The Dissertation Committee for Wenjia Zhang Certifies that this is the approved version of the following dissertation:

Theoretical and Empirical Investigations of Excessive Congestion as a Result of Market and Planning Failures

Committee:

Ming Zhang, Supervisor

Kara M. Kockelman

Michael Oden

Talia McCray

David A. Kendrick

Theoretical and Empirical Investigations of Excessive Congestion as a Result of Market and Planning Failures

by

Wenjia Zhang, B.S., M.S.

Dissertation

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

The University of Texas at Austin

December 2015

Dedication

This work is dedicated to my wife and parents for their love, support and encouragement.

Acknowledgements

The sincerest thanks are owed to a number of people helping me to complete this dissertation. First and foremost, I am grateful to my advisor, Ming Zhang, for giving me enough freedom to conduct this research, helping me throughout its completion, and mentoring and funding my doctoral studies. Special thanks to Kara Kockelman for bringing me back to the field of urban economics and her relentless support and enthusiasm, and the countless hours devoted to discussing, reading, and editing my work. I thank other committee members, Michael Oden, Talia McCray, and David Kendrick, for their valuable comments and suggestions, and Bjorn Sletto for his coordination and encouragement throughout my doctoral study. I also thank my colleagues Kwangyul Choi, Binbin Chen, and Jane Winslow for their supports, sharing, and helping during my doctoral study and living at Austin.

I have developed parts of my dissertation into several articles under review for publication in several journals. A major part of Chapters 3 and 4 in this dissertation is from two articles coauthored with Kara Kockelman and a part of Chapter 6 refers to an article coauthored with Ming Zhang. I specifically thank my coauthors and several reviewers, such as Daniel McMillen, Gilles Duranton, Alex Anas, Jan Brueckner, and Esteban Rossi-Hansberg, for their inspired questions and comments.

This research was funded through the graduate school named/endowed continuing fellowship from the University of Texas at Austin. My doctoral study was supported by other scholarship, including the Snell Sustainable Cities Doctoral Fellowship granted by the Center of Sustainable Development and research assistantship of research projects from the Southwest Region University Transportation Centre.

Without the support of my family, all my efforts would be in vain. Finally, I am eternally indebted to my wife Xiao for her love and patience.

Theoretical and Empirical Investigations of Excessive Congestion as a Result of Market and Planning Failures

Wenjia Zhang, Ph.D.

The University of Texas at Austin, 2015

Supervisor: Ming Zhang

This dissertation research places traffic congestion in a broader context of land use and economic linkages and contends that optimal congestion relief requires both land use and pricing policies. Congestion is considered excessive when the individually desirable (or privately optimal) amount of auto travel exceeds the socially optimal level. Two underlying causes of excessive congestion are discussed in detail here: market failures from congestion and agglomeration externalities and planning failures from exclusionary zoning and low-density zoning practices. This research is among the first to connect the economics of planning failure with excessive congestion.

This research first specifies a spatial general equilibrium framework that reflects congestion delays, agglomeration economies, and planning failures. Simulation findings suggest that anti-congestion policies might erode agglomeration economies, causing a net social loss. Pricing policies need to balance the benefits from congestion reduction with the losses that come from weakening agglomeration tendencies. The congestion diseconomy is only a small share (5%-23%) of the total cost of congestion; policies seeking to produce free-flow speeds may lead to substantial welfare loss.

Simulations demonstrate how that, when planning failures dominate region, even the first-best pricing is not so effective since planning failures are insensitive to (market) pricing signals. Incorporating land use and economic policies are found to be socially optimal when both planning and market failures exist. This research also examines practical policies. Application of a mileage tax or a cordon toll partially reduces excessive congestion and decentralizes jobs. Urban growth boundaries are relatively inefficient and may distort land markets, causing worse congestion. Firm cluster zoning is more effective since allowing for jobs decentralization. Densification policies can alleviate the excessive congestion caused by low-density zoning regulation. Under exclusionary zoning regulations, building an employment subcenter can greatly improve welfare and alleviate congestion. Planning for new subcenters, however, requires a subsidy or incentive to trigger the firms' relocations.

This research also presents an empirical study using the 2006 Household Travel Survey data obtained for Austin, Texas. This study develops a multilevel multinomial logit model to investigate the interaction effects between land use and travel cost variables on travel mode choice. Results suggest that road-pricing policies are more efficient in reducing driving in neighborhoods with better walkability and easier access to activity centers. The impacts of land use patterns on driving are stronger when driving costs rise. These findings suggest that an incorporation of both land use policies and road pricing policies benefits a region's residents more than the either policy on its own.

Table of Contents

List of Tables

List of Figures

CHAPTER 1: INTRODUCTION

PROBLEM STATEMENT

Traffic congestion[1] is one of the most obstinate problems plaguing many cities and regions around the world. Road users experience direct negative effects due to congestion, such as travel delays, extra fuel costs, and personal stress. There is also an external impact on non-road users, such as residents and businesses near congested roads. A recent report by the Texas Transportation Institute (TTI) estimates that in 2011, each U.S. commuter saved 38 hours per year and 19 gallons of gasoline, summing to $820 per commuter annually, when all roads were free of congestion. Congestion also increases environmental and social costs of travel: namely, air pollution, greenhouse gas emissions, and energy dependency. Traffic congestion in 498 urban areas in the United States caused 5.5 billion hours of travel delay, 2.9 billion gallons of excess fuel consumption, and 56 billion pounds of additional CO_2 emissions. These total $120 billion in losses due to congestion, or almost $400 per capita per year (TTI, 2012).

Traffic congestion seems to be an unavoidable consequence of contemporary city life. Intensifying congestion could be evidence of social prosperity and economic growth (Downs, 2005; Taylor, 2002). Cities with empty roads during peak hours may well indicate the presence of an economic recession. A certain degree of congestion is individually or socially desirable. Road users can tolerate travel delay as long as they obtain other benefits from living and working in a congested area (e.g., high wages, shorter commutes, and easy access [OECD, 2007]). Congestion is excessive only when

[1]Traffic congestion is a concept bounded by space and time. Congestion can occur on not only the street and highway network but also the public transit system, railway system, and airport slots. This dissertation primarily focuses on road traffic congestion raised by automobile vehicles in the citywide highway network system. Without specific notations, traffic congestion and congestion used in the dissertation represent road traffic congestion.

its marginal social cost (MSC) of travel exceed the marginal social benefit (MSB) (OECD, 2007; Victorian Competition and Efficiency Commission [VCEC], 2006).

Congestion ties closely to agglomeration. Heavy congestion often takes place in large cities with strong agglomeration of knowledge and production. The spatial concentration of activities and traffic could simultaneously generate negative congestion externalities and positive agglomeration externalities. Congestion will not negatively affect a city's economy if the agglomeration benefits fully compensate for the congestion losses. Many empirical studies have examined the link between congestion and urban economies (e.g., Boarnet, 1997; Graham, 2007; Hymel, 2009; Sweet, 2011; Weisbrod, Vary, & Treyz, 2001). Although their findings remain inconclusive, several studies have demonstrated that the city economy is impaired only when congestion exceeds a threshold level of crowding (e.g., OECD, 2007; Sweet, 2014). This finding implies that *congestion* is not always bad but *excessive congestion*[2] is inefficient. It is thus important to examine policies for reducing excessive congestion, rather than eliminating *all* congestion.

Few existing studies distinguish excessive congestion from congestion (e.g., OECD, 2007). Such studies may lead to an overestimation of the net social costs of congestion (i.e., the congestion diseconomy) and overreaction in terms of anti-congestion policies. To dissect inefficient congestion, this dissertation tackles three important questions: What causes excessive driving demand, how should the term *excessive* be defined and measured, and how should excessive congestion be effectively mitigated? By

[2]As discussed in the following chapters, traffic congestion may be insufficient, rather than excessive, after agglomeration externalities are accounted for. However, it is better to explain insufficient congestion as insufficient agglomeration. Social optimum exactly needs more agglomeration and concentration of production and consumption activities, rather than more traffic on the roads, although congestion is an inevitable consequence of increasing levels of agglomeration. This dissertation emphasizes the issues of excessive, rather than insufficient, congestion. Policies for resolving the problem of insufficient congestion probably need to encourage agglomeration enhancement.

resolving these issues, this study aims to find a much-needed bridge between planning and economics to inform efficient investment decisions and effective policymaking for congestion relief.

The research articulates two underlying causes of excessive congestion: market failures – in the current market setting where road users pay only the personal cost of travel, which is lower than the social costs of trip making (Arnott, 1979; Brueckner, 2000; Kono & Joshi, 2012; Pines & Sadka, 1985), and planning failures – where government planning interventions cause less efficient development (e.g., low-density sprawl) than would occur without that intervention (Cervero, 1996; Levine, 2006). Either failure tends to result in excessive travel and congestion. In an unregulated market, road users who only pay the private costs of using limited road space would generate excess travel demand, leading to more congestion. Some planning interventions, such as density-constrained, single-use, and minimum-lot-size zoning, restrict alternative development that the market desires and increase trip length and automobile dependence. The increased automobile travel demand adds to regional congestion in the long term.

Although much literature has explained congestion under the framework of market failure (see a review by Anas & Lindsey [2011]), less has scrutinized how planning failure shapes excessive congestion in our living metropolitan areas. It is important to recognize that excessive travel demand caused by market failures differ from planning failures. Excessive congestion is the amount of congestion exceeding the socially optimal level, in which additional trips produce more costs than benefits to the community. Excessive congestion is caused by underpaid travel costs in the perspective of market failures and unnecessary regulatory policies in the perspective of planning failures. Many studies have shown the adverse effects of land use regulation, such as low-density zoning's effects on housing affordability, employment, and urban productivity

(Glaeser & Gyourko, 2003; Hsieh & Moretti, 2015; Quigley & Rafael, 2005; Turner, Haughwout, & van der Klaauw, 2014). However, the effects of land use regulations on mobility and accessibility, along with traffic congestion, are less examined and need further investigation.

The intricate causes of excessive congestion make its measurement complex. The analytical framework for excessive congestion estimation needs to internalize not only social costs specific to travel but also social benefits outside the transport system. The values of daily travel are often not from trips themselves but activities performed at the destinations. People benefit from driving to shopping malls for daily consumption and to the workplace for wage income. The evaluation of travelers' benefits thus needs a systematic framework that integrates interactions among transportation, land use, and labor and consumption markets. Also, the framework should enable researchers to tackle tensions between congestion and agglomeration as well as be responsive to land use regulations. Only limited studies have developed such integrated models (e.g., Anas & Liu, 2007; Anas & Xu, 1997; Wheaton, 2004). Most models only recognize either congestion externalities (e.g., Anas & Xu, 1999; Arnott, 1979; Brueckner, 2007; Pines & Sadka, 1985; Solow, 1972; Wheaton, 1998;) or agglomeration externalities (Berliant, Peng, & Wang, 2002; Borck & Wrede, 2009; Fujita & Ogawa, 1982; Lucas & Rossi-Hansberg, 2002; Rossi-Hansberg, 2004).

