_x}{\beta_h})^{\frac{1}{\rho-1}}}$$

Dividing the numerator and the denominator with $(p\dfrac{\beta_x}{\beta_h})^{\frac{1}{\rho-1}}$, we obtain:

$$h_d = \frac{y - r}{(p\frac{\beta_x}{\beta_h})^{\frac{1}{1-\rho}} + p} \tag{A-15}$$

Substituting $h_d$ into the utility function to obtain the indirect utility function, $\tilde{V}_i(p_i, g_i, g_j)$:

$$
\begin{aligned}
\tilde{V}_i(.) &= [\beta_x(y-r-h_d)^\rho + \beta_h h_d^\rho + \alpha g_i^\rho + \alpha\kappa g_j^\rho]^{1/\rho} \\
&= \left[\beta_x(y-r-\left\{\frac{(p\frac{\beta_x}{\beta_h})^{\frac{1}{\rho-1}}(y-r)}{1+(p^\rho\frac{\beta_x}{\beta_h})^{\frac{1}{\rho-1}}}\right\})^\rho + \beta_h\left\{\frac{(p\frac{\beta_x}{\beta_h})^{\frac{1}{\rho-1}}(y-r)}{1+(p^\rho\frac{\beta_x}{\beta_h})^{\frac{1}{\rho-1}}}\right\}^\rho + \alpha g_i^\rho + \alpha\kappa g_j^\rho\right]^{1/\rho} \\
&= \left[\beta_x(y-r-\left\{\frac{(p\frac{\beta_x}{\beta_h})^{\frac{1}{\rho-1}}(y-r)}{1+(p^\rho\frac{\beta_x}{\beta_h})^{\frac{1}{\rho-1}}}\right\})^\rho + \beta_h\frac{(p\frac{\beta_x}{\beta_h})^{\frac{\rho}{\rho-1}}}{[1+(p^\rho\frac{\beta_x}{\beta_h})^{\frac{1}{\rho-1}}]^\rho}(y-r)^\rho + \alpha g_i^\rho + \alpha\kappa g_j^\rho\right]^{1/\rho} \\
&= \left[\beta_x((y-r)(1-\frac{(p\frac{\beta_x}{\beta_h})^{\frac{1}{\rho-1}}}{1+(p^\rho\frac{\beta_x}{\beta_h})^{\frac{1}{\rho-1}}}))^\rho + \beta_h\frac{(p\frac{\beta_x}{\beta_h})^{\frac{\rho}{\rho-1}}}{[1+(p^\rho\frac{\beta_x}{\beta_h})^{\frac{1}{\rho-1}}]^\rho}(y-r)^\rho + \alpha g_i^\rho + \alpha\kappa g_j^\rho\right]^{1/\rho} \\
&= \left[\beta_x(y-r)^\rho(1-\frac{(p\frac{\beta_x}{\beta_h})^{\frac{1}{\rho-1}}}{1+(p^\rho\frac{\beta_x}{\beta_h})^{\frac{1}{\rho-1}}})^\rho + \beta_h\frac{(p\frac{\beta_x}{\beta_h})^{\frac{\rho}{\rho-1}}}{[1+(p^\rho\frac{\beta_x}{\beta_h})^{\frac{1}{\rho-1}}]^\rho}(y-r)^\rho + \alpha g_i^\rho + \alpha\kappa g_j^\rho\right]^{1/\rho} \\
&= \left[\beta_x(y-r)^\rho(\frac{1}{1+(p\frac{\beta_x}{\beta_h})^{\frac{1}{\rho-1}}})^\rho + \beta_h\frac{(p\frac{\beta_x}{\beta_h})^{\frac{\rho}{\rho-1}}}{[1+(p^\rho\frac{\beta_x}{\beta_h})^{\frac{1}{\rho-1}}]^\rho}(y-r)^\rho + \alpha g_i^\rho + \alpha\kappa g_j^\rho\right]^{1/\rho} \qquad \text{(A-16)}
\end{aligned}
$$

We can then rewrite A-16 as:

$$\tilde{V}(.) = [(y-r)^\rho\Phi(p) + \alpha g_i^\rho + \alpha\kappa g_j^\rho]^{1/\rho} \qquad \text{(A-17)}$$

where $\Phi(p) = \frac{\beta_x}{1+(p\frac{\beta_x}{\beta_h})^{\frac{1}{\rho-1}}} + \frac{\beta_h(p\frac{\beta_x}{\beta_h})^{\frac{\rho}{\rho-1}}}{[1+(p^\rho\frac{\beta_x}{\beta_h})^{\frac{1}{\rho-1}}]^\rho}$.

The boundary locus between jurisdiction $i$ and $j$ satisfies:

$$\tilde{V}_i(p_i, g_i, \kappa g_j) = \tilde{V}_j(p_j, g_j, \kappa g_i)$$

$$\Rightarrow [(y - r_i)^\rho \Phi(p_i) + \alpha g_i^\rho + \alpha\kappa g_j^\rho]^{1/\rho} = [(y - r_j)^\rho \Phi(p_j) + \alpha g_j^\rho + \alpha\kappa g_i^\rho]^{1/\rho}$$

$$\Rightarrow \alpha = \frac{(y - r_j)^\rho \Phi(p_j) - (y - r_i)^\rho \Phi(p_i)}{g_i^\rho - g_j^\rho + \kappa(g_j^\rho - g_i^\rho)} \tag{A-18}$$

$$\Rightarrow \ln\, \alpha_{ij}(y) = \ln\, \Big[\frac{(y - r_j)^\rho \Phi(p_j) - (y - r_i)^\rho \Phi(p_i)}{g_i^\rho - g_j^\rho + \kappa(g_j^\rho - g_i^\rho)}\Big] \tag{A-19}$$

To calculate community populations and incomes, we use the result from Epple and Sieg (1999) to get the standardized value of the joint distribution $(\ln y, \ln \alpha)$, $Z$, by utilizing A-19:

$$Z_{ij}(y) = \frac{\ln\Big[\frac{(y-r_j)^\rho \Phi(p_j) - (y-r_i)^\rho \Phi(p_i)}{g_i^\rho - g_j^\rho + \kappa(g_j^\rho - g_i^\rho)}\Big] - \mu_{\ln\alpha} - \lambda\sigma_{\ln\alpha}\frac{\ln y - \mu_{\ln y}}{\sigma_{lny}}}{\sigma_{\ln\alpha}\sqrt{1-\lambda^2}} \tag{A-20}$$

Following Epple and Sieg (1999) for Matlab function *zproca*:

$$\mu_{\ln(\alpha)|\ln(y)} = \mu_{\ln(\alpha)} + \lambda\sigma_{\ln(\alpha)}\frac{lny - \mu_{ln(y)}}{\sigma_{\ln(y)}} \tag{A-21}$$

$$\sigma_{\ln(\alpha)|\ln(y)} = \sigma_{\ln(\alpha)}\sqrt{1-\lambda^2} \tag{A-22}$$

$$Z_i = \frac{\ln\alpha - \mu_{\ln\alpha|\ln y} - \sigma^2_{\ln\alpha|\ln y}}{\sigma_{\ln\alpha|\ln y}} \tag{A-23}$$

*Deriving the pivotal voter locus*

In this model, we assume that households maximize their utility by consuming housing with respect to public good provision in their own jurisdiction since they vote only for taxation in the jurisdiction where they live.

Denoting $\bar{h}_d$ as the pivotal voters housing demand in equilibrium, then the utility function in equilibrium is:

$$U_i = [\beta_x(y - r_i - p_i\bar{h}_d^i)^\rho + \beta_h \bar{h_d^i}^\rho + \alpha g_i^\rho + \alpha\kappa g_j^\rho]^{1/\rho} \tag{A-24}$$

$$FOC(g_i) : \frac{dU_i}{dg_i} = \frac{1}{\rho}[\beta_x(y - r_i - p_i\bar{h_d^i})^\rho + \beta_h \bar{h_d^i}^\rho + \alpha g_i^\rho + \alpha\kappa g_j^\rho]^{\frac{1}{\rho}-1}$$

$$\times [\rho\beta_x(y - r_i - p_i\bar{h_d^i})^{\rho-1}(-\frac{dr_i}{dg_i} - \frac{dp_i}{dg_i}) + \rho\alpha g_i^{\rho-1}] = 0$$

Assuming $\frac{dr_i}{dg_i} = 0$ and $\frac{dp_i}{dg_i} = 1$, then:

$$
\begin{aligned}
&-\rho\beta_x(y - r_i - p\bar{h}_d^i)^{\rho-1} + \rho\alpha g_i^{\rho-1} = 0 \\
&\Rightarrow \alpha_i(y) = \frac{\beta_x(y - r_i - p_i\bar{h}_d^i)^{\rho-1}}{g_i^{\rho-1}}
\end{aligned}
$$

substituting A-15 for $\bar{h}_d^i$ :

$$
\begin{aligned}
\alpha_i(y) &= \frac{\beta_x\left(y - r_i - p_i\left\{\frac{(p_i\frac{\beta_x}{\beta_h})^{\frac{1}{\rho-1}}(y-r)}{1+(p_i^\rho\frac{\beta_x}{\beta_h})^{\frac{1}{\rho-1}}}\right\}\right)^{\rho-1}}{g_i^{\rho-1}} \\
&= \frac{\beta_x\left(y - r_i - p_i(p_i\frac{\beta_x}{\beta_h})^{\frac{1}{\rho-1}}\left\{\frac{(y-r)}{1+(p_i^\rho\frac{\beta_x}{\beta_h})^{\frac{1}{\rho-1}}}\right\}\right)^{\rho-1}}{g_i^{\rho-1}} \\
&= \frac{\beta_x\left(y - r_i - (p_i^\rho\frac{\beta_x}{\beta_h})^{\frac{1}{\rho-1}}\left\{\frac{(y-r)}{1+(p_i^\rho\frac{\beta_x}{\beta_h})^{\frac{1}{\rho-1}}}\right\}\right)^{\rho-1}}{g_i^{\rho-1}} \\
&= \frac{\beta_x\left(\frac{y[1+\cancel{(p_i^\rho\frac{\beta_x}{\beta_h})^{\frac{1}{\rho-1}}}]}{[1+(p_i^\rho\frac{\beta_x}{\beta_h})^{\frac{1}{\rho-1}}]} - \frac{r[1+\cancel{(p_i^\rho\frac{\beta_x}{\beta_h})^{\frac{1}{\rho-1}}}]}{[1+(p_i^\rho\frac{\beta_x}{\beta_h})^{\frac{1}{\rho-1}}]} - \frac{\cancel{(p_i^\rho\frac{\beta_x}{\beta_h})^{\frac{1}{\rho-1}}(y-r_i)}}{[1+(p_i^\rho\frac{\beta_x}{\beta_h})^{\frac{1}{\rho-1}}]}\right)^{\rho-1}}{g_i^{\rho-1}}
\end{aligned}
$$

$$
\alpha_i(y) = \frac{\beta_x\left(\frac{y-r}{1+(p_i^\rho\frac{\beta_x}{\beta_h})^{\frac{1}{\rho-1}}}\right)^{\rho-1}}{g_i^{\rho-1}} \tag{A-25}
$$

and consequently:

$$
\begin{aligned}
\ln\alpha(y) &= \ln\left[\frac{\beta_x\left(\frac{y-r}{1+(p_i^\rho\frac{\beta_x}{\beta_h})^{\frac{1}{\rho-1}}}\right)^{\rho-1}}{g_i^{\rho-1}}\right] \\
&= \ln\beta_x - (\rho-1)[\ln(1 + (p^\rho\frac{\beta_x}{\beta_h})^{\frac{1}{\rho-1}})) + \ln g + (\rho-1)\ln(y-r)]
\end{aligned}
$$

or we can relate to the procedure used in Calabrese et al. (2012):

$$\ln \alpha = \tilde{K}_i + (\rho - 1)\ln(y - r)$$

where

$$\tilde{K}_i = \ln(\beta_x) - (\rho - 1)\Big\{ \ln \Big(1 + (p^{\rho}\frac{\beta_x}{\beta_h})^{\frac{1}{\rho-1}}\Big) + \ln g_i\Big\} \tag{A-26}$$

Standardized values for pivotal voter based on Epple and Sieg (1999) then:

$$\begin{aligned}\tilde{Z}_i(y) &= \frac{\ln \alpha - \mu_{\ln \alpha} - \lambda \sigma_{\ln \alpha} \frac{\ln y - \mu_{\ln y}}{\sigma_{lny}}}{\sigma_{\ln y}\sqrt{1-\lambda^2}} \\ &= \frac{\tilde{K}_i + (\rho - 1)\ln(y - r) - \mu_{\ln \alpha} - \lambda \sigma_{\ln \alpha} \frac{\ln y - \mu_{\ln y}}{\sigma_{lny}}}{\sigma_{\ln y}\sqrt{1-\lambda^2}}\end{aligned} \tag{A-27}$$

**Tiebout allocation**

Assuming a CES utility function:

$$U_i = [\beta_x x^\rho + \beta_h h^\rho + \beta_{g_i}(\alpha) g_i^\rho + \beta_{g_j}(\alpha)\kappa g_j^\rho]^{1/\rho}, \tag{A-28}$$

where $\rho < 0$ and $\beta'_{g_i}, \beta_{g_j} > 0$. Let $\beta_g(\alpha) \equiv \alpha$ and $\beta_{g_j} \equiv \alpha_j$. Note that we suppress subscript $i$ on $x$ and $h$ for simplicity. Assuming the household preferences on public goods are the same for $i$ and $j$ then we have:

$$U_i = [\beta_x x^\rho + \beta_h h^\rho + \alpha g_i^\rho + \alpha\kappa g_j^\rho]^{1/\rho} \tag{A-29}$$

*Deriving indirect utility function*

$$\max_h \, [\beta_x x^\rho + \beta_h h^\rho + \alpha g_i^\rho + \alpha\kappa g_j^\rho]^{1/\rho} \tag{A-30}$$

As an exponentiation is a monotone transformation, then we can change the problem into:

$$\max_h \, \beta_x x^\rho + \beta_h h^\rho + \alpha g_i^\rho + \alpha\kappa g_j^\rho. \tag{A-31}$$

First-order conditions:

$$\begin{aligned} -p\beta_x\rho(y - ph)^{\rho-1} + \rho\beta_h h^{\rho-1} &= 0 \\ -p\beta_x(y - ph)^{\rho-1} + \beta_h h^{\rho-1} &= 0 \end{aligned} \tag{A-32}$$

Solving for $h_d$ (i.e. Marshallian demand for housing)— from A-32:

$$\beta_h h^{\rho-1} = p\beta_x(y-ph)^{\rho-1}$$

$$\beta_h^{\frac{1}{\rho-1}} h = p^{\frac{1}{\rho-1}}\beta_x^{\frac{1}{\rho-1}}(y-ph)$$

$$\beta_h^{\frac{1}{\rho-1}} h + php^{\frac{1}{\rho-1}}\beta_x^{\frac{1}{\rho-1}} = (p\beta_x)^{\frac{1}{\rho-1}} y$$

$$h(\beta_h^{\frac{1}{\rho-1}} + p^{\frac{\rho}{\rho-1}}\beta_x^{\frac{1}{\rho-1}}) = (p\beta_x)^{\frac{1}{\rho-1}} y$$

$$h_d = \frac{(p\beta_x)^{\frac{1}{\rho-1}} y}{\beta_h^{\frac{1}{\rho-1}} + p^{\frac{\rho}{\rho-1}}\beta_x^{\frac{1}{\rho-1}}} \times \left[\frac{(1/\beta_h)^{\frac{1}{\rho-1}}}{(1/\beta_h)^{\frac{1}{\rho-1}}}\right]$$

$$h_d = \frac{(p\frac{\beta_x}{\beta_h})^{\frac{1}{\rho-1}} y}{1 + (p^{\rho}\frac{\beta_x}{\beta_h})^{\frac{1}{\rho-1}}}$$

Dividing the numerator and the denominator with $(p\dfrac{\beta_x}{\beta_h})^{\frac{1}{\rho-1}}$, we obtain:

$$h_d = \frac{y}{(p\frac{\beta_x}{\beta_h})^{\frac{1}{1-\rho}} + p} \tag{A-33}$$

Substituting $h_d$ into the utility function to obtain the indirect utility function, $\tilde{V}_i(p_i, g_i, g_j)$:

$$
\begin{aligned}
\tilde{V}_i(.) &= [\beta_x(y-h_d)^\rho + \beta_h h_d^\rho + \alpha g_i^\rho + \alpha\kappa g_j^\rho]^{1/\rho} \\
&= \left[\beta_x(y - \left\{\frac{(p\frac{\beta_x}{\beta_h})^{\frac{1}{\rho-1}} y}{1+(p^\rho\frac{\beta_x}{\beta_h})^{\frac{1}{\rho-1}}}\right\})^\rho + \beta_h\left\{\frac{(p\frac{\beta_x}{\beta_h})^{\frac{1}{\rho-1}} y}{1+(p^\rho\frac{\beta_x}{\beta_h})^{\frac{1}{\rho-1}}}\right\}^\rho + \alpha g_i^\rho + \alpha\kappa g_j^\rho\right]^{1/\rho} \\
&= \left[\beta_x(y - \left\{\frac{(p\frac{\beta_x}{\beta_h})^{\frac{1}{\rho-1}} (y-r)}{1+(p^\rho\frac{\beta_x}{\beta_h})^{\frac{1}{\rho-1}}}\right\})^\rho + \beta_h\frac{(p\frac{\beta_x}{\beta_h})^{\frac{\rho}{\rho-1}}}{[1+(p^\rho\frac{\beta_x}{\beta_h})^{\frac{1}{\rho-1}}]^\rho} y^\rho + \alpha g_i^\rho + \alpha\kappa g_j^\rho\right]^{1/\rho} \\
&= \left[\beta_x(y(1 - \frac{(p\frac{\beta_x}{\beta_h})^{\frac{1}{\rho-1}}}{1+(p^\rho\frac{\beta_x}{\beta_h})^{\frac{1}{\rho-1}}}))^\rho + \beta_h\frac{(p\frac{\beta_x}{\beta_h})^{\frac{\rho}{\rho-1}}}{[1+(p^\rho\frac{\beta_x}{\beta_h})^{\frac{1}{\rho-1}}]^\rho} y^\rho + \alpha g_i^\rho + \alpha\kappa g_j^\rho\right]^{1/\rho} \\
&= \left[\beta_x y^\rho(1 - \frac{(p\frac{\beta_x}{\beta_h})^{\frac{1}{\rho-1}}}{1+(p^\rho\frac{\beta_x}{\beta_h})^{\frac{1}{\rho-1}}})^\rho + \beta_h\frac{(p\frac{\beta_x}{\beta_h})^{\frac{\rho}{\rho-1}}}{[1+(p^\rho\frac{\beta_x}{\beta_h})^{\frac{1}{\rho-1}}]^\rho} y^\rho + \alpha g_i^\rho + \alpha\kappa g_j^\rho\right]^{1/\rho} \\
&= \left[\beta_x y^\rho(\frac{1}{1+(p\frac{\beta_x}{\beta_h})^{\frac{1}{\rho-1}}})^\rho + \beta_h\frac{(p\frac{\beta_x}{\beta_h})^{\frac{\rho}{\rho-1}}}{[1+(p^\rho\frac{\beta_x}{\beta_h})^{\frac{1}{\rho-1}}]^\rho} y^\rho + \alpha g_i^\rho + \alpha\kappa g_j^\rho\right]^{1/\rho} \\
&= \left[y^\rho\frac{\beta_x}{[1+(p\frac{\beta_x}{\beta_h})^{\frac{1}{\rho-1}}]^\rho} + y^\rho\frac{\beta_h(p\frac{\beta_x}{\beta_h})^{\frac{\rho}{\rho-1}}}{[1+(p^\rho\frac{\beta_x}{\beta_h})^{\frac{1}{\rho-1}}]^\rho} + \alpha g_i^\rho + \alpha\kappa g_j^\rho\right]^{1/\rho}
\end{aligned} \tag{A-34}
$$

We can then rewrite A-34 as:

$$
\tilde{V}(.) = [y^\rho\Phi(p) + \alpha g_i^\rho + \alpha\kappa g_j^\rho]^{1/\rho} \tag{A-35}
$$

where $\Phi(p) = \frac{\beta_x}{1+(p\frac{\beta_x}{\beta_h})^{\frac{1}{\rho-1}}} + \frac{\beta_h(p\frac{\beta_x}{\beta_h})^{\frac{\rho}{\rho-1}}}{[1+(p^\rho\frac{\beta_x}{\beta_h})^{\frac{1}{\rho-1}}]^\rho}$.

The boundary locus between jurisdiction $i$ and $j$ satisfies:

$$\tilde{V}_i(p_i, g_i, \kappa g_j) = \tilde{V}_j(p_j, g_j, \kappa g_i)$$

$$\Rightarrow [y^\rho \Phi(p_i) + \alpha g_i^\rho + \alpha\kappa g_j^\rho]^{1/\rho} = [y^\rho \Phi(p_j) + \alpha g_j^\rho + \alpha\kappa g_i^\rho]^{1/\rho}$$

$$\Rightarrow \alpha_{ij} = \frac{y^\rho \Phi(p_j) - y^\rho \Phi(p_i)}{g_i^\rho - g_j^\rho + \kappa(g_j^\rho - g_i^\rho)} = \frac{y^\rho (\Phi(p_j) - \Phi(p_i))}{g_i^\rho - g_j^\rho + \kappa(g_j^\rho - g_i^\rho)} \tag{A-36}$$

$$\Rightarrow \ln \alpha_{ij}(y) = \ln \Big[\frac{y^\rho (\Phi(p_j) - \Phi(p_i))}{g_i^\rho - g_j^\rho + \kappa(g_j^\rho - g_i^\rho)}\Big] \tag{A-37}$$

We can rewrite as:

$$\ln \alpha_{ij}(y) = \rho \ln y + K_{ij} \tag{A-38}$$

$$\text{where } K_{ij} = \ln \Big[\frac{(\Phi(p_j) - \Phi(p_i))}{g_i^\rho - g_j^\rho + \kappa(g_j^\rho - g_i^\rho)}\Big] \tag{A-39}$$

$K_{ij}$ is the community intercept or variable *kint* in the matlab program.

Following Epple and Sieg (1999) to get the standardized value of the joint distribution $(\ln y, \ln \alpha)$, $Z(y)$:

$$Z_{ij}(y) = \Omega_j + \omega \ln(y) \tag{A-40}$$

$$\text{where } \omega_j = \omega_1 + \omega_2$$

$$\omega_1 = \frac{\rho}{\sigma_\alpha \sqrt{1 - \lambda^2}}$$

$$\omega_2 = \frac{-\lambda}{\sigma_{\ln y} \sqrt{1 - \lambda^2}}$$

$$\Omega = \frac{K_j - \mu_{\ln \alpha} + \lambda \sigma_{\ln \alpha} \frac{\mu_{\ln y}}{\sigma_{\ln y}}}{\sigma_{\ln y} \sqrt{(1 - \lambda^2)}}$$

*Deriving the pivotal voter locus*

The next step is deriving the pivotal voter locus. In this model, we assume that households maximize their utility by consuming houses with respect to public good provision in their own jurisdiction since they vote only for taxation in the jurisdiction where they live.

Denoting $\bar{h}_d$ as the pivotal voters housing demand in equilibrium, then the utility function in equilibrium is:

$$U_i = [\beta_x(y - p_i\bar{h}_d^i)^\rho + \beta_h \bar{h_d^i}^\rho + \alpha g_i^\rho + \alpha\kappa g_j^\rho]^{1/\rho} \tag{A-41}$$

$$\begin{aligned}
FOC(g_i) : \frac{dU_i}{dg_i} &= \frac{1}{\rho}[\beta_x(y - p_i\bar{h_d^i})^\rho + \beta_h \bar{h_d^i}^\rho + \alpha g_i^\rho + \alpha\kappa g_j^\rho]^{\frac{1}{\rho}-1} \\
&\times [-\rho\beta_x(y - p_i\bar{h_d^i})^{\rho-1}\frac{dp_i}{dg_i}\bar{h_d^i} + \rho\alpha g_i^{\rho-1}] = 0 \\
&\Rightarrow [-\rho\beta_x(y - p_i\bar{h_d^i})^{\rho-1}\frac{dp_i}{dg_i}\bar{h}_d + \rho\alpha g_i^{\rho-1}] = 0 \\
&\Rightarrow \frac{dp}{dg} = \frac{\alpha g^{\rho-1}}{\beta_x(y - p\bar{h_d^i})^{\rho-1}\bar{h_d^i}}
\end{aligned} \tag{A-42}$$

Substituting A-33 to A-42:

$$
\begin{aligned}
\frac{dp}{dg} &= \frac{\alpha g^{\rho-1}}{\beta_x (y - p\left[\frac{(p\frac{\beta_x}{\beta_h})^{\frac{1}{\rho-1}} y}{1+(p^{\rho}\frac{\beta_x}{\beta_h})^{\frac{1}{\rho-1}}}\right])^{\rho-1}\left[\frac{(p\frac{\beta_x}{\beta_h})^{\frac{1}{\rho-1}} y}{1+(p^{\rho}\frac{\beta_x}{\beta_h})^{\frac{1}{\rho-1}}}\right]} \\
&= \frac{\alpha g^{\rho-1}}{\beta_x (y - \left[\frac{(p^{\rho}\frac{\beta_x}{\beta_h})^{\frac{1}{\rho-1}} y}{1+(p^{\rho}\frac{\beta_x}{\beta_h})^{\frac{1}{\rho-1}}}\right])^{\rho-1}\left[\frac{(p\frac{\beta_x}{\beta_h})^{\frac{1}{\rho-1}} y}{1+(p^{\rho}\frac{\beta_x}{\beta_h})^{\frac{1}{\rho-1}}}\right]} \\
&= \frac{\alpha g^{\rho-1}}{\beta_x (y[1 - \left[\frac{(p^{\rho}\frac{\beta_x}{\beta_h})^{\frac{1}{\rho-1}}}{1+(p^{\rho}\frac{\beta_x}{\beta_h})^{\frac{1}{\rho-1}}}\right]])^{\rho-1}\left[\frac{(p\frac{\beta_x}{\beta_h})^{\frac{1}{\rho-1}} y}{1+(p^{\rho}\frac{\beta_x}{\beta_h})^{\frac{1}{\rho-1}}}\right]} \\
&= \frac{\alpha g^{\rho-1}}{\beta_x (y\left[\frac{1}{1+(p^{\rho}\frac{\beta_x}{\beta_h})^{\frac{1}{\rho-1}}}\right])^{\rho-1}\left[\frac{(p\frac{\beta_x}{\beta_h})^{\frac{1}{\rho-1}} y}{1+(p^{\rho}\frac{\beta_x}{\beta_h})^{\frac{1}{\rho-1}}}\right]} \\
&= \frac{\alpha g^{\rho-1}}{\beta_x y^{\rho-1}\frac{1}{[1+(p^{\rho}\frac{\beta_x}{\beta_h})^{\frac{1}{\rho-1}}]^{\rho-1}}\left[\frac{(p\frac{\beta_x}{\beta_h})^{\frac{1}{\rho-1}}}{1+(p^{\rho}\frac{\beta_x}{\beta_h})^{\frac{1}{\rho-1}}}\right] y} \\
&= \frac{\alpha g^{\rho-1}}{\beta_x y^{\rho}\left[\frac{(p\frac{\beta_x}{\beta_h})^{\frac{1}{\rho-1}}}{1+(p^{\rho}\frac{\beta_x}{\beta_h})^{\frac{1}{\rho-1}}}\right]} = \frac{\alpha g^{\rho-1}}{\frac{\beta_x y^{\rho}}{1+(p^{\rho}\frac{\beta_x}{\beta_h})^{\frac{1}{\rho-1}}}(p\frac{\beta_x}{\beta_h})^{\frac{1}{\rho-1}}} \\
&= \frac{\alpha g^{\rho-1}}{\left[\frac{y}{1+(p^{\rho}\frac{\beta_x}{\beta_h})^{\frac{1}{\rho-1}}}\right]^{\rho}\beta_x^{\frac{\rho}{\rho-1}}\left(\frac{\beta_h}{p}\right)^{\frac{1}{1-\rho}}} \\
&= \frac{\alpha g^{\rho-1}}{\left[\frac{y}{1+(p^{\rho}\frac{\beta_x}{\beta_h})^{\frac{1}{\rho-1}}}\right]^{\rho}\beta_x^{\frac{\rho}{\rho-1}}\left(\frac{\beta_h}{p}\right)^{\frac{\rho}{1-\rho}}\frac{\beta_h}{p}} \\
&= \frac{\alpha g^{\rho-1}}{\left[\frac{y}{1+(p^{\rho}\frac{\beta_x}{\beta_h})^{\frac{1}{\rho-1}}}\left(\frac{\beta_h}{p}\right)^{\frac{1}{1-\rho}}\beta_x^{\frac{1}{\rho-1}}\right]^{\rho}\frac{\beta_h}{p}} \\
&= \frac{\alpha g^{\rho-1}}{\left[\frac{y}{1+(p^{\rho}\frac{\beta_x}{\beta_h})^{\frac{1}{\rho-1}}}\left(\frac{\beta_h}{p\beta_x}\right)^{\frac{1}{1-\rho}}\right]^{\rho}\frac{\beta_h}{p}}
\end{aligned}
\tag{A-43}
$$

$$= \frac{\alpha g^{\rho-1}}{\left[\frac{y}{[1+(p\frac{\beta_x}{\beta_h})^{\frac{1}{\rho-1}}p](p\frac{\beta_x}{\beta_h})^{\frac{1}{1-\rho}}}\right]^{\rho}\frac{\beta_h}{p}}$$

$$= \frac{\alpha g^{\rho-1}}{\left[\frac{y}{(p\frac{\beta_x}{\beta_h})^{\frac{1}{1-\rho}}+(p\frac{\beta_x}{\beta_h})^{0}p}\right]^{\rho}\frac{\beta_h}{p}}$$

$$= \frac{\alpha g^{\rho-1}}{\left[\frac{y}{p+(p\frac{\beta_x}{\beta_h})^{\frac{1}{1-\rho}}}\right]^{\rho}\frac{\beta_h}{p}} \tag{A-44}$$

The condition of the pivotal voter's indifference curves and the government-services possibility frontier (GPF) when households are myopic is:

$$\left.\frac{dp}{dg}\right|_{u=\bar{u}} = \left.\frac{dp}{dg}\right|_{gpf} \tag{A-45}$$

$$\frac{\alpha g^{\rho-1}}{\left[\frac{y}{p+(p\frac{\beta_x}{\beta_h})^{\frac{1}{1-\rho}}}\right]^{\rho}\frac{\beta_h}{p}} = \left.\frac{dp}{dg}\right|_{gpf}$$

$$\Rightarrow \ln\alpha = \tilde{K_j} + \rho\ln y \tag{A-46}$$

$$\text{where } \tilde{K_j} = -\rho\ln(p+(p\frac{\beta_x}{\beta_h})^{\frac{1}{1-\rho}}) + \ln\beta_x - \ln p - (\rho-1)\ln g + \ln\frac{dp}{dg}$$

$\frac{dp}{dg}$ in the last term is variable *dpdgv* in the Matlab function, *ktil*, fed by another function called *myopic*.

## A.3 Matlab codes

HEAD TAX

This Matlab program is written based on GAUSS program used by Calabrese et al. (2012). There are 5 communities in the metropolitan area that impose an efficient tax (head-tax). Although we assume *absentee landlords*, we take into account their welfare in the CV calculation.