Therefore, this dissertation builds a spatial general equilibrium framework to analyze interactions among land use, transportation, and agglomeration of business. Relying on computational simulations and empirical studies, I investigate the effectiveness of two types of anti-congestion policies: congestion pricing and land use planning. The major argument is that either congestion pricing or land use planning policies may reduce excessive driving demand and congestion; however, neither pricing nor planning alone can effectively correct both market and planning failures.

The congestion pricing strategy is an economic approach to addressing market failures by applying a toll, equal to the external cost of travel, on all road users who create the cost. Optimal tolls equal the marginal external costs of congestion—that is, the gap between marginal social and private expenses. Such an optimal tolling policy is difficult to implement in practice, despite the fact that it, in theory, can fully correct the market failure from congestion externalities and adjust traffic to the socially optimal level. There are many more practical but second-best policies, such as cordon charges, area-wide pricing, and variable-rate highway tolling. Pioneering examples around the world include Singapore's cordon charge in the early 1970s, Norway's toll rings in the mid-1980s, London's area-based pricing in 2003 (Ieromonachou, 2006; Santos, 2005), and high-occupancy toll (HOT) lanes and expressways in the United States (U.S. GAO, 2012).

Land use planning strategies are regularly embraced by planners who believe that command-and-control regulations, seeking ideal or desired land use patterns, are an effective solution to congestion. Although most economists regard land use planning as a second-best substitute for optimal congestion pricing policies (e.g., Brueckner, 2000; Wheaton, 1998), many planners argue that land use planning serves more as a strategy for correcting planning failures, rather than market failures (Cervero, 1996; Knaap, Talen, Olshansky, & Forrest, 2000; Levine, 2006). These planners recommend alternatives (e.g., mixed-use development, transit-oriented development, and smart growth) to low-density sprawl, the latter of which is notorious for nurturing an auto-oriented lifestyle and auto dependence (Knaap et al., 2000).

Land use and transportation studies have widely discussed the efficiency of both congestion pricing and land use planning as tactics for driving congestion reduction and relief. However, most studies recognize them as independent or substitutable, rather than complementary, policies. Although a few empirical studies have detected potential benefits from complementary land use and pricing policies (e.g., Guo, Agrawal, & Dill, 2011; Lee & Lee, 2013), no theoretical interpretations have been advanced. This research aims to fill the gap and investigate how incorporating land use and economic policies could be more efficient than either congestion pricing alone or land use planning alone policies.

RESEARCH OBJECTIVES AND QUESTIONS

In this dissertation, I propose an integrated approach to understanding how market systems and planning regulations encourage or discourage auto travel and congestion. This study first recognizes the difference between the economist's view of market failure and the city planner's view of planning failure. I then explore the combination of both views in theory and practice. Specifically, I aim to achieve the following objectives:

(1) To examine how agglomeration externalities as a source of market failure, other than congestion externalities, affect traffic congestion (Chapters 2 and 3). This requires developing a new urban economic model with endogenously determined congestion and agglomeration externalities. In this research, I will adopt computational simulations to identify both the socially optimal and excessive levels of congestion, to measure diseconomies of excessive congestion from market failures, and to examine interactions between congestion and agglomeration.

(2) To theorize how *planning failures* cause excessive congestion (Chapters 2 and 3). I aim at bridging economic and planning analytics for congestion

studies by placing planning failures into the economic framework for market failure analysis. The framework is applied to compare the role of planning versus market failures in shaping excessive congestion and social inefficiency and to articulate when land use planning or economic policies are superior and when incorporating land use and economic policies are socially preferred.

(3) To identify and evaluate *socially optimal* policies for reducing the excessive congestion caused by market failures (e.g., from congestion and agglomeration externalities) and planning failures (e.g., from exclusionary zoning and low-density zoning regulations), respectively (Chapter 3). This research compares land use planning- and congestion pricing-alone policies with complementary land use and pricing policies, with focuses on their impacts on congestion relief, land use, and social welfare.

(4) To assess the effectiveness of *practical* (i.e., second-best) congestion pricing (e.g., vehicle miles of travel [VMT] tax and cordon toll) and land use planning policies (e.g., urban growth boundaries [UGBs], firm cluster zoning, residential densification, and building suburban employment centers), rather than unrealistic *optimal* policies, as a strategy to reduce excessive congestion and improve social welfare (Chapters 4 and 5).

(5) To empirically investigate the interaction effects of land use characteristics and travel costs on travel mode choice in the Austin, Texas area and to provide evidence on the benefits of incorporating land use and economic policies for reducing auto travel demand in Austin (Chapter 6).

(6) To foster methodological innovation in studies addressing urban congestion, planning failure, and land use–transportation–economy integration.

For achieving these research goals, this dissertation will investigate the following questions:

(1) How do anti-congestion policies affect agglomeration economies? Alternatively, in which situations could these policies benefit or harm agglomeration economies? These questions can help decision makers to define and estimate excessive congestion and avoid misestimating the efficiency of anti-congestion policies. Although negative congestion externality is the primary source of market failure leading to excessive congestion, positive agglomeration externality could be a source of market failure that causes insufficient crowding and congestion. To identify excessive congestion, we need to investigate the tension between congestion, as an important centrifugal force of urban growth, and agglomeration, as an important centripetal force.

(2) How do planning failures from exclusionary and low-density zoning regulations affect congestion, land use, and social efficiency? When do planning failures play a more important role in determining excessive congestion than market failure? These questions saliently lack sufficient studies, either in theory or empirics. Proponents of congestion pricing strategies, such as many economists, often overlook the roles of planning failures in the excessive congestion occurrence. In contrast, proponents of land use planning strategies, such as many planners, often neglect the salient role of market failure leading to excessive congestion. Although there is reasoning behind both perspectives, little research has been successful in connecting them together. The primary barriers to bridging economic and planning studies are the inconsistent analytical frameworks used by them.

This research thus will develop a framework for simulating mechanisms of planning and market failures.

(3) In the presence of market and planning failures, can an optimal pricing or land use policy alone eliminate all excessive congestion? This question investigates the necessity of a complementary land use and economic policy. We can break this issue down into four sub-questions: (a) If congestion pricing is ignored, what is the effectiveness of land use planning on congestion reduction and welfare improvement; (b) if land use planning is ignored, what is the effectiveness of congestion pricing on congestion reduction and welfare improvement; (c) comparing these two policies, which one will bring more desired outcomes on congestion reduction, land use, and well-being; and (d) what are the effects of combining congestion pricing and land use planning? These questions are important for land use and transportation planning practice.

(4) What practical pricing and land use policies are effective in reducing excessive congestion? Because optimal policies are often idealistic and infeasible in planning practice, this research will examine several practical, second-best policies, with a focus on their influences on congestion reduction, land use, and social efficiency as well as their potential side effects.

(5) Is incorporating land use and economic policies effective to reduce auto-travel demand in empirical studies? Empirical studies have long investigated the anti-congestion impact of either land use or pricing policies. However, few examine the “mutually supportive” effects between them (Guo et al., 2011). Despite the fact that the empirical study does not directly measure

congestion, it can still provide evidence to examine the findings from the theoretical and simulation research used in this dissertation.

DISSERTATION OUTLINE

The dissertation includes seven chapters. This Chapter 1 states the research problems, key objectives, and questions in this dissertation and describes the organization and structure of the dissertation. Chapter 2 reviews the literature on theories related to excessive congestion, market failure, and planning failure and the empirical studies on anti-congestion policies, including land use planning and congestion pricing strategies. Chapter 2 helps to theorize how congestion connects with agglomeration and how market and planning failures shape excessive congestion. It also identifies the gaps between theories and empirics in addressing congestion issues.

Chapter 3 relies on urban economic theories to develop a novel spatial general equilibrium model that accommodates market failure from congestion and agglomeration externalities. This model is extended to internalize planning failures from exclusionary zoning and low-density zoning regulations. This chapter first investigates how congestion and agglomeration externalities cause market failures leading to excessive congestion and identifies the socially optimal pricing policies to reduce excessive congestion from market failure. Next, the chapter examines how planning failures cause excessive congestion and explores the optimal land use remedies for planning failures. Finally, this chapter emphasizes the importance of incorporating land use and economic policies in cities with both planning and market failures.

Based on the modeling and simulation frameworks created in Chapter 3, Chapter 4 focuses on the effectiveness of practical pricing policies: VMT taxes and cordon tolls. Chapter 5 focuses on practical land use policies, including UGBs, firm cluster zoning, and residential densification. Although practical policies are less efficient and second-best or even third-best compared to the optimal policies, they are often much more politically and financially feasible than first-best policies.

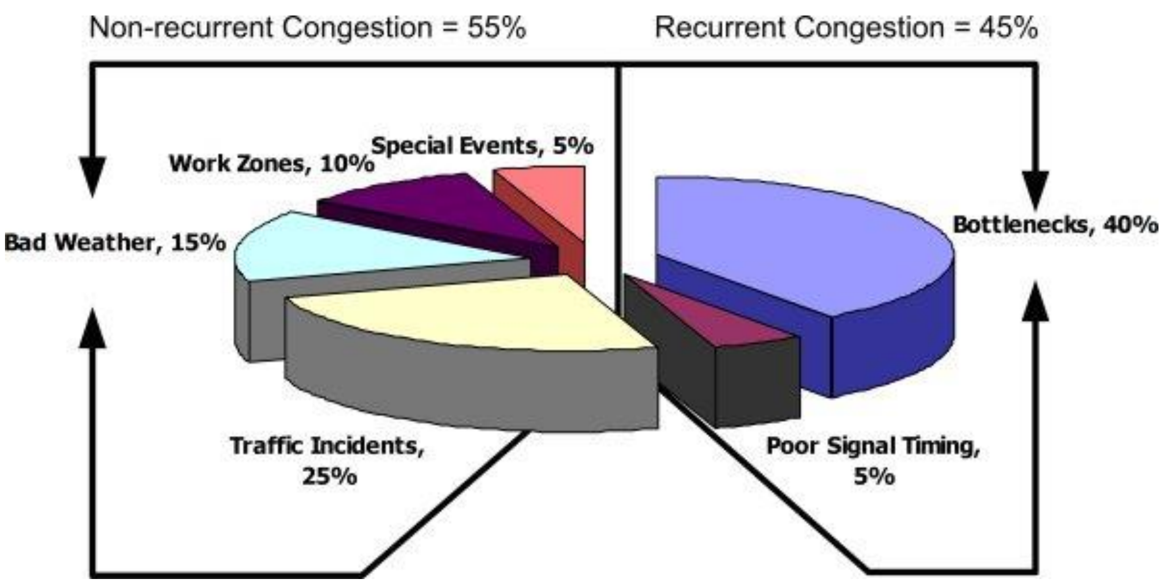
Chapter 6 presents an empirical study of Austin, Texas region area to investigate the interaction effects of land use and travel cost variables on travel mode choice. It develops a multilevel logit model to identify whether neighborhood-level land use characteristics can serve as a spatial context modifying the pricing effect on reducing driving. Despite the fact that this study does not directly model congestion, it can provide evidence of the benefit of incorporating land use and economic policies.

Chapter 7 summarizes key findings of this dissertation research. Primary contributions to literature are outlined, and future work opportunities (theoretical, empirical, and methodological) are discussed.

CHAPTER 2: THEORY OF EXCESSIVE CONGESTION AND EVIDENCE OF ANTI-CONGESTION POLICIES

This chapter contains two sections. The first section disentangles the causes of excessive congestion, articulates theories of market and planning failures, and connects them with traffic congestion. The second section elucidates the underlying reasoning of land use planning and congestion pricing policies for reducing congestion and summarizes the effectiveness of these policies as found in empirical studies.

THEORY OF EXCESSIVE CONGESTION



Sources: http://international.fhwa.dot.gov/pubs/pl07012/images/figure_1.cfm

Figure 2.1 Causes of Congestion in the United States

Congestion arises when travel demand for road space exceeds the available supply of road capacity and can derive from both the supply and demand sides. A report from the U.S. Department of Transportation (DOT) categorizes national highway congestion into nonrecurrent and recurrent congestion and summarizes seven types of causes (Freeway

Management Handbook, 1997). Nonrecurrent congestion results from nonpredictable or one-time-only events, such as traffic incidents, bad weather, maintenance and work zones, and special events. Recurrent congestion comes instead from predictable events, such as signal lights and physical bottlenecks. Nonrecurrent factors generate 55% of highway congestion in the U.S., and recurrent factors account for 45% (see Figure 2.1). Also, socioeconomic dynamics may be significant sources aggravating traffic congestion, including growing population and employment, rising incomes, and decreasing costs of driving (Falcocchio & Levinson, 2015).

Factors causing traffic congestion are not necessarily triggers to excessive congestion. According to the *Oxford English Dictionary* (2015), excessive is defined as "more than is necessary, normal, or desirable." The definition of the "desirable" demand on auto travel is the key point for excessive congestion estimation. Excessive congestion could be a subjective concept in which road users apply their own standards to define which parts of congestion are normal and tolerable. Also, the concept of desirable demand is a consequence of comparison. For example, the individually desirable level of travel demand and congestion might not be the socially desirable level.

In this dissertation, the definition of desirable refers to a concept widely used in economics: socially optimal. From an economic perspective, the socially optimal level of congestion occurs when the marginal social benefit (MSB) of travel equals the marginal social cost (MSC). At this optimal level, societies cannot achieve a larger net benefit by adding or removing a trip; the net profit to the whole society is at its maximum. Measuring the socially optimal level of congestion requires knowing not only the marginal private costs (MPCs) and benefits (MPBs) of travel but also the marginal external costs and benefits imposed on others. According to the first theorem of welfare economics, if no externalities of congestion exist in the society, the competitive allocation of travel-related activities and resources in the free market without any price interventions by government policy can achieve the socially optimal level of congestion. In the presence of externalities of crowding, if road users account for all of their impacts on others, the resulting level of congestion is socially optimal. This economic definition of what is the socially optimal level of congestion suggests that societies needs a certain degree of congestion, and the optimal level of congestion represents the maximum amount of traffic volume desired by the whole society.

Excessive congestion arises when the congestion level exceeds the socially optimal level. In this case the MSC of travel surpasses the MSB, leading to a net social cost; that is, a welfare loss. This research defines such a welfare loss as *congestion diseconomy*, which equals the total social cost minus the benefit of travel. In contrast, congestion can be insufficient when the congestion level is below the socially optimal level. In this case, the society could ask for more trips to support increasing economic activities. Insufficient congestion is better defined as insufficient agglomeration because agglomerated activities determine congestion rather than the reverse. The following sections present several graphical analyses theorizing when congestion is excessive, and they investigate what policies are efficient for adjusting congestion to the socially optimal level.

Market Failure from Congestion Externalities

> *"In principle, it is possible to determine whether congestion is excessive by examining the impact of adding an extra vehicle to a road or allowing an additional passenger onto a train. When the net benefits derived by the additional traveler are greater than the additional costs imposed on existing travelers, adding further cars onto the road or people onto trains will make the community better off overall. Increasing congestion would be consistent with increasing welfare, but when the net benefits derived by an additional driver or a train passenger are less than the costs imposed on existing road users or train*

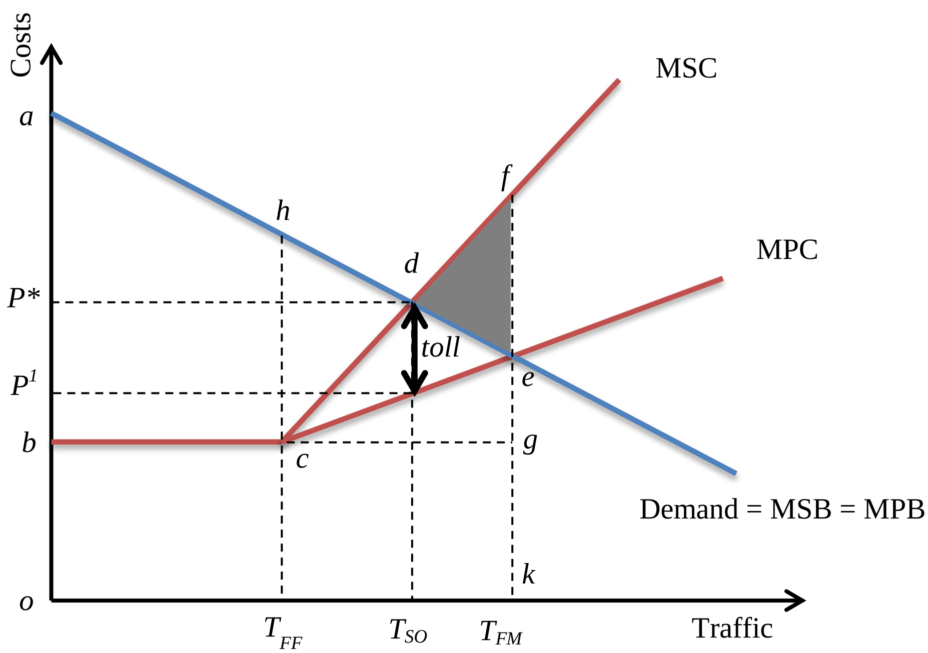
passengers, society as a whole is worse off due to the extra travel. When this occurs, congestion is said to be excessive." (VCEC, 2006: p55)

The quotation above is from a report of the Victorian Competition and Efficiency Commission in Australia (VCEC, 2006). It explains that excessive congestion arises when the MSC of adding extra drivers to a road exceeds the MSB (MSC > MSB). The key question is why such a situation would happen in the free market. In a socially optimal situation, MSC equals MSB, and the transportation market is efficient. No road users could be made better off without other road users becoming worse off. The situation in which MSC exceeds MSB can occur when the market fails to price the external cost of traffic congestion, that is, the congestion externality. Thus, underpriced travel is a crucial cause of excessive congestion.

Figure 2.2 illuminates how congestion externality as a source of market failure incurs excessive congestion. The graph analysis assumes that identical road users are expected to drive on a particular section of the road network during peak hours. The horizontal axis represents traffic flow passing the section within a fixed period, while the vertical axis represents costs for crossing the road section. The driving demand curve demonstrates that driving demand decreases with driving costs, and that the MSB of adding an extra vehicle to the road decreases with the traffic volume. Because no external benefits of driving are present here, the marginal private benefit (MPB) equals the MSB. In contrast, the MSC curve is the same as the MPC curve when traffic volume is low, for example, under the free-flow volume (T_{FF})[3]. The underlying assumption is that road users have less impact on other users when they can travel at the free-flow speed. The MPC curve represents the private cost of each vehicle for using the road, including time cost,

[3] The free-flow volume can be understood as the maximum traffic volume under free-flow speeds, which are often estimated below the speed limit. In this dissertation, we mainly consider travel delay as the external costs of congestion. If other external costs of travel are included, such as costs of air and noise pollution and crashes, the MSC curve should be above the MPC curve even when the traffic volume is low.

fuel cost, and vehicle maintenance cost. MPC is constant and relatively low when road users can drive at the free-flow speed and then increases with upward slopes when the traffic volume exceeds T_{FF}. After passing T_{FF}, MSC diverts from MPC because MSC internalizes congestion costs imposed on other road users and nonroad users, such as the cost of time delay.



Notes: MSC is the curve of marginal social costs; MPC is the curve of marginal private costs. The demand curve also represents the marginal private benefit (MPB). Since no external benefits of travel are present here, the marginal social benefit (MSB) equals MPB. T_{FF} is the traffic volume in the free-flow situation; T_{SO} is the socially optimal level of traffic volume; T_{FM} is the equilibrium level of traffic in the free-market without accounting for congestion externalities. The shaded area, triangle *def*, represents the magnitude of net costs to the society caused by excessive congestion, i.e., total congestion diseconomies, in the free-market equilibrium. P*- P1 is the Pigouvian toll to move MPC to intersect at the socially desirable point d.

Figure 2.2 Congestion Externalities as a Source of Market Failure Leading to Excessive Congestion

In a free market without pricing regulations, road users pay only for the private costs of their driving, not for external costs imposed on others. Free-market equilibrium occurs at the intersection of MPC and MSB, that is, the point where the MSB equals the

MPB. The free-market equilibrium level of traffic, T_{FM}, is the individually desirable level of congestion rather than the socially desirable level. After accounting for the external costs, the socially optimal level of congestion occurs at the intersection of MSC and MSB. The socially optimal level of traffic is T_{SO}, smaller than T_{FM}.