The main script below will call ten functions (procedures): func.m, kcalc.m, zproc.m, zproca.m, summrs.m,poproc.m, inproc.m, cvcalc.m, montc.m, and nlsrch.m.

```
%% Main script

clear all;
clc;

%% set sample size for Monte Carlo simulation
smpl=1000000;
%% generate repeatable random number with the Mersenne twister generator
s = RandStream('mt19937ar','Seed',1); % MATLAB's start-up settings
RandStream.setDefaultStream(s);
%% FOR MATLAB2013 users:
% rng(1,'twister');
% s=rng;

eps1=randn(smpl,1);
eps2=randn(smpl,1);

kappa_vec=0:0.1:1; %kappa is a parameter of public good spillovers
pot_gain=zeros(1,length(kappa_vec));

for iii=1:length(kappa_vec)
        kappa=kappa_vec(iii);

%for ii=1:length(pot_gain)
```

```
warning off; % turning off Warnings on GLOBAL VARIABLE
ncom=5;    % Number of communities

Montind=1; %Montind=1 to do Monte Carlo comparison


%% Parameters
muy= 10.51710439;
sigy= 0.88623;
mual= -2.36406786786;
sigal= 0.00990949382558 ;
lam = 0;
gm=.25;
ALx= 1;
wq =1;
Theta = (1-gm)/gm;
ALh =0.355945238972 ;
rho = -.01;

%% I use "for loop" to create a plot of kappa vs. potential welfare gain


land=[.4 .15 .15 .15 .15]';

zmin =ones(1,48)*(-1000);
lnalphamax = ones(1,48)*10;
hsuply=ones(ncom,1);
iter=0;
bigind=0;
iwrite=0;
nlfac=1;
big=100;

global ncom kappa muy sigy mual sigal lam gm ALx wq Theta ALh rho yvec...
lyvec lnalphamax diff wvec iwrite tol bignum maxits nlfac big ibig bigind...
land mual zmin wvec smpl eps1 eps2 gc;

incnu=ones(ncom,1);
popn=ones(ncom,1);
hd=ones(ncom,1);
%dpdgv=ones(ncom,1);

meanincandalpha=[exp(muy+sigy^2/2) exp(mual+sigal^2/2)];
display('mean income and alpha'); disp(num2str(meanincandalpha))
```

```
%% From Calabrese et.al. (2012):
%% evec is a vector of ordinates for numerical integration where the values
%% are normalized to the interval (-1,+1). If a function of a variable x
%% is to be integrated over an interval [xl,xh], the ordinates
%% normalized to that interval are obtained by
%% x=(0.5*((xl+xh)+(diff.*evec'))) where diff=xh-xl.

evec=[...
-0.9987710 -0.9935302 -0.9841246 -0.9705916 -0.9529877,-0.9313867,...
-0.9058791 -0.8765720 -0.8435883 -0.8070662 -0.7671590,-0.7240341,...
-0.6778724 -0.6288674 -0.5772247 -0.5231610 -0.4669029 -0.4086865,...
-0.3487559 -0.2873625 -0.2247638 -0.1612224 -0.0970047 -0.03238017,...
0.03238017 0.09700470 0.1612224 0.2247638 0.2873625 0.3487559,...
0.40868650 0.46690290 0.5231610 0.5772247 0.6288674 0.6778724,...
0.72403410 0.76715900 0.8070662 0.8435883 0.8765720 0.9058791,...
0.93138670 0.95298770 0.9705916 0.9841246 0.9935302 0.9987710]' ;

%% Wvec is a vector of interval widths normalized to conform to the
%% normalization of evec. To renormalize to an interval [xl,xh],
%% multiply by diff/2 where diff=xh-xl.


wvec=[...
0.00315335 0.00732755 0.01147723 0.01557932 0.01961616 0.02357076,...
0.02742651 0.03116723 0.03477722 0.03824135 0.04154508 0.04467456,...
0.04761666 0.05035904 0.05289019 0.05519950 0.05727729 0.05911484,...
0.06070444 0.06203942 0.06311419 0.06392424 0.06446616 0.06473770,...
0.06473770 0.06446616 0.06392424 0.06311419 0.06203942 0.06070444,...
0.05911484 0.05727729 0.05519950 0.05289019 0.05035904 0.04761666,...
0.04467456 0.04154508 0.03824135 0.03477722 0.03116723 0.02742651,...
0.02357076 0.01961616 0.01557932 0.01147723 0.00732755 0.00315335]' ;


xl=-5;
xh=3;
diff=xh-xl;
lyvec=muy +sigy*(0.5*((xl+xh)+(diff.* evec')));

yvec=exp(lyvec);

kmin=xl;

yvec=repmat(yvec,ncom,1); %% transform yvec from 1x48 into 5x48
```

```
%% initial guess
 x0=[10 10 10 10 10 1832 4623 6544 9626 17371 1832 4623 6544 9626 17371]';



%% Centralized property-tax equilibrium values

pc =  [17.1321740284 17.1321740284 17.1321740284 17.1321740284...
17.1321740284]';
capqc = (((1+((pc.^rho)*ALx/ALh).^(1/(rho-1))).^rho).^(-1))*ALx +...
ALh*(pc*ALx/ALh).^(rho/(rho-1)))./(1+((pc.^rho)*ALx/ALh).^(1/...
(rho-1))).^rho;
tc =[.35    .35     .35     .35     .35]';
gc =[3829.71688669  3829.71688669  3829.71688669  3829.71688669...
3829.71688669]';


iwrite=0;
tol=.000001;
maxits=1000;
nlfac=1;
dh=1e-4;
iter=1000;
big=1000;
ibig=0;
bignum=.125;
bigind=0;
nladj=1.05;


%%======================Loop========================%%

xsave = x0;
x=xsave;
count=ones(size(x,1),1);
icnt=0;
while icnt<size(count,1);
icnt=icnt+1;
count(icnt,1)=icnt;
end;

allfun=abs(func(x));
maxfun=max(allfun);% create a row vector with max values for each column
nladj=1.05;
if maxfun<20;
```

```
nladj=1.5;
end
dh=.0001;
nlfac=.05;


x = nlsrch(x,tol,bignum,maxits,nlfac);

xsave=x;
funcout=func(x);

pr=[ncom muy sigy mual sigal lam ALh ALx rho gm Theta];
disp('parameters:');
disp({'ncom' 'muy' 'sigy' 'mual' 'sigal' 'lam' 'ALh' 'ALx' 'rho'...
'gm' 'Theta'
 'kappa'; ncom muy sigy mual sigal lam ALh ALx rho gm Theta kappa});


disp('result p, g t are'); %% t is a headtax or r in in this program
p=x(1:ncom,1);
g=x(ncom+1:2*ncom,1);
r=x(2*ncom+1:3*ncom,1);

disp(num2str([p g r]))

kv=kcalc(p,g,r,yvec,lyvec);

zmat=zproc(kv,lyvec);
inc=inproc(zmat);
popn=poproc(zmat);


display('Incomes, Populations, Mean Incomes')
disp(num2str([inc popn inc./popn]))

display('Sum of Community Incomes Compared to Total Metropolitan Income')
disp(num2str([sum(inc) exp(muy+sigy^2/2)]))


hd = (p+(ALx*p./ALh).^(1/(1-rho))).^(-1).*(inc-r.*popn);
hsuply=land.*p.^Theta*((1-gm)/wq).^Theta;

display('Housing Demands and Supplies, housing share of income,...
Land Areas')
disp(num2str([hd,hsuply,p.*hd./inc,land]))
```

```
if Montind==1;
[kvrnd,yave,houave]=montc([p;g;r],capqc);
monres=[kvrnd yave houave];

display('Compare Monte Carlo and quadrature')

display('Community Sizes')
disp(num2str([monres(:,1) popn]))

display('community mean incomes')
disp(num2str([monres(:,2) (inc./popn)]))

display('mean housing demand')
disp(num2str([monres(:,3),hd./popn]))


[avgcvad,distbetad] = cvcalc([p;g;r],capqc);
cvdist=[avgcvad,distbetad];

disp('avgcv and %better off');
disp(num2str([avgcvad; distbetad]))

rentsa =gm*((1-gm)/wq).^((1-gm)/gm)*sum(land.*(pc./(1+tc)).^(1/gm))';

rentsb = gm*((1-gm)/wq).^((1-gm)/gm)*sum(land.*p.^(1/gm))';

display(' Rents a, b ')
disp(num2str( [rentsa,rentsb]))
Rentsch= rentsb - rentsa;

display(' Change in Rents=')
disp(num2str(Rentsch))

display('aggregate cv + change in rents');
disp(num2str(cvdist(1,1)-Rentsch))

pot_gain(iii)=-((cvdist(1,1)-Rentsch));


end
h=plot(kappa_vec,pot_gain)
set(h,'linewidth',2);
title('Spillover vs. Potential Gain, A Head-Tax Case (US)')
xlabel('\kappa')
```

```
ylabel('Potential gain (US$)')

%% this is the end of the main script
```

```
%% func.m

function[fun]=func(x)
global ncom gm ALx wq Theta ALh rho yvec lyvec iwrite nlfac big bigind land;

xs=x;

fun=zeros(3*ncom,1);

p=x(1:ncom,1);
g=x(ncom+1:2*ncom,1);
r=x(2*ncom+1:3*ncom,1);

kv=kcalc(p,g,r,yvec,lyvec);
zmat=zproc(kv,lyvec);
zmata = zproca(kv,lyvec);

popn=poproc(zmat);
incnu=inproc(zmat);

hd = (p+(ALx*p./ALh).^(1/(1-rho))).^(-1).*(incnu-r.*popn);

if iwrite <100;
iwrite=iwrite+1;
elseif iwrite>09;
iwrite=0;
nlfac=min([1; 1.05*nlfac],[],1); % it will return a min value
                                 % between 1 and 1.05*nlfac

display({'nlfac big ibig' nlfac big bigind});
end

hsuply=land.*p.^Theta.*((1-gm)/wq).^Theta;
summrsxg = summrs(p,g,r,yvec,lyvec,zmata);


fun(1:ncom,1)=hd-hsuply;
fun(ncom+1:2*ncom,1)=r-g;
fun(2*ncom+1:3*ncom,1)=summrsxg-popn;
x=xs;
fun;
end
%% this is the end of func.m
```

```
%% kcalc.m

function[lnalpha]=kcalc(pvect,gvec,rvec,yvec,lyvec)
global  ALx ALh rho yvec lyvec lnalphamax ncom kappa;
rvec=repmat(rvec,1,48);
pvect=repmat(pvect,1,48);
aty = (yvec-rvec);

%%capq is ncomx48
capq=(aty.^rho).*(ALx./(1+(pvect.^rho*ALx/ALh).^(1/(rho-1))).^rho +...
(ALh*(pvect*ALx/ALh).^(rho/(rho-1)))./(1+(pvect.^rho*ALx/ALh).^(1/...
(rho-1))).^rho);

%% alpha is 4x48
alpha=(capq(2:ncom,:)-capq(1:(ncom-1),:))./repmat((gvec(1:...
(ncom-1),1).^rho-gvec(2:ncom,1).^rho)+kappa*((gvec(2:ncom,1).^rho-gvec...
(1:(ncom-1),1).^rho)),1,48);
% Repmat replicate the denominator from 4x1 into 4x48;
% divide numerator with the denominator element by element.

lnalpha = log(alpha);
lnalpha(imag(lnalpha)~=0)= 10;
lnalpha = [lnalpha; lnalphamax];

end

%% this is the end of kcalc.m
```

```
%% zproc.m

function[zv]= zproc(kv,lyv)

global ncom muy sigy mual sigal lam;

%% I use repmat to generate a (ncomx48) matrix with uniform values.
omgv = (kv./(sigal*(1-lam^2).^(.5)))-repmat((mual/(sigal*(1-lam^2).^(.5...
))),ncom,48)-lam*((repmat(lyv,ncom,1)-repmat(muy,ncom,48...
))./sigy)./(1-lam^2).^(.5);
zv=omgv;
end
%% this the end of zproc.m
```

```
%% zproca.m

function[zv]= zproca(kv,lyv)

global ncom muy sigy mual sigal lam ;

mualy = repmat(mual,1,48)+lam*sigal.*(lyv-repmat(muy,1,48)./sigy); %% 1x48
sigaly = sqrt(1-lam.^2)*sigal; %% scalar
zv= (kv-repmat(mualy,ncom,1)-repmat(sigaly.^2,ncom,48))./sigaly;
zv;
end
%% this is the end of zproca.m
```

```
%% summrs.m
function[summrsxg]=summrs(p,g,r,yv,lyv,zmata)

global mual ncom lam sigal muy sigy rho ALx ALh zmin diff wvec;

mualy = mual+lam*sigal.*((lyv-muy)./sigy);
sigaly = sqrt(1-lam^2)*sigal;

r=repmat(r,1,48);
p=repmat(p,1,48);
g=repmat(g,1,48);
capq = g.^(rho-1)./((ALx.*((yv-r)-(p.*(yv-r))./(p+(p.*ALx/ALh).^(1/...
(1-rho)))).^(rho-1)));
capq(isnan(capq))= 0;
capq(imag(capq)~=0)=0;

zmatb = [zmin;zmata(1:(ncom-1),:)];


summrsxg= (diff/2)*((repmat(exp(mualy+sigaly^2/2),ncom,1).*...
(normcdf(zmata)-normcdf(zmatb)).*capq).*repmat(normpdf((lyv-muy)/...
sigy),ncom,1)*wvec);

end
%% this is the end of summrs.m




%%poproc.m

function[popn]=poproc(zmat)

global ncom muy sigy lyvec diff wvec;

%% zmat is 5x48, where 5=ncom
cumpop=(diff/2).*(repmat(normpdf((lyvec-repmat(muy,1,48))/sigy),ncom,1)...
).*normcdf(zmat)*wvec;

cumpop=[0;cumpop];

popn=(cumpop(2:(ncom+1),1)-cumpop(1:ncom,1));

end
%% this is the end of poproc.m
```

```
%%inproc.m

function[incnu]=inproc(zmat)

global ncom muy sigy lyvec diff wvec;

cincnu=((diff/2).*repmat(exp(lyvec).*normpdf((lyvec-...
repmat(muy,1,48))/sigy),ncom,1).*normcdf(zmat)*wvec);
cincnu=[0;cincnu];
incnu=(cincnu(2:(ncom+1),1)-cincnu(1:ncom,1));
end
%% this is the end of inproc.m
```

```
%% cvcalc.m

function[avgcvad,distbetad]=cvcalc(x,capqc)

global smpl ncom ALx ALh rho muy sigy lam eps1 eps2 mual sigal gc kappa;

cv=zeros(smpl,1);

p=x(1:ncom,1);
g=x(ncom+1:2*ncom,1);
r=x(2*ncom+1:3*ncom,1);

capq=ALx./(1+((p.^rho)*ALx/ALh).^(1./(rho-1))).^rho +...
ALh*(p*ALx/ALh).^(rho/(rho-1)))./(1+((p.^rho)*ALx/ALh).^(1/(rho-1))).^rho;
yrnd=exp(muy+sigy*((1-lam^2).^(.5))*eps1+lam*sigy*eps2);


alprnd=exp(mual+sigal*eps2);


utlc = (repmat(alprnd',5,1).*repmat(gc,1,1000000).^rho+kappa*repmat...
(alprnd',5,1).*repmat(gc,1,1000000).^rho+repmat(capqc,1,1000000).*repmat...
(yrnd',5,1).^rho).^(1/rho);

utl=(repmat(alprnd',5,1).*repmat(g,1,1000000).^rho+kappa*repmat...
(alprnd',5,1).*repmat(g,1,1000000).^rho+repmat(capq,1,1000000).*(repmat...
(yrnd,1,5)-repmat(r',1000000,1))'.^rho).^(1/rho);

cvall = ((utlc.^rho - repmat(alprnd',5,1).*repmat(g,1,1000000).^rho-...
kappa*repmat(alprnd',5,1).*repmat(g,1,1000000).^rho)./repmat...
(capq,1,1000000)).^(1/rho)-(repmat(yrnd,1,5)-repmat(r',1000000,1))';

ordstat2=min(cvall,[],1)';
%% Return the minimum of matrix cvall from each column,
%% and put them as a COLUMN vector

avgcvad= sum(ordstat2)'/smpl;

distadad = logical(ordstat2<=0);

distbetad = sum(distadad)'/smpl;

end
%% this is the end of cvcalc.m
```

```
%% montc.m

function[kvrnd,yave,houave]=montc(x,capqc)

global smpl ncom ALx ALh rho muy sigy lam eps1 eps2 mual sigal gc kappa;

cv=zeros(smpl,1);

p=x(1:ncom,1);
g=x(ncom+1:2*ncom,1);
r=x(2*ncom+1:3*ncom,1);

capq=ALx./(1+((p.^rho)*ALx/ALh).^(1/(rho-1))).^rho +...
(ALh*(p*ALx/ALh).^(rho/(rho-1)))./(1+((p.^rho)*ALx/ALh).^(1/(rho-1))).^rho;
yrnd=exp(muy+sigy*((1-lam^2).^.5)*eps1+lam*sigy*eps2);

alprnd=exp(mual+sigal*eps2);

utl=-(repmat(alprnd',5,1).*repmat(g,1,1000000).^rho+kappa*repmat...
(alprnd',5,1).*repmat(g,1,1000000).^rho+repmat...
(capq,1,1000000).*(repmat(yrnd,1,5)-repmat(r',1000000,1))'.^rho);

utlc = (repmat(alprnd',5,1).*repmat(gc,1,1000000).^rho+kappa*repmat...
(alprnd',5,1).*repmat(gc,1,1000000).^rho+...
repmat(capqc,1,1000000).*repmat(yrnd',5,1).^rho).^(1/rho);

cv = ((utlc.^rho - repmat(alprnd',5,1).*repmat(g,1,1000000).^rho-...
kappa*repmat(alprnd',5,1).*repmat(g,1,1000000).^rho)./repmat...
(capq,1,1000000)).^(1/rho)-(repmat(yrnd,1,5)-repmat(r',1000000,1))';

ordstat1=min(cv,[],1)';
%% Return the minimum of a matrix cv from each column

avgcv = sum(ordstat1)'/smpl;

e1=logical(repmat(ordstat1,1,5)==cv');

kvrnd=sum(e1)'/smpl;
yave=(sum(e1.*repmat(yrnd,1,5))'/smpl)./kvrnd;

houave=(p+(ALx*p./ALh).^(1/(1-rho))).^(-1).*(sum(e1.*(repmat(yrnd,1,5)-...
repmat(r',1000000,1)))'/smpl)./kvrnd;

end
%% this is the end of montc.m
```

```
%% nlsrch.m

% Note that I have 2 local functions in this particular script:grad1 and bv
function[x1]=nlsrch(x0,tol,bignum,maxits,nlfac)

global big bigind ibig

k=size(x0,1);

fv0=func(x0);

dx=1; iter=1; big=1000; ibig=0;
option1=0; option2=1; option3=1; option4=1;

if option4 == 0;

disp('******** Solution of Non-Linear Equation System *********** ');

end;

tic;

while (abs(dx)>=tol & iter <= maxits &  big >= tol)
jc=grad1(fv0,x0,k);


x1=x0 - nlfac*inv(jc)*fv0;

% evaluate the vector function with the new estimates
 fv1=func(x1);

%  Compute the largest function value in the vector of functions
 big=max(abs(fv1));


% Define the index of the largest function value in  the vector of functions.
[saka,Indices]=max(abs(fv1)); ibig=Indices' ;

%%Define index for which function is largest.
[arana,index]=max(abs(fv1)); bigind=index';


if big<bignum;
nlfac=max([((bignum-.95*big)/bignum);nlfac])';
  end;
```

```
if option1 == 1;

disp(['Iteration = ' num2str(iter)])

disp(['f(x0) = ']); disp([ num2str(fv0)]);
disp(['New estimates = ']); disp([num2str(x1)]);

end;

if option2 == 1;

   if abs(fv1) > abs(fv0);

    disp('The procedure is diverging. The current values are: ')

    disp(['Iteration = ' num2str(iter)])

    disp(['f(x0) = ']); disp([ num2str(fv0)])
    disp(['f(x1) = ']); disp([ num2str(fv1)])
    disp(['f(x0) = ']); disp([ num2str(fv0)])
    disp(['x1 = ']); disp([ num2str(x1)])

    disp('  Try new starting values. ');

    end;
end;

dx=x1-x0;   x0=x1;


 fv0=func(x0);

  iter=iter + 1;

end;

big=max(abs(fv1))';

if option4 == 0;  %

disp(['Largest element of Abs(f(x)) is: ' num2str(big)])

end;

if iter >  maxits;
```

```
disp('Iteration limit exceeded:');

disp(['Largest element of Abs(f(x)) is: ' num2str(big)])

disp(['Final estimates for x are: ']); disp([num2str(x1)])
end;

if option3 == 1;
   tme=toc

    if option4==0;
    disp(['Total time required: ' num2str(tme) 'seconds'])
    disp(['Total number of iterations:' num2str(iter-1)])
    end;

end;
end

%%Jacobian matrix

function[g]=grad1(gf0,gx0,k);

dh=1e-4;
g=zeros(k,k);

i=1;
    while i<=k;

g(:,i)=(func(bv(gx0,1,i,k))-gf0)/dh;

i=i+1;
    end;
end

function[bv_output]=bv(b,a,i,k)

gee=eye(k); %%% k=15 , gee is an identity matrix 15x15
dh=1e-4;
bv_output=b+(a*dh)*gee(:,i);
end
%% this is the end of nlsrch.m
```

PROPERTY TAX

This Matlab program is written based on GAUSS program used by Calabrese et al. (2012). There are 5 communities in the metropolitan area with property tax. Although we assume *absentee landlords*, we take into account their welfare in the CV calculation. We assume lognormal distributions for both income and preference parameter for public good. As in Calabrese et al. (2012) we also assume that voters are myopic.