Figure 2.2 provides an intuitive approach to measuring excessive congestion from underpriced driving. The excessive driving demand equals the difference between the socially and individually desirable levels of traffic, i.e., $T_{FM} - T_{SO}$. In the free market, the equilibrium traffic produces a net benefit equaling the area of *abcd* minus the area of *def*. At the socially optimum level of congestion, the traffic level reaches equilibrium at the socially desirable level, creating a net benefit; that is, the area of *abcd*. Thus, excessive congestion can bring a net social cost up to the shaded area, the triangle *def*. This shaded area also represents the diseconomy of congestion.

These findings demonstrate that humans prefer to live and work in cities with a moderate level of congestion rather than without congestion at all. When traffic volume increases from T_{FF} to T_{SO}, congestion becomes worse but the total consumer surplus increases up to the area *cdh* (see Figure 2.2). Also, the congestion diseconomy is just a part of the full cost of congestion. When driving demand reaches the T_{FM} level, driving at the free-flow speed generates a cost of the area *bokg* and congestion brings a total cost equaling the area *cfg*. After accounting for the benefit of travel, the exact diseconomy of congestion, *def*, is smaller than the full social cost of congestion, *cfg*.

Because market allocations with excessive congestion are inefficient, improving the social efficiency requires eliminating excessive congestion or reducing those driving trips that are valued less than their social costs. An efficient policy is to adjust price signals that road users receive. This can be done by imposing a toll equaling the optimal marginal external cost (e.g., $P^* - P^1$ in Figure 2.2) on all road users. The toll makes road users face the full social cost of travel, including both private and external costs. Such a congestion pricing policy is also called Pigouvian congestion pricing.

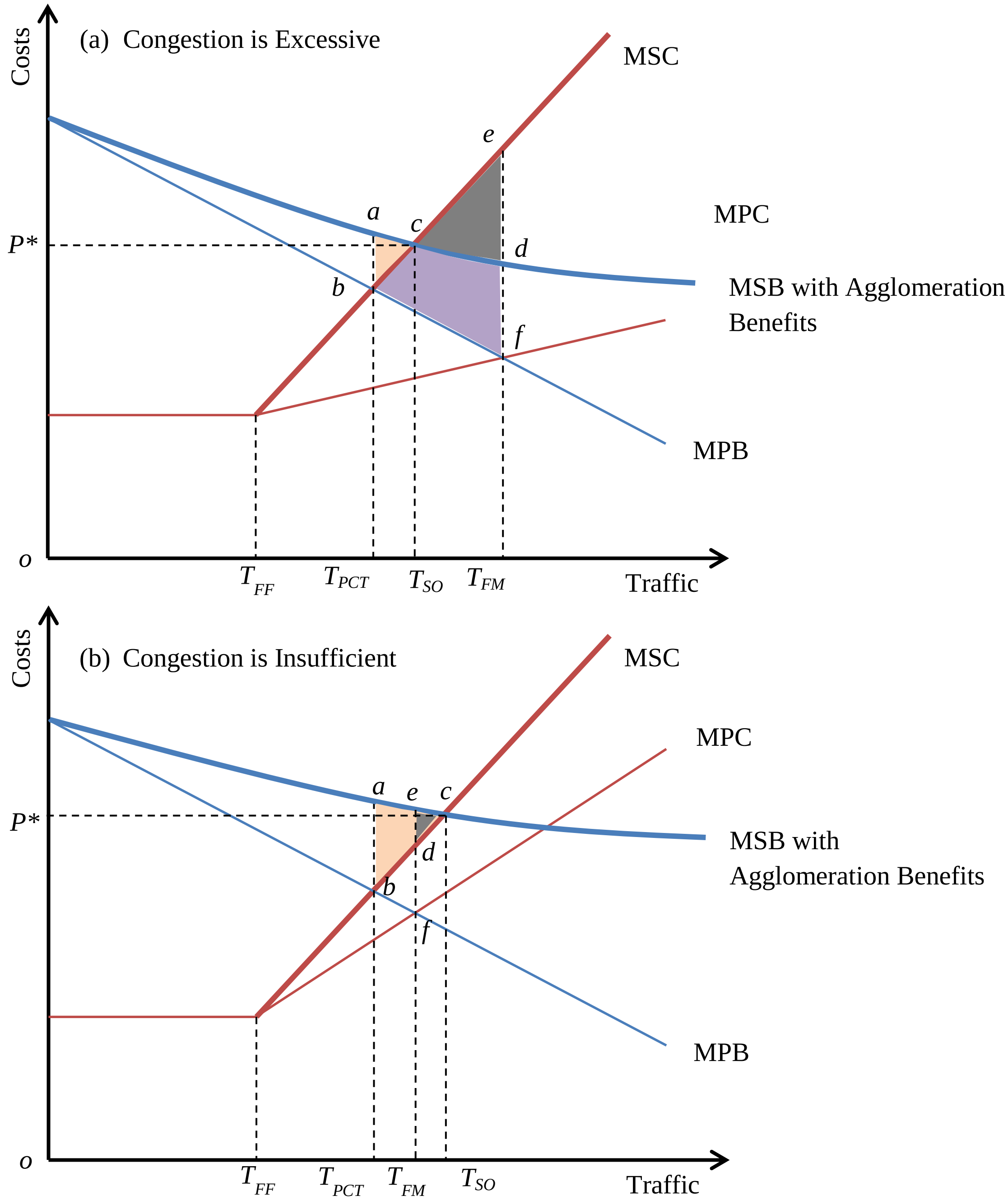
Market Failure from Congestion and Agglomeration Externalities

Previous analysis of market failure assumes that there is only one externality in the market, that is, the negative congestion externality. This section extends the discussion of market failure to a more realistic context by adding agglomeration externalities. Locations with agglomerated business companies and grocery stores, such as downtown or suburban centers, often generate and attract crowded traffic. Congestion is born with agglomeration; both congestion and agglomeration come from the spatial concentration of activities and traffic. Both residents and businesses can benefit from agglomeration. For example, people can benefit from living and working closer to each other. Firms can benefit from locating closer to each other for easier access to intermediate inputs and labor, lower transaction costs, easier job-worker matching, and knowledge spillovers (Fujita & Thisse, 2002; Puga, 2010; Rosenthal & Strange, 2004). Consequently, the congestion diseconomy is a type of agglomeration diseconomy, while the agglomeration economy can be seen as a kind of congestion benefit. Humans desire a certain degree of congestion, probably largely due to a need for strong agglomeration economies. Congestion is efficient if the agglomeration economy can fully compensate the congestion diseconomy.

Figure 2.3 provides a graph analysis of how congestion and agglomeration externalities together affect excessive congestion. While agglomeration externalities are on the production side, Figure 2.3 is a simplified interpretation assuming that increasing traffic volumes indicate increasing agglomeration benefits. This assumes that the external benefits to firms will be transferred to workers and that workers will use these benefits to make their travel decisions. The existence of agglomeration externality thus makes the

MSB greater than the MPB. The market reaches an optimum when the MSB equals the MSC—i.e., at point *c*—and T_{SO} is the socially optimal level of traffic when both externalities are fully corrected. TFM is the equilibrium traffic volume in the free market, and TPCT is the equilibrium level of traffic when only congestion externalities are corrected (e.g., by using Pigouvian congestion toll policies). The optimal level of traffic differs from the level when only congestion externalities are accounted for, i.e., at point b.

After internalizing both externalities, driving demand and congestion in the free market are not always excessive (Figure 2.3a) but are insufficient (Figure 2.3b) compared to those in the social optimum category. As shown in Figure 2.3a, if agglomeration economies are not considered excessive, traffic volume is T_{FM} -T_{PCT} and the congestion diseconomy is the triangle area *bef*. After considering agglomeration externalities, the excessive traffic volume becomes T_{FM}-T_{SO} and the congestion diseconomy decreases to the field of *cde*. These findings demonstrate that a part of congestion diseconomy is compensated by the agglomeration economy, i.e., the area *bcdf*. Excessive congestion may be overestimated if researchers recognize the cost and benefit of travel in the transportation market but overlook the extra benefit of helping to shape agglomeration. It is essential to estimate excessive congestion and congestion diseconomies within a framework that internalizes both congestion and agglomeration externalities.



Notes: MSC is the curve of marginal social costs; MPC is the curve of marginal private costs; MPB is the marginal private benefit; MSB is the marginal social benefit after accounting potential agglomeration economies. T_{FF} is the traffic volume in free-flow situation; T_{SO} is the socially optimal level of traffic volume; T_{FM} is the equilibrium level of traffic in the free market; T_{PCT} is the equilibrium level of traffic when only congestion externalities are fully priced using Pigouvian congestion toll policies.

Figure 2.3 Congestion and Agglomeration Externalities Leading to Excessive or Insufficient Congestion

After recognizing potential extra benefits from crowding, the socially optimal policy should not aim for eliminating all excessive congestion but for reducing excessive traffic to the optimal level. According to Figure 2.3, the optimal level occurs when the marginal congestion externality (MCE) equals the marginal agglomeration externality (MAE). When anti-congestion policies reduce traffic volumes from T_{FM} to T_{PCT}, they can bring welfare gains (i.e., net social benefits) at first, but eventually result in losses (e.g., from T_{SO} to T_{PCT}) because the agglomeration economy reduced by these policies exceeds the reduced congestion diseconomy, leading to a net social loss.

Figure 2.3b provides an example of an efficient market that may desire more, rather than less, travel and more congestion. Although increasing congestion will raise the amount of congestion diseconomy, this increased diseconomy can be offset by a rise in agglomeration economy. For example, when traffic volumes increase from T_{PCT} to T_{FM} the increased congestion diseconomy *bdf* is compensated by the increased agglomeration economy *abfe*, leading to a net social benefit of *abde.* In this case, efficient policies are those subsidizing agglomeration or travel rather than those against congestion. This finding is not to suggest that creating more congestion is necessary; it is to argue that a greater travel demand could be socially desirable. Even though congestion is insufficient, planners cannot rely on supply-side policies (e.g., narrowing down road space) but must instead rely on demand-side policies like subsidizing firm innovation and agglomeration.

When congestion is "insufficient," intensifying congestion may be a sign of economic growth. For example, when traffic levels increase from T_{SOLR} to T_{SO}, the city economy improves. In this process, despite the increases in congestion diseconomy, the agglomeration economy also rises at an even faster rate. Only when additional traffic leads to a larger diseconomy than the economy (i.e., MCE>MAE) can increasing congestion slow the city's economic development. Empirical studies often support the latter findings and report that traffic congestion can harm the city economy through slowing employment growth (Hymel, 2009), decreasing gross output (Boarnet, 1997), and reducing marginal agglomeration benefits (Graham, 2007; Weisbrod, Vary, & Treyz, 2001). While these studies assume that congestion is thoroughly negative to the city economy, fewer recognize the potential benefit of a certain level of congestion in economically healthy cities. Sweet (2014) provided an empirical study relying on the panel data of 88 U.S. metropolitan areas. His findings suggest that congestion will slow employment growth only when the levels of congestion measured by travel delay or daily traffic per lane exceed a threshold level.