The main script below will call ten functions (procedures): myopic.m, trans.m, untrans.m, func.m, kcalc.m, ktil.m, zproc.m, voproc.m, poproc.m, inproc.m, cvcalc.m, montc.m, and nlsrch.m. Note that nlsrch.m used in this program is the same as nlsrch.m used in the head tax case.

```
clear all;
clc;

warning off; %%% turning off Warnings on GLOBAL VARIABLE

ncom=5;     % Number of communities

%% creating repeatable random number with the mersenne twister generator
s = RandStream('mt19937ar','Seed',1); % MATLAB's start-up settings
RandStream.setDefaultStream(s);
%% Set following to 1 to do Monte Carlo comparison%%

kappa_vec=0:0.1:0.9;
pot_gain=zeros(1,length(kappa_vec));
tax_vec=zeros(ncom, length(kappa_vec));
perc_better=zeros(1,length(kappa_vec));
popn_vec=zeros(ncom, length(kappa_vec));
meaninc_vec=zeros(ncom, length(kappa_vec));


for iii=1:length(kappa_vec)
        kappa=kappa_vec(iii);
```

```
Montind=1;


%%Initial Parameters%%

muy=10.51710439;
sigy=0.88623;
mual = -2.36406786786;
sigal=   0.00990949382558 ;
lam = 0;
gm=.25;
ALx= 1;
wq =1;
Theta = (1-gm)/gm;
ALh =0.355945238972 ;
rho = -.01;


land=[.4 .15 .15 .15 .15]';


omg1=0;
omg2=0;
gex=0;
kmax=10;
hsuply=ones(ncom,1);
iter=0;
iloop=0;
mxt=100*ones(ncom,1);
bigind=0;
parad=0;
iwrite=0;
nlfac=1;
bigpct=100;
big=100;
taxbal=.2*ones(ncom,1);

global kappa ncom muy sigy mual sigal lam gm ALx wq Theta ALh rho yvec...
lyvec diff wvec iwrite nlfac tol bignum maxits big ibig bigind land mual...
wvec smpl eps1 eps2 gc mxt kmax;

incnu=ones(ncom,1);
popn=ones(ncom,1);
hd=ones(ncom,1);
```

```
dpdgv=ones(ncom,1);

meanincandalpha=[exp(muy+sigy^2/2) exp(mual+sigal^2/2)];
display('mean income and alpha'); disp(num2str(meanincandalpha))

%% From Calabrese et.al. (2012):
% evec is a vector of ordinates for numerical integration.  These
% ordinates are normalized to the inteval (-1,+1).
% if a function of a variable x is to be integrated over an interval
% [xl,xh], the ordinates normalized to that interval are obtained by
% x=(0.5*((xl+xh)+(diff.*evec'))) where diff=xh-xl.

evec=[...
-0.9987710 -0.9935302 -0.9841246 -0.9705916 -0.9529877,-0.9313867,...
-0.9058791 -0.8765720 -0.8435883 -0.8070662 -0.7671590,-0.7240341,...
-0.6778724 -0.6288674 -0.5772247 -0.5231610 -0.4669029 -0.4086865,...
-0.3487559 -0.2873625 -0.2247638 -0.1612224 -0.0970047 -0.03238017,...
0.03238017 0.09700470 0.1612224 0.2247638 0.2873625 0.3487559,...
0.40868650 0.46690290 0.5231610 0.5772247 0.6288674 0.6778724,...
0.72403410 0.76715900 0.8070662 0.8435883 0.8765720 0.9058791,...
0.93138670 0.95298770 0.9705916 0.9841246 0.9935302 0.9987710]' ;


%% Wvec is a vector of interval widths normalized to conform to the
% normalization of evec above.  To renormalize to an interval [xl,xh],
% 5 multiply by diff/2 where diff=xh-xl.


wvec=[...
0.00315335 0.00732755 0.01147723 0.01557932 0.01961616 0.02357076,...
0.02742651 0.03116723 0.03477722 0.03824135 0.04154508 0.04467456,...
0.04761666 0.05035904 0.05289019 0.05519950 0.05727729 0.05911484,...
0.06070444 0.06203942 0.06311419 0.06392424 0.06446616 0.06473770,...
0.06473770 0.06446616 0.06392424 0.06311419 0.06203942 0.06070444,...
0.05911484 0.05727729 0.05519950 0.05289019 0.05035904 0.04761666,...
0.04467456 0.04154508 0.03824135 0.03477722 0.03116723 0.02742651,...
0.02357076 0.01961616 0.01557932 0.01147723 0.00732755 0.00315335]' ;


xl=-5;
xh=3;
diff=xh-xl;
lyvec=muy +sigy*(0.5*((xl+xh)+(diff.* evec')));

yvec=exp(lyvec);
```

```
kmin=xl;


%% Initial guess

                 x0=[ 14.258...
                      16.185...
                      17.175...
                      18.391...
                      20.797...
                      1880.980...
                      3063.992...
                      3849.785...
                      5006.582...
                      8016.818...
                     0.3488...
                     0.3507...
                     0.3518...
                     0.3529...
                     0.3545]';



%% Centralized property-tax equilibrium values

pc =  [17.1321740284 17.1321740284 17.1321740284 17.1321740284...
17.1321740284]';
capqc = (((1+((pc.^rho)*ALx/ALh).^(1/(rho-1))).^rho).^(-1))*ALx +...
(ALh*(pc*ALx/ALh).^(rho/(rho-1)))./...
(1+((pc.^rho)*ALx/ALh).^(1/(rho-1))).^rho;
tc =[.35    .35     .35     .35     .35]';
gc =[3829.71688669  3829.71688669  3829.71688669  3829.71688669...
3829.71688669]';

% Set sample size for Monte Carlo simulation
% generate repeatable random number with the Mersenne twister generator
s = RandStream('mt19937ar','Seed',1); % MATLAB's start-up settings
RandStream.setDefaultStream(s);

%% FOR MATLAB2013 users:
% rng(1,'twister');
% s=rng;

smpl=1000000;
```

```
eps1=randn(smpl,1);
eps2=randn(smpl,1);

iwrite=0;
tol=1e-8;
maxits=100;
nlfac=1;
dh=1e-4;
iter=1000;
big=1000;
ibig=0;
bignum=.125;
bigind=0;
nladj=1.05;



%% ============Loop==============

xsave=trans(x0);

x=xsave;
count=ones(size(x,1),1);
icnt=0;

while icnt<size(count,1);
icnt=icnt+1;
count(icnt,1)=icnt;
end;

allfun=abs(func(x));
maxfun=max(allfun);
nladj=1.05;

if maxfun<20;
nladj=1.5;
end
dh=.001;
nlfac=.05;


x = nlsrch(x,tol,bignum,maxits,nlfac);

xsave=x;
funcout=func(x);
```

```
x=untrans(x);

pr=[ncom muy sigy mual sigal lam ALh ALx rho gm Theta];

disp('parameters:');
disp({'ncom' 'muy' 'sigy' 'mual' 'sigal' 'lam' 'ALh' 'ALx' 'rho'...
'gm' 'Theta' 'kappa';
ncom muy sigy mual sigal lam ALh ALx rho gm Theta kappa});

disp('result p, g, t are');
p=x(1:ncom,1);
g=x(ncom+1:2*ncom,1);
tax=x(2*ncom+1:3*ncom,1);

disp(num2str([p g tax]))

kv=kcalc(p,g);

zmat=zproc(kv,lyvec);
inc=inproc(zmat);
popn=poproc(zmat);


display('Incomes, Populations, Mean Incomes')
disp(num2str([inc popn inc./popn]))
%% ?;
display('Sum of Community Incomes Compared to Total Metropolitan Income')
disp(num2str([sum(inc) exp(muy+sigy^2/2)]))



hd = (p+(ALx*p./ALh).^(1/(1-rho))).^(-1).*inc;
hsuply=land.*(p./(1+tax)).^Theta*((1-gm)/wq).^Theta;
%%?;
display('Housing Demands and Supplies, housing share of income,...
Land Areas')
disp(num2str([hd,hsuply,p.*hd./inc, land]))


if Montind==1;
[kvrnd,yave,houave]=montc([p;g]);
monres=[kvrnd yave houave];

display('Compare Monte Carlo and quadrature')
```

```
display('Community Sizes')
disp(num2str([monres(:,1) popn]))

display('community mean incomes')
disp(num2str([monres(:,2) (inc./popn)]))

display('mean housing demand')
disp(num2str([monres(:,3),hd./popn]))

[avgcvad,distbetad] = cvcalc([p;g],capqc);
cvdist=[avgcvad,distbetad];

disp('aggregate cv and %better off');
disp(num2str([avgcvad; distbetad]))


rentsa =gm*((1-gm)/wq).^((1-gm)/gm)*sum(land.*(pc./(1+tc)).^(1/gm))';

rentsb = gm*((1-gm)/wq).^((1-gm)/gm)*sum(land.*(p./(1+tax)).^(1/gm))';

display(' Rents a, b ')
disp(num2str( [rentsa,rentsb]))
Rentsch= rentsb - rentsa;

display(' Change in Rents=')
disp(num2str(Rentsch))

display('aggregate cv + change in rents');
disp(num2str(cvdist(1,1)-Rentsch))
pot_gain(iii)=-((cvdist(1,1)-Rentsch));
tax_vec(:,iii)=tax;
perc_better(iii)=distbetad;
popn_vec(:,iii)=popn;
meaninc_vec(:,iii)=inc./popn;

end

%% Plot kappa vs.potential gain
figure;
h0=plot(kappa_vec,pot_gain);
set(h0,'linewidth',2);
ylim([-600 0]);
xlim([0 0.9]);
title('Spillover vs. Potential Gain, A Property Tax Case (US)');
xlabel('\kappa');
```

```
ylabel('Potential gain (US$)');


%% Plot kappa vs. property tax
figure;
tax_vec_conv=0.07*tax_vec;
h=plot(kappa_vec,tax_vec_conv);
set(h,'linewidth',2);
title('Spillover vs. Property tax rate (US)');
xlabel('\kappa');
ylabel('Property tax');
legend('Central City','Suburban1','Suburban2','Suburban3','Suburban4',
'Location', 'SouthWest');



%% Plot kappa vs. distbetad
better=100*perc_better;
figure;
h2=plot(kappa_vec,better);
set(h2,'linewidth',2);
ylim([-100 100]);
xlim([0 0.9]);
title('Spillover vs. Better-off, a property tax case (US)');
xlabel('\kappa');
ylabel('Better-off population (%)');


%% Plot kappa vs. Population
figure;
h3=plot(kappa_vec,popn_vec);
set(h3,'linewidth',2);
title('Spillover vs. Population, a property tax case (US)');
xlabel('\kappa');
ylabel('Population (normalized to 1)')
legend('Central City','Suburban1','Suburban2','Suburban3','Suburban4',
'Location','Northeast');



%% Plot kappa vs. Mean Incomes
figure;
h4=plot(kappa_vec,meaninc_vec);
set(h4,'linewidth',2);
title('Spillover vs. Mean incomes, a property tax case (US)');
```

```
xlabel('\kappa');
ylabel('Mean Incomes');
legend('Central City','Suburban1','Suburban2','Suburban3','Suburban4',
'Location','SouthWest');

%% this is the end of the main script for property tax
```

```
%% kcalc.m

function [kint] = kcalc(pvect,gvec)

global  kappa ALx ALh rho kmax;
capq=ALx./(1+(pvect.^rho*ALx/ALh).^(1/(rho-1))).^rho+...
(ALh*(pvect*ALx/ALh).^(rho/(rho-1)))./(1+(pvect.^rho*ALx/ALh).^(1/...
(rho-1))).^rho;

kint=log((capq(2:5,1)-capq(1:4,1))./(gvec(1:4,1).^rho-gvec(2:5,1).^rho+...
kappa*(gvec(2:5,1).^rho-gvec(1:4,1).^rho)));
kint=[kint;kmax];

kint(imag(kint)~=0)= 10;
kint(isnan(kint)) = 10;
kint(isinf(kint)) = 10;
end
%% this is the end of kcalc.m
```

```
%% ktil.m

function [ktilde] = ktil(p,g,dpdgv)

global rho ALx ALh
ktilde=-rho*log(p+(p*(ALx/ALh)).^(1/(1-rho)))+log(ALh)-log(p)-...
(rho-1)*log(g)+log(dpdgv);
ktilde(imag(ktilde)~=0)= 10;
ktilde(isnan(ktilde)) = 10;
ktilde(isinf(ktilde)) = 10;
end
%% this is the end of ktil.m
```

```
%% zproc.m

function [zv] = zproc(kv,lyv)

global muy sigy mual sigal lam rho ;

omg1=rho./(sigal*(1-lam^2)^.5);
omg2=-lam./(sigy*(1-lam^2)^.5);
omgv=(kv-mual+lam*sigal*muy/sigy)./(sigal*(1-lam^2)^.5);
zv=repmat(omgv,1,48)+(omg1+omg2)*repmat(lyv,5,1);
%% zv is 5x48. it's zmat in the main script
end
%% this is the end of zproc.m
```

```
%% voproc.m
function [ vlow ] = voproc(zmat,ztmat)

global ncom diff lyvec muy sigy wvec;

zlolim=-5*ones(1,size(zmat,2));
zlobnd=[zlolim;zmat(1:(ncom-1),:)];

vlow=((diff/2).*(repmat((normpdf(((lyvec-muy)/sigy))),5,1).*...
(normcdf(ztmat)-normcdf(zlobnd))*wvec));

end

%% this is the end of voproc.m
```

```
%% poproc.m

function [popn] = poproc(zmat)

global ncom muy sigy lyvec diff wvec;
% zmat is 5x48, where 5=ncom

cumpop=(diff/2).*(repmat(normpdf((lyvec-repmat(muy,1,48))/...
sigy),ncom,1)).*normcdf(zmat)*wvec;

cumpop=[0;cumpop];

popn=(cumpop(2:6,1)-cumpop(1:5,1));

end

%% this is the end of poproc.m



%% inproc.m

function[incnu]=inproc(zmat)

global ncom muy sigy lyvec diff wvec;

cincnu=((diff/2).*repmat(exp(lyvec).*normpdf((lyvec-repmat(muy,1,48))/...
sigy),ncom,1).*normcdf(zmat)*wvec);
cincnu=[0;cincnu];
incnu=(cincnu(2:6,1)-cincnu(1:5,1));

end

%% this is the end of inproc.m


%% myopic.m

function [myop] = myopic(hd,popn)
% This proc calculates the slope of the GPF when voters are myopic

myop=popn./hd;
end

%% this is the end of myopic.m
```

```
%% cvcalc.m
function [avgcv,distbet,avgcvad,distbetad] = cvcalc(x,capqc)

global kappa smpl ncom ALx ALh rho muy sigy lam eps1 eps2 mual sigal gc;

p=x(1:ncom,1);
g=x(ncom+1:2*ncom,1);

capq=ALx./(1+((p.^rho)*ALx/ALh).^(1./(rho-1))).^rho +...
(ALh*(p*ALx/ALh).^(rho/(rho-1)))./(1+((p.^rho)*ALx/ALh).^(1/(rho-1))).^rho;
yrnd=exp(muy+sigy*((1-lam^2).^(.5))*eps1+lam*sigy*eps2);

alprnd=exp(mual+sigal*eps2);

utlc = (repmat(alprnd',5,1).*repmat(gc,1,1000000).^rho-kappa*...
repmat(alprnd',5,1).*repmat(gc,1,1000000).^rho+repmat(capqc,1,1000000).*...
repmat(yrnd',5,1).^rho).^(1/rho);

utl=(repmat(alprnd',5,1).*repmat(g,1,1000000).^rho-kappa*...
repmat(alprnd',5,1).*repmat(g,1,1000000).^rho+repmat(capq,1,1000000).*...
(repmat(yrnd,1,5))'.^rho).^(1/rho);

cvall = ((utlc.^rho - repmat(alprnd',5,1).*repmat(g,1,1000000).^rho...
-kappa*repmat(alprnd',5,1).*repmat(g,1,1000000).^rho)./...
repmat(capq,1,1000000)).^(1/rho)-(repmat(yrnd,1,5))';

ordstat1=max(utl)';

cv=zeros(smpl,1);

icmax =smpl;
icnt = 0;

while icnt < icmax;
icnt = icnt + 1;
jvar=0;

while jvar<5;
jvar=jvar+1;

    if utl(jvar,icnt) == ordstat1(icnt,1);

    cv(icnt,1) = ((utlc(jvar,icnt).^rho -...
     alprnd(icnt,1)*g(jvar,1).^rho)/capq(jvar,1)).^(1/rho)-yrnd(icnt,1);
    end
```

```
end;
end;

avgcv= sum(cv)'./smpl;

dist=logical(cv<0);
distbet = sum(dist)'/smpl;

ordstat2=min(cvall,[],1)';
%% Return the minimum of matrix cvall from each column

avgcvad= sum(ordstat2)'/smpl;

distadad = logical(ordstat2<=0);

distbetad = sum(distadad)'/smpl;

end

%% this is the end of cvcalc.m
```

```
%% montc.m

function [kvrnd,yave,houave] = montc(x,capqc)

global kappa smpl ncom ALx ALh rho muy sigy lam eps1 eps2 mual sigal gc;

p=x(1:ncom,1);
g=x(ncom+1:2*ncom,1);

capq=ALx./(1+((p.^rho)*ALx/ALh).^(1/(rho-1))).^rho +...
(ALh*(p*ALx/ALh).^(rho/(rho-1)))./(1+((p.^rho)*ALx/ALh).^(1/(rho-1))).^rho;
yrnd=exp(muy+sigy*((1-lam^2).^.5)*eps1+lam*sigy*eps2);

alprnd=exp(mual+sigal*eps2);

utl=-(repmat(alprnd',5,1).*repmat(g,1,1000000).^rho-kappa*...
repmat(alprnd',5,1).*repmat(g,1,1000000).^rho+repmat(capq,1,1000000).*...
(repmat(yrnd,1,5))'.^rho);

ordstat1=max(utl)';

e1=logical(repmat(ordstat1,1,5)==utl');

kvrnd=sum(e1)'/smpl;
yave=(sum(e1.*repmat(yrnd,1,5))'/smpl)./kvrnd;

%% this is the end of montc.m
```

We use the same program and procedures for Indonesian case with different parameters and initial guess for head-tax case. To calibrate the Indonesian property tax case we use the results from the head tax case and calibrate the housing prices to get a condition where the housing market is in an equilibrium.

```
%PARAMETER FOR INDONESIAN CASE

muy= 9.861407405811761; %Statistic Indonesia, Jakarta 2011
sigy= 0.8310;
```

```
%% Initial guess:
x0=[10 10 10 10 10 952 2403 3402 5005 9032 952 2403 3402 5005 9032]';

%%Centralized property-tax equilibrium values%%
pc =  [13.1321740284 13.1321740284 13.1321740284 13.1321740284...
13.1321740284]';
tc=[0.1 0.1 0.1 0.1 0.1]';
gc =[2129.71688669  2129.71688669  2129.71688669  2129.71688669...
2129.71688669]';
```

## A.4 Comparison between Monte Carlo and Quadrature Results

### A.4.1 Efficient Tax

Table A.1: Comparison: Community size

| Jurisdiction | Monte Carlo | Quadrature |
|---|---|---|
| 1 | 0.54 | 0.53 |
| 2 | 0.12 | 0.16 |
| 3 | 0.16 | 0.13 |
| 4 | 0.13 | 0.11 |
| 5 | 0.06 | 0.07 |

Table A.2: Comparison: Mean incomes

| Jurisdiction | Monte Carlo | Quadrature |
|---|---|---|
| 1 | 21836.59 | 21514.20 |
| 2 | 46122.16 | 47369.65 |
| 3 | 65445.85 | 69464.26 |
| 4 | 107345.43 | 103512.53 |
| 5 | 232143.27 | 203572.87 |

Table A.3: Comparison: Mean housing demand

| Jurisdiction | Monte Carlo | Quadrature |
|---|---|---|
| 1 | 482.07934 | 474.46896 |
| 2 | 875.58096 | 900.97789 |
| 3 | 1190.347 | 1268.9082 |
| 4 | 1879.7208 | 1808.0607 |
| 5 | 3791.4179 | 3296.2473 |

### A.4.2 Property Tax

Table A.4: Comparison: Community size

| Jurisdiction | Monte Carlo | Quadrature |
|---|---|---|
| 1 | 0.42 | 0.39 |
| 2 | 0.15 | 0.15 |
| 3 | 0.15 | 0.15 |
| 4 | 0.14 | 0.15 |
| 5 | 0.14 | 0.16 |

Table A.5: Comparison: Mean incomes

| Jurisdiction | Monte Carlo | Quadrature |
|---|---|---|
| 1 | 27873.14 | 26954.37 |
| 2 | 45695.17 | 43692.87 |
| 3 | 57352.55 | 54753.88 |
| 4 | 74584.73 | 71008.76 |
| 5 | 122386.04 | 113309.55 |

Table A.6: Comparison: Housing demand

| Jurisdiction | Monte Carlo | Quadrature |
|---|---|---|
| 1 | 524.60 | 507.30 |
| 2 | 759.64 | 726.35 |
| 3 | 899.58 | 858.82 |
| 4 | 1093.76 | 1041.32 |
| 5 | 1589.32 | 1471.45 |

### A.5 The household's maximization problem

$$\max_{x_a^i, x_b^i, l_i} \pi_u^i[U_u^i(x_a^i, x_b^i, l_i)] + \pi_d^i[U_d^i(x_a^i, x_b^i, l_i)]$$

subject to

$$q_a x_a^i + q_b x_b^i = w_i l_i (1 - t_i)\sigma_i + w_i l_i (1 - \sigma_i) - T_i$$

Where $q_a = p_a(1 + \tau_a)$ then we can rewrite the constraint as:

$$p_a(1 + \tau_a) + q_b x_b^i = w_i l_i (1 - \sigma_i t_i) - T_i$$

Assuming:

$$U_u^i > U_d^i$$

$$U_u^i - U_d^i = G(1 - \sigma_i) w_i l_i (1 - \sigma_i)$$

$$\tau_b = 0$$

$$\begin{aligned}
\Rightarrow \pi_u^i[U_u^i(x_a, x_b, l_i)] + \pi_d^i[U_d^i(x_a, x_b, l_i)] &= (1 - \pi_d^i)[U_u^i(x_a, x_b, l_i)] + \pi_d^i[U_d^i(x_a, x_b, l_i)] \\
&= U_u^i(.) - \pi_d^i[U_u^i(.) - U_d^i(.)] \\
&= U_u^i(.) - \pi_d^i G(1 - \sigma_i) w_i l_i [1 - \sigma_i]
\end{aligned}$$

Define $g(1-\sigma_i) \equiv G(1-\sigma_i)(1-\sigma_i)$ as cost of being detected per unit income, we have:

$$\pi_u^i[U_u^i(.)] + \pi_d^i[U_d^i(.)] = U_u^i(.) - \pi_d^i g(1-\sigma_i)w_i l_i$$

The maximization problem:

$$\max_{x_a^i, x_b^i, l_i} U_u^i(.) - \pi_d^i g(1-\sigma_i)w_i l_i \text{ s.t. } p_a(1+\tau_a)x_a^i + q_b x_b^i = w_i l_i(1-t_i\sigma_i) - T_i$$

It gives us the Lagrangian function as follows:

$$\mathcal{L} = U_u^i(x_a, x_b, l_i) - \pi_d^i g(1-\sigma_i)w_i l_i + \alpha[w_i l_i(1-t_i\sigma_i) - T_i - (1+\tau_a)x_a^i - q_b x_b^i]$$

By Roy's identity we obtain:

$$\begin{aligned}
V_{\tau_a}^i &= \frac{dU}{d\tau_a} = \frac{\partial \mathcal{L}}{\partial \tau_a}\bigg|_* = -\alpha_i x_a^i \\
V_{t_i}^i &= \frac{dU}{dt_i} = \frac{\partial \mathcal{L}}{\partial t_i}\bigg|_* = -\alpha_i w_i l_i \sigma_i \\
V_{T_i}^i &= \frac{dU}{dT_i} = \frac{\partial \mathcal{L}}{\partial T_i}\bigg|_* = -\alpha_i \\
V_{\sigma_i}^i &= \frac{dU}{d\sigma_i} = \frac{\partial \mathcal{L}}{\partial \sigma_i}\bigg|_* = -w_i l_i(\pi_d^i g'(1-\sigma_i) + \alpha_i t_i)
\end{aligned}$$

where $\alpha_i$ is the marginal utility of income, both $p_a$ and $p_b(= q_b)$ are numéraire.

### A.6 The mimicker's maximization problem

Defining:

$$\bar{l_2} \equiv \frac{w_1 l_1}{w_2}$$

$$\bar{\sigma_2} \equiv \frac{\sigma_i w_i l_i}{w_2 \bar{l_2}}$$

$$\bar{x}_a^2(\tau_a, t_1, T_1; \bar{l_2}, \bar{\sigma_2})$$

$$\bar{x}_b^2(\tau_a, t_1, T_1; \bar{l_2}, \bar{\sigma_2})$$

$$\bar{V}^2(\tau_a, t_1, T_1) \equiv U^2(\bar{x}_a^2(.), \bar{x}_b^2(.), \frac{w_1 l_1(.)}{w_2}; \bar{\sigma}_2)$$

The maximization problem for a mimicker:

$$\max_{\bar{x}_a^2, \bar{x}_b^2, \bar{l}_2} U(\bar{x}_a^2, \bar{x}_b^2, \bar{l}_2) - \pi_d^2 g(1-\bar{\sigma}_2) w_2 \bar{l}_2 \text{ s.t. } w_2 \bar{l}_2 (1 - t_1 \bar{\sigma}_2) - T_1 - \bar{x}_a^2 (1+\tau_a) - \bar{x}_b^2$$

which gives a Lagrangian function:

$$\mathcal{L} = U(\bar{x}_a^2, \bar{x}_b^2, \frac{w_1 l_1}{w_2}) - \pi_d^2 g(1-\bar{\sigma}_2) w_2 \bar{l}_2 + \bar{\alpha}_2 [w_2 \bar{l}_2 (1 - t_1 \bar{\sigma}_2) - T_1 - \bar{x}_a^2 (1+\tau_a) - \bar{x}_b^2]$$

$$\mathcal{L} = U(.) - \pi_d^2 g(1-\bar{\sigma}_2) w_2 \frac{w_1 l_1(.)}{w_2} + \bar{\alpha}_2 [w_2 \frac{w_1 l_1(.)}{w_2} (1 - t_1 \bar{\sigma}_2) - T_1 - \bar{x}_a^2 (1+\tau_a) - \bar{x}_b^2]$$

$$\mathcal{L} = U(.) - \pi_d^2 g(1-\bar{\sigma}_2) w_1 l_1(.) + \bar{\alpha}_2 [w_1 l_1(.)(1 - t_1 \bar{\sigma}_2) - T_1 - \bar{x}_a^2 (1+\tau_a) - \bar{x}_b^2]$$

By the envelope theorem:

$$
\begin{aligned}
\bar{V}^2_{\tau_a} &= \frac{\partial \mathcal{L}}{\partial \tau_a}\bigg|_* = \frac{U_{\bar{l}_2}}{w_2}\frac{\partial w_1 l_1}{\partial \tau_a} - \pi_d^2 g(1-\bar{\sigma}_2)\frac{\partial w_1 l_1}{\partial \tau_a} + \bar{\alpha}_2\bigg[(1-t_1\bar{\sigma}_2)\frac{\partial w_1 l_1}{\partial \tau_a} - \bar{x}_a^2\bigg] \\
\bar{V}^2_{t_1} &= \frac{\partial \mathcal{L}}{\partial t_1}\bigg|_* = \frac{U_{\bar{l}_2}}{w_2}\frac{\partial w_1 l_1}{\partial t_1} - \pi_d^2 g(1-\bar{\sigma}_2)\frac{\partial w_1 l_1}{\partial t_1} + \bar{\alpha}_2\bigg[-w_1 l_1 \bar{\sigma}_2 + (1-t_1\bar{\sigma}_2)\frac{\partial w_1 l_1}{\partial t_1}\bigg] \\
\bar{V}^2_{T_1} &= \frac{\partial \mathcal{L}}{\partial T_1}\bigg|_* = \frac{U_{\bar{l}_2}}{w_2}\frac{\partial w_1 l_1}{\partial T_1} - \pi_d^2 g(1-\bar{\sigma}_2)\frac{\partial w_1 l_1}{\partial T_1} + \bar{\alpha}_2\bigg[(1-t_1\bar{\sigma}_2)\frac{\partial w_1 l_1}{\partial T_1} - 1\bigg] \\
\bar{V}^2_{\bar{\sigma}_2} &= \frac{\partial \mathcal{L}}{\partial \bar{\sigma}_2}\bigg|_* = -(\pi_d^2 g'(1-\bar{\sigma}_2)w_1 l_1 + \bar{\alpha}_2 w_1 l_1 t_1)
\end{aligned}
$$

Please note that we also assume $g(0) = g'(0) = 0$.