In brief, these findings demonstrate the importance of incorporating congestion and agglomeration externalities in addressing congestion issues. Optimal anti-congestion policies should not only aim to alleviate traffic congestion but also to avoid eroding agglomeration economies. Anti-congestion studies should examine the interactions among transportation, land use, and production (labor) markets.

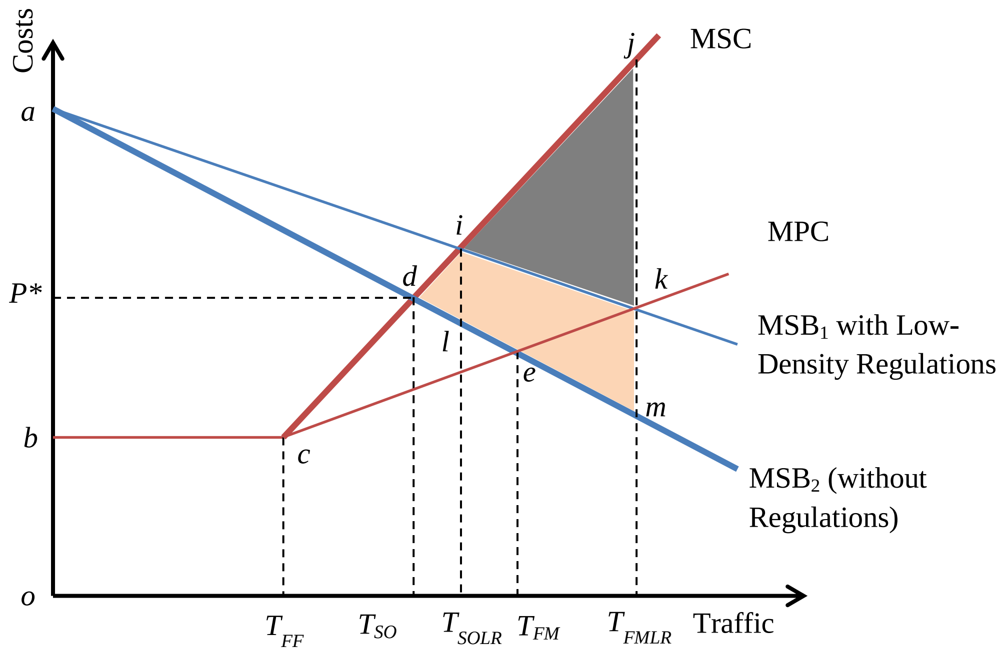
Excessive Congestion as a Result of Planning Failures

> *"There is a near-universal acknowledgment among transportation and land-use researchers that municipalities regularly employ their land-use regulatory powers to exclude denser development (e.g., Gordon and Richardson 2001; Boarnet and Crane 1997; Cervero 1989). Thus, a prerequisite to the development of alternatives is the liberalization of restrictive regulations that compel a low-density development pattern. If municipal regulations constrain development to this pattern despite market interest in alternatives, a paucity of these options is not a market failure but a product of regulatory policy, a "planning failure" (Cervero 1996). Sprawl's claim to being the market – and hence default – solution from which deviations demand justification in science would be undermined. Quite independent of travel behaviors benefits, the immediate payoff of such policies would be the expansion of transportation and land-use choice – that is, the ability of households to find the environments that fit their needs and preferences in housing type, neighborhood characteristics, and travel options." (Levine, 2006:9-10)*

Planning failure is a type of government or regulatory failure, which is a public sector analogy to market failure, and occurs when government intervention deters efficient allocation of goods and resources (Datta-Chaudhuri, 1990; Grand, 1991; Winston, 2000). While there are many planning policy departments, this research primarily tackles failures caused by inefficient land use regulations. Like market failures, these planning failures can cause significant market distortions. For example, many economic studies have explored the side effects of land use regulations by investigating their impacts on housing supply, affordability, productivity, and social welfare (Glaeser & Gyourko, 2003; Glaeser, Gyourko, & Saks, 2005; Gyourko, Mayer, & Sinai, 2006; Hsieh & Moretti, 2015; Mayer & Sommerville, 2000; Quigley & Rafael, 2005; Turner et al., 2014). Land use regulations mainly include those restricting the supply of housing and lands, such as low-density zoning, urban growth boundaries (UGBs), and other urban containment policies limiting land supply. Despite this line of research remaining inconclusive, many studies find that restrictive land use regulations can produce escalating housing prices or rent and losses of social surplus in the land market (Turner et al., 2014). Hsieh and Moretti (2015) provided an estimate that the U.S. GDP from 1964 to 2009 would likely have increased by 13.5% if high-productivity cities had removed all restrictions on housing supply.

While much economic literature has discussed the distortion in the land market caused by planning failure, less literature has extended this debate to the potential transportation impact. Few theoretical and empirical studies have assessed the effects of planning failure on mobility, accessibility, and congestion. These effects require an integrated investigation of the connection between land use and transportation, including questions of how land use regulations affect land development, how development affects accessibility and travel demand, and how all of these affect congestion on the roads. Only limited studies in the planning field have recognized and examined the potential roles of planning failure in the transportation market. For example, Cervero (1996) argued that local regulations such as single-use zoning exclude the potential for market-driven, job-housing proximity, leading to overloaded vehicle miles of travel(VMT). Levine (2006) presented a theoretical analysis of how planning failure differs from market failure in affecting travel behavior. He developed a novel paradigm to investigate and compare the potential difference between the supply and demand of alternative development with and without regulatory policies (for discussion about these empirical findings, refer to following sections). While this approach is highly likely to justify whether planning failure exists, and whether alternative development is preferred, it cannot quantify the cost and benefit of planning failure and its impact on the transportation market.

This dissertation develops an innovative approach to analyzing the connection between excessive congestion and planning failure, based on a similar supply and demand equilibrium analysis for market failure. This research primarily focuses on planning failures in the U.S. resulting from exclusionary zoning and low-density regulations that restrict maximum density, maximum height, minimum lot size, and single-use land. The underlying presupposition is that these land use regulations, mainly in the suburbs, have zoned out denser developments desired by the market. This zoned-out effect can lead to excessive urban sprawl, job-housing imbalance, lower accessibility, extra auto travel distance, and an overwhelming dependence on the vehicle. All these consequences will result in excessive demand for vehicle use and cause excessive congestion.



Notes: MSC is the curve of marginal social costs; MPC is the curve of marginal private costs; MSB1 is the marginal social benefit after implementing low-density regulations; MSB2 is the marginal social benefit without regulations. T_{FF} is the traffic volume in free-flow situation; T_{SO} is the socially optimal level of traffic volume; T_{FM} is the equilibrium level of traffic in the free market; T_{SOLR} is socially optimal level under land use regulations; T_{FMLR} is the free-market equilibrium level under regulations. The shaded area *ijk* represents the total diseconomy from market failure, and the area *dikm* represents that from planning failure.

Figure 2.4 Planning and Market Failures Together Causing Excessive Congestion

Figure 2.4 demonstrates how planning and market failures trigger excessive congestion on the roads in a city under low-density regulations. By comparing Figures 2.2 and 2.4, one can see there are two MSB curves in Figure 2.4, which represent the MSB before and after implementing low-density regulations. Because regulations could induce excessive driving demand, the MSB2 curve will shift right to MSB1 if regulations exist. The equilibrium levels of traffic are, respectively, T_{FMLR} and T_{FM} in the free market with and without low-density regulations. The congestion level is socially optimal when MSB2 intersects with MSC, making the T_{SO} the socially optimal level of traffic. In contrast, T_{SOLR} is the optimal level of traffic under the restriction of low-density zoning.

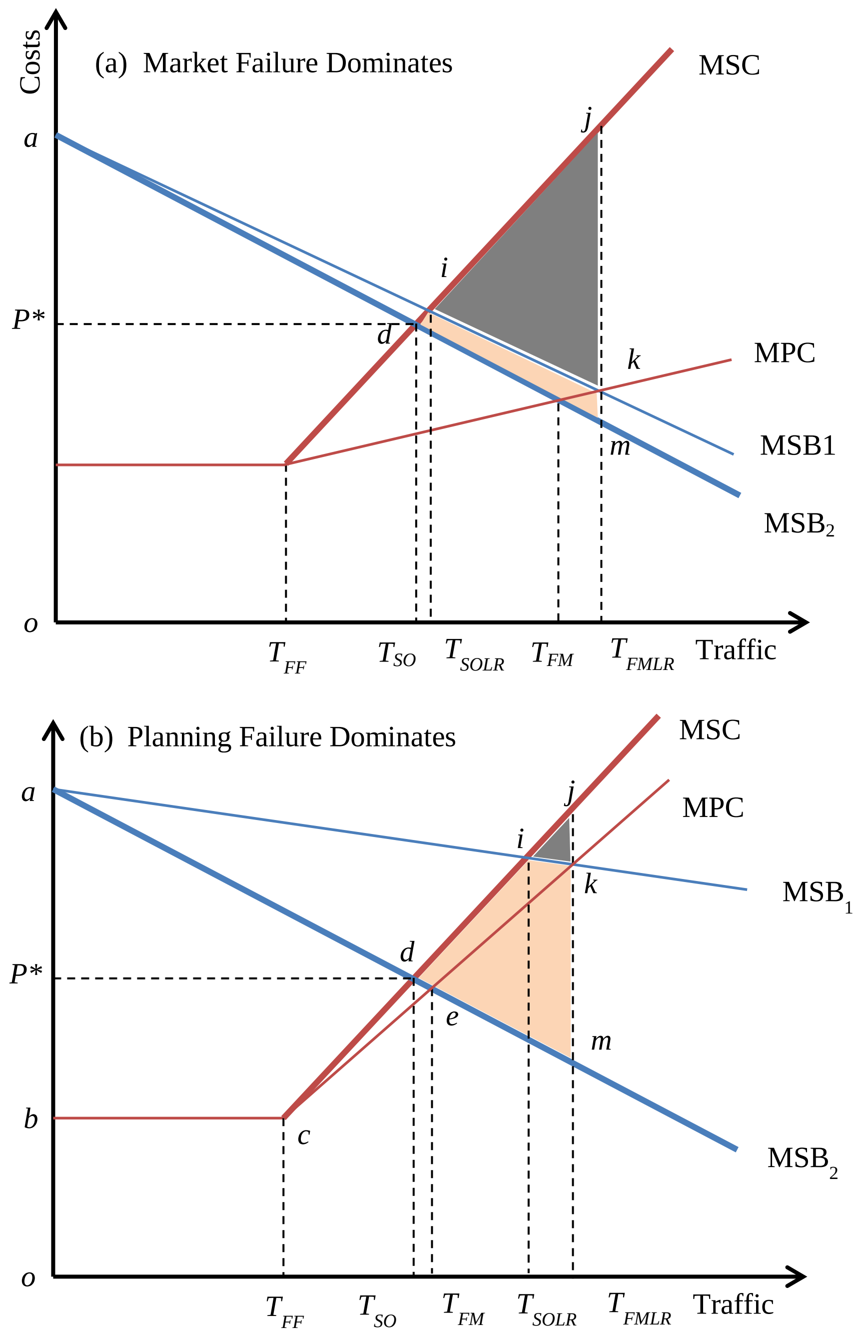
Accordingly, the individually optimal level of traffic in cities with low-density regulations is T_{FMLR} while the socially optimal level is T_{SO}. The total amount of excessive travel demand equals T_{FMLR} - T_{SO}. The excessive congestion can be divided into two parts. The first part comes from the excessive travel demand (T_{FMLR} - T_{SOLR}) due to market failure, while the second part (T_{SOLR} - T_{SO}) is caused by planning failure. The congestion diseconomy is the sum of two shaded areas – the trapezoid *idmk* due to planning failure and the triangle *ijk* due to market failure – that is, the triangle *dmj*.