We can then rewrite the equations into:

$$
\begin{aligned}
\bar{V}^2_{\tau_a} &= -\bar{\alpha}_2 \bar{x}_a^2 + \bar{\alpha}^2 \frac{\partial w_1 l_1}{\partial \tau_a}\xi \\
\bar{V}^2_{t_1} &= -\bar{\alpha}_2 w_1 l_1 \bar{\sigma}_2 + \bar{\alpha}_2 \frac{\partial w_1 l_1}{\partial t_1}\xi \\
\bar{V}^2_{T_1} &= -\bar{\alpha}_2 + \bar{\alpha}^2 \frac{\partial w_1 l_1}{\partial T_1}\xi \\
\bar{V}^2_{\bar{\tau}_2} &= -w_1 l_1(\bar{\alpha}_2 t_1 + \pi_d^2 g'(1-\bar{\tau}_2))
\end{aligned}
$$

where

$$
\xi \equiv \frac{U_{\bar{l}_2}}{w_2 \bar{\alpha}_2} + (1 - t_1\bar{\sigma}_2) - \frac{\pi_d^2 g(1-\bar{\sigma}_2)}{\bar{\alpha}_2}
$$

### A.7 The planner's problem

Before working on the planner's problem we will derive equations of compensated labor supply and demand for good $a$ (denoted by a tilde). Defining household $i$ tax payment:

$$\bar{R}_i = \tau_a x_a^i + t_1 \sigma_i w_i l_i + T_i$$
$$\Rightarrow w_i l_i = (\bar{R}_i - \tau_a x_a^i - T_i)\frac{1}{t_i \sigma_i}$$
$$\Rightarrow \frac{\partial w_i l_i}{\partial \bar{R}_i} = -\frac{\partial w_i l_i}{\partial T_i} = \frac{1}{t_i \sigma_i}$$

Defining duality and adopting Cook (1972):

$$w_i \tilde{l}_i(t_i, U^i) \equiv w_i l_i(t_i, \bar{R}_i, (t_i, U^i))$$
$$\frac{\partial w_i \tilde{l}_i}{\partial t_i} = \frac{\partial w_i l_i}{\partial t_i} + \frac{\partial w_i l_i}{\partial \bar{R}_i}.\frac{\partial \bar{R}_i}{\partial t_i}$$
$$\frac{\partial w_i \tilde{l}_i}{\partial t_i} = \frac{\partial w_i l_i}{\partial t_i} - \frac{\partial w_i l_i}{\partial T_i}.\sigma_i w_i l_i$$

Analogously, utilizing the same process for compensated demand for good $a$, where the duality is $\tilde{x}_a^i(t_i, U^i) \equiv x_a^i(t_i, \bar{R}_i(t_i, U^i))$, we can obtain:

$$\frac{\partial \tilde{x}_a^i}{\partial t_i} = \frac{\partial x_a^i}{\partial t_i} - w_i l_i \sigma_i \frac{\partial w_i l_i}{\partial T_i}$$

Now we turn to the maximization problem.

The planner faces a maximization problem as follows:

$$\max_{\tau_a, t_i, T_i} V^2(\tau_a, t_2, T_2)$$

subject to:

$$\begin{aligned}
&MUC: \quad && V^1(\tau_a, t_1, T_1) \geq U_s^1 \\
&ICH: \quad && V^2(\tau_a, t_2, T_2) \geq \bar{V}^2(\tau_a, t_1, T_1; \frac{w_1 l_1(.)}{w_2}) \\
&GBC: \quad && n_1[\tau_a x_a^1(.) + \sigma_1 t_1 w_1 l_1(.) + T_1] + n_2[\tau_a x_a^2(.) + \tau_2 t_2 w_2 l_2(.) + T_2] \geq R_0
\end{aligned}$$

where $x_a^1(\tau_a, t_1, T_1)$, $x_a^2(\tau_a, t_2, T_2)$, $l_1(\tau_a, t_1, T_1)$, and $l_2(\tau_a, t_2, T_2)$.

$\mu, \lambda > 0$ since we have monotonic preferences, and $\gamma > 0$ if $U_s^1$ is sufficiently high. It is also assumed that we have a normal case (see Stiglitz, 1982), i.e. the low type will never mimic high type.

The Lagrangian function:

$$\begin{aligned}
\mathcal{L} = V^2(\tau_a, t_2, T_2) + \mu\Big[U_s^1 - V^1(\tau_a, t_1, T_1)\Big] + \gamma\Big[\bar{V}^2(\tau_a, t_1, T_1; \frac{w_1 l_1(.)}{w_2}) \\
-V^2(\tau_a, t_2, T_2)\Big] + \lambda\Big[R_0 - n_1[\tau_a x_a^1(\tau_a, t_1, T_1) + \sigma_1 t_1 w_1 l_1(\tau_a, t_1, T_1) + T_1] - n_2[\tau_a x_a^2(\tau_a, t_2, T_2) \\
+\tau_2 t_2 w_2 l_2(\tau_a, t_2, T_2) + T_2]\Big]
\end{aligned}$$

First-order conditions:

$$
\begin{aligned}
\frac{\partial \mathcal{L}}{\partial \tau_a} =& V^2_{\tau_a} - \mu V^1_{\tau_a} + \gamma \bar{V}^2_{\tau_a} - \gamma V^2_{\tau_a} + \lambda \Bigg[ - n_1 \Bigg( \tau_a \frac{\partial x^1_a}{\partial \tau_a} + x^1_a + t_1 \sigma_1 \frac{\partial w_1 l_1}{\partial \tau_a} \Bigg) \\
& - n_2 \Bigg( \tau_a \frac{\partial x^2_a}{\partial \tau_a} + x^2_a + t_2 \sigma_2 \frac{\partial w_2 l_2}{\partial \tau_a} \Bigg) \Bigg] = 0 \\
\frac{\partial \mathcal{L}}{\partial t_1} =& - \mu V^1_{t_1} + \gamma \bar{V}^2_{t_1} + \lambda \Bigg[ - n_1 \Bigg( \tau_a \frac{\partial x^1_a}{\partial t_1} + t_1 \sigma_1 \frac{\partial w_1 l_1}{\partial t_1} + \sigma_1 w_1 l_1 \Bigg) \Bigg] = 0 \\
\frac{\partial \mathcal{L}}{\partial t_2} =& V^2_{t_2} - \gamma V^2_{t_2} + \lambda \Bigg[ - n_2 \Bigg( \tau_a \frac{\partial x^2_a}{\partial t_2} + t_2 \sigma_2 \frac{\partial w_2 l_2}{\partial t_2} + \sigma_2 w_2 l_2 \Bigg) \Bigg] = 0 \\
\frac{\partial \mathcal{L}}{\partial T_1} =& - \mu V^1_{T_1} + \gamma \bar{V}^2_{T_1} + \lambda \Bigg[ - n_1 \Bigg( \tau_a \frac{\partial x^1_a}{\partial T_1} + t_1 \sigma_1 \frac{\partial w_1 l_1}{\partial T_1} + 1 \Bigg) \Bigg] = 0 \\
\frac{\partial \mathcal{L}}{\partial T_2} =& V^2_{T_2} - \gamma V^2_{T_2} + \lambda \Bigg[ - n_2 \Bigg( \tau_a \frac{\partial x^2_a}{\partial T_2} + t_2 \sigma_2 \frac{\partial w_2 l_2}{\partial T_2} + 1 \Bigg) \Bigg] = 0
\end{aligned}
$$

Using some manipulations:

$$
\begin{aligned}
&\text{(i) foc}(t_1) - w_1 l_1 \sigma_1 \text{foc}(T_1) = 0 \\
&- \mu V_{t_1}^1 + \gamma \bar{V}_{t_1}^2 - \lambda n_1 \left( \tau_a \frac{\partial x_a^1}{\partial t_1} + t_1 \sigma_1 \frac{\partial w_1 l_1}{\partial t_1} + \sigma_1 w_1 l_1 \right) \\
&- w_1 l_1 \sigma_1 \left[ - \mu V_{T_1}^1 + \gamma \bar{V}_{T_1}^2 - \lambda n_1 \left( \tau_a \frac{\partial x_a^1}{\partial T_1} + t_1 \sigma_1 \frac{\partial w_1 l_1}{\partial T_1} + 1 \right) \right] = 0 \\
\Rightarrow &- \mu V_{t_1}^1 + \gamma \bar{V}_{t_1}^2 - \lambda n_1 \tau_a \frac{\partial x_a^1}{\partial t_1} - \lambda n_1 t_1 \sigma_1 \frac{\partial w_1 l_1}{\partial t_1} - \cancel{\lambda n_1 \sigma_1 w_1 l_1} \\
&+ w_1 l_1 \sigma_1 \mu V_{T_1}^1 - w_1 l_1 \sigma_1 \gamma \bar{V}_{T_1}^2 + w_1 l_1 \sigma_1 \lambda n_1 \tau_a \frac{\partial x_a^1}{\partial T_1} + w_1 l_1 \sigma_1 \lambda n_1 t_1 \sigma_1 \frac{\partial w_1 l_1}{\partial T_1} \\
&+ \cancel{w_1 l_1 \sigma_1 \lambda n_1} = 0 \\
\Rightarrow & t_1 \left( \lambda n_1 \sigma_1 \left[ \frac{\partial w_1 l_1}{\partial t_1} - \sigma_1 w_1 l_1 \frac{\partial w_1 l_1}{\partial T_1} \right] \right) = - \mu V_{t_1}^1 + \gamma \bar{V}_{t_1}^2 - \lambda n_1 \tau_a \frac{\partial x_a^1}{\partial t_1} + w_1 l_1 \sigma_1 \mu V_{T_1}^1 \\
&- w_1 l_1 \sigma_1 \gamma \bar{V}_{T_1}^2 + w_1 l_1 \sigma_1 \lambda n_1 \tau_a \frac{\partial x_a^1}{\partial T_1} \\
\Rightarrow & t_1 \left( \lambda n_1 \sigma_1 \frac{\partial w_1 \tilde{l}_1}{\partial t_1} \right) = - \mu ( V_{t_1}^1 - w_1 l_1 \sigma_1 V_{T_1}^1 ) + \gamma \left[ \cancel{- \bar{\alpha}_2 w_1 l_1 \sigma_1} + \bar{\alpha}_2 \frac{\partial w_1 l_1}{\partial t_1} \xi \right] \\
&- \lambda n_1 \tau_a \frac{\partial x_a^1}{\partial t_1} - w_1 l_1 \sigma_1 \gamma \left[ \cancel{- \bar{\alpha}_2} + \bar{\alpha}_2 \frac{\partial w_1 l_1}{\partial T_1} \xi \right] + w_1 l_1 \sigma_1 \lambda n_1 \tau_a \frac{\partial x_a^1}{\partial T_1}
\end{aligned}
$$