While market failure can be fully remedied by charging underpriced travel a Pigouvian congestion toll, planning failure can be fully corrected by eliminating all regulations or allowing for denser or more compact development, encouraging market outcomes. The research framework placing market and planning failures together needs to examine the interaction between land use and transportation. In theory land use policies can be used to fully correct market failure, assuming land use and densities are regulated to the optimal land use pattern (e.g., Pines & Sadka, 1985; for more discussion, refer to Chapters 3 and 5). However, pricing strategies probably cannot fully correct the planning failure, because the constraints on land use are insensitive to pricing signals. That is, pricing probably cannot remove regulations. However, pricing policies could help residents or firms leave the planning area with regulations (for more discussion, refer to Chapter 5). Therefore, congestion relief needs to incorporate both land use and pricing policies. If planning failure exists, land use policies can play a dominant role, rather than a replaceable role, as advocated by the proponents of pricing in congestion reduction.

Market versus Planning Failures

Figure 2.5 explicates in which situations planning (or market) failure plays a more important role in determining excessive congestion, as a straightforward response to my research question (2). As shown in Figure 2.5a, market failure will dominate when marginal external costs are high and when zoning regulations are less restrictive and have an insignificant impact on driving demand. When market failure dominates, pricing policies can reduce most of the excessive congestion. In contrast, land use policy is a dominant policy for alleviating most excessive congestion when planning failure dominates (Figure 2.5b). This situation occurs when land use regulations increase significant driving demand and worsen accessibility, and when marginal external costs of congestion are relatively small.

In summary, it is unfair to conclude that economic policies are always superior to land use policies or vice versa. The evaluation of anti-congestion policies should compare the congestion diseconomy caused by the market with that caused by planning failures. This will help to determine which land use and pricing policies are efficient.



Notes: T_{SO} is the socially optimal level of traffic volume; T_{FM} is the equilibrium level of traffic in the free market; T_{SOLR} is socially optimal level under land use regulations; T_{FMLR} is the free-market equilibrium level under regulations. The shaded area *ijk* represents the total diseconomy from market failure, and the area *dikm* represents that from planning failure.

Figure 2.5 Congestion Diseconomy from Market versus Planning Failures

REASONING BEHIND AND FINDINGS OF ANTI-CONGESTION POLICIES

This section reviews two anti-congestion policies: land use planning and congestion pricing. Many research fields have widely discussed the impact of these policies on travel behaviors and transportation performance, including the areas of urban planning, transport engineering, economics, geography, regional science, and public health.

Land Use Policies

There are three major lines of reasoning addressing why land use policies can be an effective strategy for reducing travel costs and congestion. These are explored below.

Land Use Design for Reducing Local Travel Times and Costs

Land use characteristics such as land use designs, diversities, and densities are assumed to affect the time, length, frequency, and cost of local trips. Many studies have provided the behavioral reasoning behind the connection between land use and travel (Boarnet & Crane, 2001; Cao, Mokhtarian, & Handy, 2009; Cervero & Kockelman, 1997; Kockelman, 1997; Zhang, 2004). For example, compared to cul-de-sac street patterns, the grid street design lowers travel costs of both walking and driving and increases the comparative advantage of walking for longer trips (Boarnet & Crane, 2001). Mixed-use neighborhoods are often more walkable, reducing local vehicle travel. They can also capture a larger share of local trips and decrease more regional travel than single-use neighborhoods, reducing total travel length and traffic volume (Ewing, Greenwald, Zhang, Walters, Feldman, et al., 2011).

Neighborhood-level land use is often summarized as the three Ds: density, diversity, and design (Cervero & Kockelman, 1997). The three Ds were later extended to five Ds by including distance to transit and destination accessibility (Ewing & Cervero, 2001), and then to seven Ds by adding demand management and demographics (Ewing et al., 2011). The effectiveness of a land use policy is often justified by a research paradigm, which generally applies a reduced-form model (e.g., regression models) to examine whether desirable land use exerts significant impact on travel behaviors (see reviews by Badoe & Miller, 2000; Crane, 1999; and Ewing & Cervero, 2001, 2010). Despite varying empirical results, most studies demonstrated that compact development with high-density, mixed-use, transit-oriented, and pedestrian-friendly built environments is a more-or-less effective strategy to reduce driving (e.g., Ewing & Cervero, 2001, 2010; Litman, 2014).

This paradigm of research can evaluate traffic benefits of land use policies and help to improve the standards of architecture, neighborhood design, and engineering design for reducing inefficient travel. However, these reduced-form studies are insufficient to determine whether land use policies are desired by residents, neighborhoods, or the market. A major challenge is that the travel benefits of land use policies are often evaluated locally while the congestion diseconomy is often assessed regionally.

Another challenge comes from residential self-selection studies, which argue that people's attitudes can affect residential location choice and related travel outcomes. As a result, land use impact on travel behaviors could be misestimated without considering travel preference (Handy, Cao, & Mokhtarian, 2006; Cao et al., 2009; van Wee, 2009). The underlying assumption is that residents will "vote with their feet" into their desirable neighborhoods based on their travel preferences. Compact neighborhoods may be desirable only for those preferring nonauto modes to driving. Thus, travel benefits of land use policies found in empirical studies are insufficient to identify the desirable demand of specific land use patterns, the desirable travel varying with modes, and the desirable level of congestion.

Land Use Planning as an Alternative Policy for Correcting Market Failures

> *"In the absence of true market-based pricing of transportation, public initiatives that reduce automobile dependence and thus help conserve finite resources must be turned to. In the jargon of economists, physical land-use planning becomes a second-best response to the inability to introduce first-best, Pareto-optimal pricing." (Cervero, 1998: 18)*

As noted in previous sections, the existence of congestion externalities[4] is the primary source of market failure, causing excessive auto travel and congestion. If road users paid only for the private cost of their driving, they would balance private costs and benefits to achieve an individually optimal level of driving. However, the private cost never includes the external cost imposed on others. When driving is underpriced, the individually optimal level of driving is greater than the socially optimal level, leading to excessive travel and congestion. To eliminate the excessive congestion, there are two approaches to moving the traffic volume from the individually desirable level to the socially desirable level. These include pricing and quantity regulation strategies. While pricing policies will be discussed in the following sections, this section focuses on land use policies as an alternative regulation for eliminating excessive congestion.

Relationships between land use and excessive congestion are complicated and require an analytical framework for connecting land use and transportation markets. This

[4] The economic justification for land use planning or zoning regulation is primarily related not to congestion externalities but to land use externalities. The first type of land use externalities comes from the publicly provided good (Tiebout, 1956) and zoning is effective to sustain an optimal community size for using public property; for example, avoiding overcrowding. Exclusionary zoning is also effective to prevent "free riders," low-income outsiders who seek to live near neighbors with higher housing consumption than themselves. Without excluding these free riders, some people will pay less on property taxes even though they share the same benefit from public goods with those paying more (Hamilton, 1975). The second type of land use externalities comes from the potential adverse effects related to the proximity of incompatible land uses. Land use regulation can be seen as a tool to correct these negative externalities, protect property rights, and enhance the system of nuisance law (Clawson, 1971; Ellickson, 1973; Fischel, 1985). These land use externalities, however, are less related to traffic congestion and thus are not considered in this dissertation research. For related discussion, refer to Fischel (1985) and Levine (2006).

line of research primarily lies in the field of urban economics. Many studies have applied the Alonso-Muth-Mills (AMM) model, that is, the monocentric model, to investigate traffic congestion (e.g., Brueckner, 2007; Kono & Joshi, 2012; Pines & Sadka, 1985; Wheaton, 1998). They demonstrate that land use allocations in the free market with underpriced driving would cause more sprawling than the optimal allocation. The Pigouvian congestion toll policy that charges each driver a toll to cover the gap between the MPC and MSC of each trip is the optimal (first-best) tolling strategy. In a closed-form region with a fixed population, a first-best congestion toll would raise residential densities near the urban core and slightly lower-edge densities near the city boundary (Pines & Sadka, 1985; Wheaton, 1998; Kono & Joshi, 2012). An appropriate policy of lot-size zoning can replace the first-best pricing to reach the social optimum, including an upward adjustment of central densities and a downward adjustment of edge densities (Figure 2.6). In theory, if a city can implement the first-best pricing policy and the optimal land use regulation, the efficiency of either policy is equivalent. While many urban economists embrace the pricing policy, they often regard land use policies as second best or substitutable.