Since the first term on the RHS equals to zero by the first-order condition of households' maximization problem, then we obtain:

$$
\begin{aligned}
\Rightarrow t_1 \left( \lambda n_1 \sigma_1 \frac{\partial w_1 \tilde{l}_1}{\partial t_1} \right) &= \gamma \xi \bar{\alpha}_2 \left( \frac{\partial w_1 l_1}{\partial t_1} - w_1 l_1 \sigma_1 \frac{\partial w_1 l_1}{\partial T_1} \right) - \lambda n_1 \tau_a \left( \frac{\partial x_a^1}{t_1} - w_1 l_1 \sigma_1 \frac{\partial w_1 l_1}{\partial T_1} \right) \\
&= \gamma \xi \bar{\alpha}_2 \frac{\partial w_1 \tilde{l}_1}{\partial t_1} - \lambda n_1 \tau_a \frac{\partial \tilde{x}_a^1}{t_1} \\
\Rightarrow t_1 &= \frac{\gamma \bar{\alpha}_2}{\lambda n_1 \sigma_1} \xi - \frac{\tau_a}{\sigma_1} \frac{\partial \tilde{x}_a^1 / \partial t_1}{\partial w_1 \tilde{l}_1 / \partial t_1}
\end{aligned}
$$

$$
\begin{aligned}
&\text{(ii) foc}(t_2) - w_2 l_2 \sigma_2 \text{foc}(T_2) = 0 \\
&V_{t_2}^2 - \gamma V_{t_2}^2 - \lambda n_2 \tau_a \frac{\partial x_a^2}{\partial t_2} - \lambda n_2 t_2 \sigma_2 \frac{\partial w_2 l_2}{\partial t_2} - \lambda n_2 \sigma_2 w_2 l_2 \\
&- w_2 l_2 \sigma_2 \left[ V_{T_2}^2 - \gamma V_{T_2}^2 - \lambda n_2 \tau_a \frac{\partial x_a^2}{\partial T_2} - \lambda n_2 t_2 \sigma_2 \frac{\partial w_2 l_2}{\partial T_2} - \lambda n_2 \right] = 0 \\
\Rightarrow &V_{t_2}^2 - \gamma V_{t_2}^2 - \lambda n_2 \tau_a \frac{\partial x_a^2}{\partial t_2} - \lambda n_2 t_2 \sigma_2 \frac{\partial w_2 l_2}{\partial t_2} - \cancel{\lambda n_2 \sigma_2 w_2 l_2} \\
&- w_2 l_2 \sigma_2 V_{T_2}^2 + w_2 l_2 \sigma_2 \gamma V_{T_2}^2 + w_2 l_2 \sigma_2 \lambda n_2 \tau_a \frac{\partial x_a^2}{\partial T_2} + w_2 l_2 \sigma_2 \lambda n_2 t_2 \sigma_2 \frac{\partial w_2 l_2}{\partial T_2} \\
&+ \cancel{w_2 l_2 \sigma_2 \lambda n_2} = 0 \\
\Rightarrow &(1-\gamma)(V_{t_2}^2 - w_2 l_2 \sigma_2 V_{T_2}^2) - \lambda n_2 \tau_a \left( \frac{\partial x_a^2}{\partial t_2} - w_2 l_2 \sigma_2 \frac{\partial w_2 l_2}{\partial T_2} \right) \\
&= t_2 \lambda n_2 \sigma_2 \left( \frac{\partial w_2 l_2}{\partial t_2} - w_2 l_2 \sigma_2 \frac{\partial w_2 l_2}{\partial T_2} \right)
\end{aligned}
$$

Again, the first term in the second parentheses equals to zero by the previous first-order conditions of households' maximization problem.

$$
\begin{aligned}
\Rightarrow &t_2 \lambda n_2 \sigma_2 \frac{\partial w_2 \tilde{l}_2}{\partial t_2} = -\lambda n_2 \tau_a \frac{\partial \tilde{x}_a^2}{\partial t_2} \\
\Rightarrow &t_2 = \frac{\partial \tilde{x}_a^2 / \partial t_2}{-\partial w_2 \tilde{l}_2 / \partial t_2} . \frac{\tau_a}{\sigma_2}
\end{aligned}
$$

(iii) $\text{foc}(\tau_a) - x_a^1\text{foc}(T_1) - x_a^2\text{foc}(T_2) = 0$

$$V_{\tau_a}^2 - \mu V_{\tau_a}^1 + \gamma \bar{V}_{\tau_a}^2 - \gamma V_{\tau_a}^2 - \lambda n_1\left[\tau_a \frac{\partial x_a^1}{\partial \tau_a} + x_a^1 + t_1\sigma_1\frac{\partial w_1 l_1}{\partial \tau_a}\right] - \lambda n_2\left[\tau_a\frac{\partial x_a^2}{\partial \tau_a} + x_a^2\right.$$

$$\left. + t_2\sigma_2\frac{\partial w_2 l_2}{\partial \tau_a}\right] + x_a^1\left[\mu V_{T_1}^1 - \gamma\bar{V}_{T_1}^2 + \lambda n_1\tau_a\frac{\partial x_a^1}{\partial T_1} + \lambda_1 n_1 t_1\sigma_1\frac{\partial w_1 l_1}{\partial T_1} + \lambda n_1\right]$$

$$- x_a^2\left[V_{T_2}^2 - \gamma\bar{V}_{T_2}^2 - \lambda n_2\tau_a\frac{\partial x_a^2}{\partial T_2} - \lambda n_2 t_2\sigma_2\frac{\partial w_2 l_2}{\partial T_2} + \lambda n_2\right] = 0$$

$$\Rightarrow \cancel{-\alpha_2 x_a^2} + \cancel{\mu\alpha_1 x_a^1} + \gamma\left[-\bar{\alpha}_2\bar{x}_a^2 + \bar{\alpha}_2\frac{\partial w_1 l_1}{\partial \tau_a}\xi\right] + \cancel{\gamma\alpha_2 x_a^2} - \lambda n_1\tau_a\frac{\partial x_a^1}{\partial \tau_a} - \cancel{\lambda n_1 x_a^1} - \lambda n_1 t_1\sigma_1\frac{\partial w_1 l_1}{\partial \tau_a}$$

$$- \lambda n_2\tau_a\frac{\partial x_a^2}{\partial \tau_a} - \cancel{\lambda n_2 x_a^2} - \lambda n_2 t_2\sigma_2\frac{\partial w_2 l_2}{\partial \tau_a} - \cancel{x_a^1\mu\alpha_1} - x_a^1\gamma\left[-\bar{\alpha}_2 + \bar{\alpha}_2\frac{\partial w_1 l_1}{\partial T_1}\xi\right] + x_a^1\lambda n_1\tau_a\frac{\partial x_a^1}{\partial T_1}$$

$$+ x_a^1\lambda n_1 t_1\sigma_1\frac{\partial w_1 l_1}{\partial T_1} + \cancel{x_a^1\lambda n_1} + \cancel{x_a^2\alpha_2} + x_a^2\gamma[-\bar{\alpha}_2] + x_a^2\lambda n_2\tau_a\frac{\partial x_a^2}{\partial T_2}$$

$$+ x_a^2\lambda n_2 t_2\sigma_2\frac{\partial w_2 l_2}{\partial T_2} + \cancel{x_a^2\lambda n_2} = 0$$

$$\Rightarrow -\gamma\bar{\alpha}_2\bar{x}_a^2 + \gamma\bar{\alpha}_2\frac{\partial w_1 l_1}{\partial \tau_a}\xi - \lambda n_1\tau_a\frac{\partial x_a^1}{\partial \tau_a} - \lambda n_1\sigma_1\frac{\partial w_1 l_1}{\partial \tau_a}\left[\frac{\gamma\bar{\alpha}_2}{\lambda n_1\sigma_1}\xi - \frac{\tau_a}{\sigma_1}\frac{\partial\tilde{x}_a^1/\partial t_1}{\partial w_1\tilde{l}_1/\partial t_1}\right]$$

$$- \lambda n_2\tau_a\frac{\partial x_a^2}{\partial \tau_a} - \lambda n_2\sigma_2\frac{\partial w_2 l_2}{\partial \tau_a}\left[\frac{\partial\tilde{x}_a^2/\partial t_2}{-\partial w_2\tilde{l}_2/\partial t_2}.\frac{\tau_a}{\sigma_2}\right] + x_a^1\gamma\bar{\alpha}_2 - x_a^1\bar{\alpha}_2\frac{\partial w_1 l_1}{\partial T_1}\xi + x_a^1\lambda n_1\tau_a\frac{\partial x_a^1}{\partial T_1}$$

$$+ x_a^1\lambda n_1\sigma_1\frac{\partial w_1 l_1}{\partial T_1}\left[\frac{\gamma\bar{\alpha}_2}{\lambda n_1\sigma_1}\xi - \frac{\tau_a}{\sigma_1}\frac{\partial\tilde{x}_a^1/\partial t_1}{\partial w_1\tilde{l}_1/\partial t_1}\right] + x_a^2\lambda n_2\tau_a\frac{\partial x_a^2}{\partial T_2}$$

$$+ x_a^2\lambda n_2\sigma_2\frac{\partial w_2 l_2}{\partial T_2}\left[\frac{\partial\tilde{x}_a^2/\partial t_2}{-\partial w_2\tilde{l}_2/\partial t_2}.\frac{\tau_a}{\sigma_2}\right] = 0$$

$$\Rightarrow -\gamma\bar{\alpha}_2\bar{x}_a^2 + \gamma\bar{\alpha}_2\frac{\partial w_1 l_1}{\partial \tau_a}\xi - \lambda n_1\tau_a\frac{\partial x_a^1}{\partial \tau_a} + \lambda n_1\cancel{\sigma_1}\frac{\partial w_1 l_1}{\partial \tau_a}\frac{\tau_a}{\cancel{\sigma_1}}\frac{\partial\tilde{x}_a^1/\partial t_1}{\partial w_1\tilde{l}_1/\partial t_1} - \cancel{\lambda n_1}\sigma_1\frac{\partial w_1 l_1}{\partial \tau_a}\frac{\gamma\bar{\alpha}_2}{\cancel{\lambda n_1}\sigma_1}\xi$$

$$- \lambda n_2\tau_a\frac{\partial x_a^2}{\partial \tau_a} + \lambda n_2\cancel{\sigma_2}\frac{\partial w_2 l_2}{\partial \tau_a}\frac{\partial\tilde{x}_a^2/\partial t_2}{\partial w_2\tilde{l}_2/\partial t_2}.\frac{\tau_a}{\cancel{\sigma_2}} + x_a^1\gamma\bar{\alpha}_2 - x_a^1\bar{\alpha}_2\frac{\partial w_1 l_1}{\partial T_1}\xi + x_a^1\lambda n_1\tau_a\frac{\partial x_a^1}{\partial T_1}$$

$$- x_a^1\lambda n_1\cancel{\sigma_1}\frac{\partial w_1 l_1}{\partial T_1}\frac{\tau_a}{\sigma\sigma_1}\frac{\partial\tilde{x}_a^1/\partial t_1}{\partial w_1\tilde{l}_1/\partial t_1} + x_a^1\cancel{\lambda n_1}\sigma_1\frac{\partial w_1 l_1}{\partial T_1}\frac{\gamma\bar{\alpha}_2}{\cancel{\lambda n_1}\sigma_1}\xi$$

$$+ x_a^2\lambda n_2\tau_a\frac{\partial x_a^2}{\partial T_2} - x_a^2\lambda n_2\cancel{\sigma_2}\frac{\partial w_2 l_2}{\partial T_2}\frac{\partial\tilde{x}_a^2/\partial t_2}{\partial w_2\tilde{l}_2/\partial t_2}.\frac{\tau_a}{\cancel{\sigma_2}} = 0$$

$$
\begin{aligned}
&\Rightarrow -\gamma\bar{\alpha}_2\bar{x_a}^2+\gamma\bar{\alpha}_2\cancel{\frac{\partial w_1 l_1}{\partial\tau_a}}\xi-\lambda n_1\tau_a\left[\frac{\partial x_a^1}{\partial\tau_a}-x_a^1\frac{\partial x_a^1}{T_1}-\frac{\partial\tilde{x}_a^1/\partial t_1}{\partial w_1\tilde{l}_1/\partial t_1}\left(\frac{\partial w_1 l_1}{\partial\tau_a}-x_a^1\frac{\partial w_1 l_1}{\partial T_1}\right)\right]\\
&-\cancel{\frac{\partial w_1 l_1}{\partial\tau_a}}\gamma\bar{\alpha}_2\xi-\lambda n_2\tau_a\left[\frac{\partial x_a^2}{\partial\tau_a}-x_a^2\frac{\partial x_a^2}{T_2}-\frac{\partial\tilde{x}_a^2/\partial t_2}{\partial w_2\tilde{l}_2/\partial t_2}\left(\frac{\partial w_2 l_2}{\partial\tau_a}-x_a^2\frac{\partial w_2 l_2}{\partial T_2}\right)\right]+x_a^1\gamma\bar{\alpha}_2\\
&-x_a^1\bar{\alpha}_2\cancel{\frac{\partial w_1 l_1}{\partial T_1}}\xi+x_a^1\bar{\alpha}_2\cancel{\frac{\partial w_1 l_1}{\partial T_1}}\xi=0
\end{aligned}
$$

Following Nava et.al (1996) that:

$$
\begin{aligned}
&\frac{\partial\tilde{x}_a^i}{\partial\tau_a}\equiv\frac{\partial x_a^i}{\partial\tau_a}-x_a^i\frac{\partial x_a^i}{T_i}\\
&\frac{\partial w_i\tilde{l}_i}{\partial\tau_a}\equiv\frac{\partial w_i l_i}{\partial\tau_a}-x_a^i\frac{\partial w_i l_i}{\partial T_i}\\
&\left.\frac{\partial\tilde{x}_a^i}{\partial\tau_a}\right|_{l_i}\equiv\frac{\partial\tilde{x}_a^i}{\partial\tau_a}-\frac{\partial\tilde{x}_a^i/\partial t_i}{\partial w_i\tilde{l}_i/\partial t_i}\frac{\partial w_i\tilde{l}_i}{\partial\tau_a},
\end{aligned}
$$

then we can obtain:

$$
\begin{aligned}
&\Rightarrow -\lambda\tau_a\left[n_1\left.\frac{\partial\tilde{x}_a^1}{\partial\tau_a}\right|_{l_1}+n_2\left.\frac{\partial\tilde{x}_a^2}{\partial\tau_a}\right|_{l_2}\right]=\gamma\bar{\alpha}_2(\bar{x}_a^2-x_a^1)\\
&\Rightarrow \tau_a=\frac{\gamma}{\lambda}\bar{\alpha}_2(x_a^1-\bar{x}_a^2)\left[n_1\left.\frac{\partial\tilde{x}_a^1}{\partial\tau_a}\right|_{l_1}+n_2\left.\frac{\partial\tilde{x}_a^2}{\partial\tau_a}\right|_{l_2}\right]^{-1}
\end{aligned}
$$