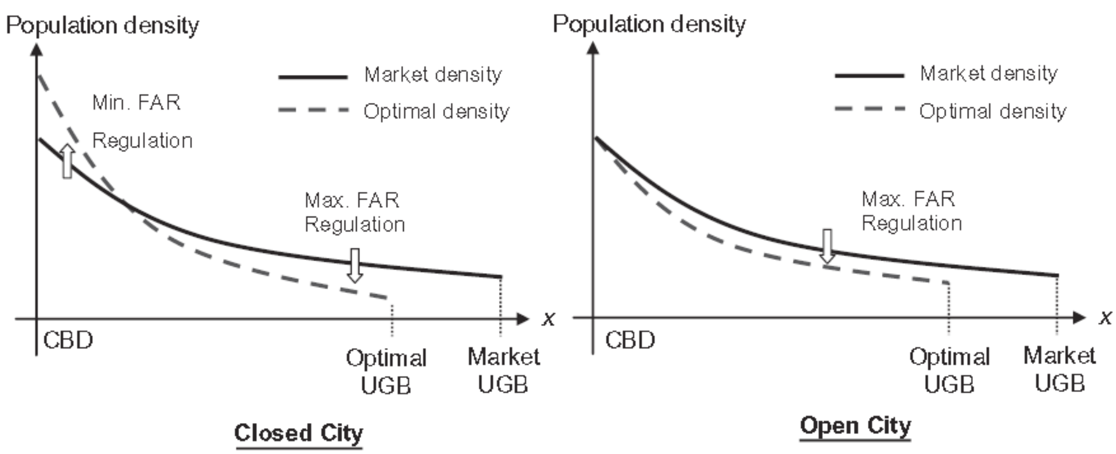


Figure 2.6 Market density and optimal density (under the first-best tolling) in a monocentric model (Kono and Joshi, 2012)

While the search for optimal land use policies for eliminating excessive congestion largely relies on analytical and simulation studies, optimal land use planning policies are never observed in practice. Only second-best or third-best land use policies are implemented in reality. For example, an imposition of an UGB may be an effective second-best policy to reduce excessive congestion because a UGB increases densities. However, Brueckner (2007) argued that UGBs achieve much lower welfare improvements than first-best tolling strategies. His monocentric spatial equilibrium model found that the best UGB offered just 0.8% of the welfare gain from levying congestion tolls. Presumably, the UGB could not foster strong central densification. Similar results can be found in Kono, Joshi, Kato, and Yokoi (2012), who discovered that a UGB policy alone is a poor substitute for first-best tolling, and that optimal regulation of building size for higher central densities, plus a suitable UGB, is an effective second-best remedy. It seems that welfare gains from restrictive UGBs are largely offset by welfare losses from their side effects, such as land rent escalation and reduced areas for development.

Many empirical studies have also reported that housing rents or prices inside the UGBs rise faster than outside the UGBs (Cho, Poudyal, & Lambert, 2008; Staley, Edgens, & Mildner, 1999). While the confines of a boundary, higher-density development on pace with population growth and immigration is important, it can be challenging. Speculation is also problematic in various settings. For example, London, England and Auckland, New Zealand have reportedly experienced major rent escalations due to relatively low housing supply from the release of land for new development (Cox, 2010). Home affordability is thus a critical topic for debate in growth-management discussions (Downs, 2004). Moreover, development activities in 95 relatively contained U.S. metro areas (as contained by city limits, greenbelts, and/or UGBs) are more agglomerated near their central cities than those in uncontained areas (Nelson, Arthur, Raymond, et al., 2004).

Therefore, land use policies could be an alternative strategy for reducing excessive congestion. Further studies are needed to investigate how first-best and second-best congestion pricing affect land use, how second-best land use policies affect excessive congestion, and whether there are potentially negative consequences of second-best land use policies.

Land Use Planning as a Prerequisite Policy for Correcting Planning Failures

To reduce excessive congestion by correcting planning failure, researchers often discuss two practical policies: liberalizing land use regulation via regulatory reform and promoting market-desired denser development. These policies often promote development patterns as alternatives to auto-oriented or sprawling development, such as compact development, mixed-use development (MUD), and transit-oriented development (TOD). Land use regulations, including low-density and single-use zoning, can produce low-density urban sprawl and more auto dependence than desired by the market, leading to an inefficient land market (Bogart, 1998; Fischel, 1985; Gordon & Richardson, 2001; Pendall, 1999; Talen & Knaap, 2003).

Rather than providing scientific proof of the traffic benefits of land use policies, Levine (2006) advocated a new research paradigm. This paradigm only needs to investigate whether people demand alternative development, whether the supply of such development is below people's demand, and whether existing land use regulations cause an undersupply of alternative development. Levine and Inam (2004) conducted a stated preference survey of 676 U.S. developers and found that most developers believe there is an oversupply of auto-oriented development and an undersupply of alternative development, such as TOD and MUD. They ascribed the mismatch between supply and demand to low-density zoning regulation.

On the demand side, the land use supply under low-density regulations may not fit a household's demand for their expected residential land use. In this case, two groups are assumed to emerge. The first is the matched group whose current residential land use corresponds to their residential preference. The other is the mismatched group whose current living environment conflicts with their residential preference. For instance, Schwanen and Mokhtarian (2004) discovered that 23.6% of workers in the San Francisco Bay Area can be classified as part of the mismatched group. These workers lived in urban (versus suburban) neighborhoods but had a low (versus high) preference of denser neighborhoods. Frank, Saelens, Powell, & Chapman (2007)'s study in Atlanta showed that 51% of surveyed residents lived in neighborhoods with high walkability and liked walking, while 28% lived in neighborhoods with low walkability and preferred less walking. In contrast, 17% lived in neighborhoods with low walkability but preferred high walkability, and only 5% lived in neighborhoods with high walkability and preferred low walkability. Thus, a total of 22% of households had mismatched residential preferences. A relaxation of regulatory barriers to alternative development is suggested to reduce the number of mismatched groups (Levine, Inam, &Torng, 2005).

While these studies demonstrate the necessity of promoting regulatory reform and alternative development for correcting planning failure, they do not estimate benefits and costs of these policies. When they recommend for more alternative development projects, they can neither tell how many projects of this kind are socially desirable nor evaluate the effectiveness of an alternative development project on reducing congestion.

Congestion Pricing Policies

Congestion pricing strategy is an economic approach to correcting market failures by charging the external cost of congestion as a toll to whomever causes it. Ideally, the congestion toll should equal the marginal external cost, that is, the gap between the marginal social and private costs. As a result, the Pigouvian congestion toll scheme would fully correct the market failure of traffic congestion and adjust the market to its optimal level (first-best). However, the optimal toll is difficult to estimate and charge in reality. Many practical pricing schemes adopt alternative pricing methods (second-best), for example, in the form of cordon charges, area-wide pricing, and variable-rate highway tolling. Pioneering examples around the world include Singapore's cordon charge in the early 1970s, Norway's toll rings in the mid-1980s, and London's area-based toll in 2003 (Santos, 2005; Ieromonachou, 2006).

In the U.S. increasing numbers of metropolitan areas have built or are building toll roads. A recent report from the United States Government Accountability Office (GAO) summarized all congestion pricing projects that receive federal funding including either high-occupancy toll (HOT) lanes or expressways (USGAO, 2012). Most tolls in these managed lanes are variably priced across traffic periods and locations and range from 25 cents to $14 around the U.S. (USGAO, 2012). Twelve HOT facilities were operated in 10 metropolitan areas until late 2011, when 13 HOT lanes were under construction or extension (USGAO, 2012). According to the Federal Highway Administration (FHWA), from 1998 to 2010 tolled miles in urbanized areas jumped 36% (2012). We can foresee a future of booming toll-road construction. The federal government, which traditionally prohibited federal funding for toll roads, has turned to permitting federal participation in tolling projects, authorized in the transportation bill Moving Ahead for Progress in the 21st Century Act (MAP-21), passed in 2012.

Pricing's Effects on Travel and Traffic

Congestion pricing is regularly regarded as a policy affecting various aspects of travel behaviors, including route choice between priced lanes or unpriced lanes, travel departure time trading off between avoiding tolls or saving time in congested periods, travel mode, destination, trip chains, trip frequency, activity selection, and car ownership (Deakin, Harvey, Pozdena, & Yarema, 1996; Giuliano, 1994). Numerous empirical studies have explored the effects of existing and potential congestion pricing projects or policies on travel demand and congestion (Table 2.1).

Early studies such as Bhatt (1993) and Harvey (1994) relied on very limited aggregate data of travel to conduct ex-ante research, such as predicting pricing elasticities of travel demand (traffic volume or VMT). They discovered that the pricing elasticities of travel demand may vary across cities and locations, but are often negative and range between 0 and –1. This range indicates that pricing effects are relatively inelastic compared to other goods or services, and that doubling travel costs cannot bring a double reduction of peak-hour travel demand. Yet, pricing policies can indeed decrease travel demand, and a small change of motorist's behaviors may bring much improvement in travel flow (TRB, 1994).

Several studies directly explore the pricing impact on the regional or citywide aggregate travel demand based on different travel demand models. Using a systematic travel demand analysis, Deakin et al. (1996) provided a more realistic ex-ante travel demand analysis under several pricing scenarios. The scenarios with congestion pricing as one important component suggest that appropriate pricing policies can lead to an abundance of benefits, including a reduction in VMT, fuel use, and emissions. Two recent studies in Austin, Texas, relying on a more rigorous travel demand model with a joint destination-mode choice model, suggested that a Pigouvian congestion toll scheme on either major or all freeways could significantly reduce average peak travel time and total system VMT in the short and long term (Gulipalli & Kockelman, 2008; Kalmanje & Kockelman, 2004).

After the 1990s, because a considerable number of congestion pricing projects were built around the world, much recent research has begun to conduct ex-post evaluations (Matas & Raymond, 2002; Odeck & Brathen, 2008). Although most ex-post analyses focus on local impact rather than the regional effects of toll road projects, many findings are comparable to the ex-ante studies. For example, they have found that higher road tolls may result in a decline in travel demand in the short and long term. A study of 19 toll road projects in Norway revealed that the average long-term pricing elasticity (–0.82) is about twice as high as the short-term one (–0.45) (Odeck & Brathen, 2008). A report by the U.S. GAO (2012) presented a more direct comparison of travel change before and after pricing was imposed. This report used data from the Department of Transportation and summarized 14 congestion pricing projects in 5 HOTs and 9 peak-period roads. The findings demonstrated that congestion pricing facilities can improve traffic conditions and reduce congestion. Specifically, HOT lane projects generate a significant reduction in travel time and vehicle throughput and an increase of travel speed in both priced and unpriced lanes. Although peak-period pricing projects tend to exert no impact on aggregate traffic demand, they probably cause drivers to shift trips from on-peak to off-peak periods.

Besides the effects on travel time and congestion, a small number of empirical studies have explored the impact of anti-congestion policies on local and regional effects on transit use or ridership. However, their findings are mixed. Among four HOT lane projects evaluated in the U.S. GAO's report, only Interstate 95 in Miami generated a significant increase in transit ridership, a 57% increase within 2 years after the toll road was opened. Gulipalli and Kockelman (2008) predicted the effects of the marginal cost pricing, i.e., the Pigouvian congestion toll, on travel mode shifting in the Dallas-Fort Worth (DFW) region of Texas. They found no significant evidence of the relationship between pricing and mode shifts. However, the effect of pricing on transit ridership relates to not only the pricing mechanism but also the availability of public transit. That is the reason many successful congestion pricing schemes occur in metropolitan areas with excellent public transit systems. Improved public transport and significant use of public transport are potentially important policy complements to congestion pricing (Anas & Lindsey, 2011).

In summary, most of the empirical studies, either from ex-ante or ex-post analyses, have shown significant effects of congestion pricing on the reduction of travel time (particularly peak-hour travel time), congestion, VMT, and traffic throughput in both local and regional areas. More ex-post studies of tolling effects are needed, especially of the local impact, because toll roads in most U.S. cities would mainly generate local rather than regional effects at first. However, less empirical research has paid attention to the potential local impacts on non–road users who live or work in the neighborhoods affected by toll roads. These people might not use adjacent toll roads. However, these toll roads or tolling policies may affect these people's travel decisions in the short run and their residential and job location choices in the long term. Such indirect but likely important effects of congestion pricing need more studies with a comprehensive framework accounting for the relationship between congestion pricing, neighborhood land use, and job markets.

Table 2.1 Empirical Studies of Pricing Effects on Travel Behaviors, Land Use, and Social Welfare

Source	**Cases**	**Role of Congestion Pricing**	**Quantitative Effects**	**Methodology**
		Travel & Traffic		
Bhatt(1993)	Metropolitan Washington Region	A price increase would result in a decline in travel demand.	Estimated the pricing elasticity of travel demand: -0.01to -.015 at the low end, -0.3 to -.4 at the high end	Projected elasticity analysis
Harvey(1994)	San Francisco and Los Angeles	Regional wide pricing would reduce VMT, and largely shift trips to off-peak hours	San Francisco: 1.8% reduction in VMT Los Angeles: 5% reduction in VMT 10-15 minutes round-trip saving in peak-hour travel	Projection
Deakin et al. (1996)	Los Angeles, Bay Area, San Diego, and Sacramento metropolitan areas	Transportation pricing measures could effectively relieve congestion, lower pollutant emissions, reduce energy use, and raise revenues.	A combination of congestion pricing, employee parking charges, a 50 cent gas tax, and mileage and emissions feeds would reduce VMT and trips by 5-7% and cut fuel use and emissions by 12-20%, varying with region.	Simulation using a travel demand modeling package, STEP
Matas and Raymond (2002)	18-year panel data of 72 road sections in Spain	Travel demand is relatively sensitive to toll changes	Elasticities: -0.21 to -0.83	Elasticity analysis by panel data model
Kalmanje and Kockelman (2004)	Austin, TX	Pigouvian congestion toll reduces peak-hour travel time.	Average peak travel times decreased by roughly 1.6% (3.3% on main roads), with tolls averaging roughly 1.5 cents per mile on main roads during peak congestion.	A standard four-step travel demand with joint destination-mode choice models
Gulipalli and Kockelman (2008)	Dallas-Fort Worth (DFW) region of Texas	Both two MCP scenarios (MCP-on-freeways and MCP-on all-roads) reduced total VMT. No significant mode shifts were found.	Total system VMT for both the MCP scenarios would fall by about 6-7% in the short term and 7–8% in the long run.	Forecast model with joint destination-mode choice models
Odeck & Brathen (2008)	19 toll road projects in Norway	Pricing elasticities of travel demand vary with road type and project location.	A mean short-run elasticity at -0.45. A mean long-run elasticity at -0.82.	Elasticity analysis by empirical data

Table 2.1 (*Continued*)

Source	Cases	Role of Congestion Pricing	Quantitative Effects	Methodology
USGAO(2012)	14 congestion pricing projects (5 HOTs and nine peak-period expressways) in the US. Data from FHWA&DOT	Both travel time and speed improve on at priced and/or unpriced lanes of all five HOT projects. No effects are found on nine peak-hour expressways. Vehicle throughput is increased in 5 HOT lane projects while no effects are found on expressways. Peak-hour pricing projects motivates a driver to take trips at an off-peak time. Mixed results exist in the HOT's effects transit ridership while no effects are found on the peak-hour projects	Settle SR167 HOT: 19% increase of peak-hour speed on unpriced lanes San Diego I-15 HOT: 20 minutes less in HOT lanes than unpriced lanes in congested time Miami I-95 HOT: 14 mins less in the HOT lanes and 11 mins less in the adjacent unpriced lanes per trip Minneapolis I-394 HOT: 9-13% increases in vehicle throughput in the HOT lanes, 5% increases in the unpriced lands Orange County SR 91 HOT: 21% increases in vehicle throughput on the entire roadway Miami I-95 HOT: 57% increase in weekday ridership	Before-and-after comparison
		Land Use (households' & firms' locations)		
Boarnet and Chalermpong (2001)	Toll roads in Orange County, CA,	New highways may raise house prices, and home buyers are willing to pay for the increased access that the new roads provide.	No specific effects of tolling are studied	Before-and-after study
Kalmanje and Kockelman (2004)	Austin, TX	Home values are predicted to fall slightly in almost all areas when all roads are the price. Residential property prices are estimated to fall marginally in the most area when pricing only major roads.	Pricing on all roads: Home values are estimated to fall between 1.5% and 6.4% in southwest Austin, but other regions (including the CBD) are predicted to experience lesser drops. Pricing on the main roads: Home values again dropped slightly in most areas, but home values in some CBD areas were predicted actually to increase marginally.	A standard four-step travel demand relying on logit models of departure time, mode, and destination

Table 2.1 (*Continued*)

Source	Cases	Role of Congestion Pricing	Quantitative Effects	Methodology
Gupta, Kalmanje, & Kockelman (2006)	Austin, TX	Toll roads are not found to impact Austin's land use pattern significantly, but they are predicted to initiate some development in localized areas along the toll corridors	75% of Austin's population is predicted to experience welfare gains of less than 3 ¢ per day under a toll road policy of 10 ¢/mile Drivers in 14 zones suffer from welfare loss of less than 2 ¢ per day per individual.	Simulation in DRAM-EMPAL models
Ubbels & Verhoef (2008)	Dutch	Three policy schemes with flat kilometer charge reveal that congestion pricing may have a considerable effect on car use, road usage, car ownership, and residential/job relocation	Car use: 6-15% fewer of car-based trips Car ownership: about 2% of the respondents would sell at least one of their cars, and 1.6% would consider giving up using cars. Relocation: 4% of households would probably change their residence location, and 11% would probably change their job location.	State preference survey on the Internet
Tillema, van Wee, Rouwendal, & van Ommeren (2008)	Dutch	A kilometer charge may change firm's behaviors in trip decision,	30-40% of firms would change decisions in firm-related travel. 30% of the employers would reimburse their employees for the loss due to pricing. 7.8% of firms would probably relocate.	State preference questionnaire
Pugh & Fairburn (2008)	The Staffordshire part of the M6 Toll corridor, UK	The toll road has caused a positive industrial land development effect at the sub-regional level.	Increased industrial land development of 3.01 hectares were found in location within a five-minute drive time of an M6 Toll Junction; 1.24 hectares for those within a 10-minute drive time; and no effects for within a 15-minute drive time.	Before-after panel data models
Vadali (2008)	Several toll road segments in Dallas County, Texas	Toll roads increased property values at the 0.25-1-mile areas even in several years before the toll road was open. Both toll road extension and new toll roads can bring development.	The spatial effects of toll roads on property value would decrease by distance.	Before-after panel data models

Table 2.1 (*Continued*)

Source	Cases	Role of Congestion Pricing	Quantitative Effects	Methodology
Anas (2013)	Chicago MSA	Tolling only the major roads decentralizes jobs and residences out of the City and the inner suburbs to the outer suburbs causing land development to increase.	When all roads are tolled then more employment and residents move out of the inner suburbs and concentrate much more in the CBD and the rest of the City and increase less in the outer suburbs, and land development is also higher.	Simulation based on RELU-TRAN2

Pricing's Effects on Land Use and Development

Empirical research has paid little attention to congestion pricing land use impact, despite much literature endorsing pricing as an effective management policy for reducing congestion, as summarized in the previous section. Congestion pricing probably differs from other sources of transport revenue such as fuel tax, sales tax, and income tax in its potential to affect decision making about land use development (Urban Land Institute [ULI], 2013). Tolling and related schemes to charge a tax or fee for every mile driven will influence land use decisions much more directly (Deakin et al., 1996; ULI, 2013) because tolls affect travelers' budget constraints and lead to mode switching and redistribution of trips. Tolls may also affect firms' labor costs and production and service demand, and can result in geographic redistribution of businesses (Deakin et al., 1996; Santos & Shaffer, 2004; Zhang & Kockelman, 2014).

According to a 2012 report by the ULI, more than 35 experts in the fields of transportation and land use planning believed that a VMT tax may accelerate new development in compact, mixed-use, and walkable nodes and may affect land use for industry, office, and especially retail (ULI, 2013). The land use along toll roads is perceived to be more compact than that along highways (Litman, 2011), although the practical expressway's effects on real estate remain blurry. For example, in Austin mixed-use development is found along the toll road 183A in the fast-developing north suburbs, while less development has emerged along the toll roads in the southern suburbs (Spivak, 2013).

Although casual relationships between pricing and land use remain ambiguous, this section summarizes four primary arguments in the literature. First, pricing schemes on toll roads may promote more compact development along toll roads than unpriced highways. For example, Ubbels and Verhoef (2008) and Tillema et al. (2008) conducted a stated preference survey in Netherlands and discovered that a significant number of residents and firms would relocate for closer job-housing proximity after the imposition of a linear congestion toll. Road tolling can partially correct market failure from congestion externalities, reduce urban sprawling due to the unpriced highway, and raise surrounding population and employment densities and land use mixtures. Litman (2014) suggested that tolling may lead to more compact development and more traffic improvement than enhancing public transit infrastructure. Similar results were found in a simulation study in the Austin area by Gupta, Kalmanje, and Kockelman (2006). Their findings suggested that congestion pricing could catalyze land development surrounding toll roads but have less influence on most other areas. Also, tolling may cause job decentralization. For example, Anas (2013) discovered that tolling on major roads can decentralize employment and residence from the inner suburbs to the outer suburbs, increasing land development in suburban areas.

Second, congestion pricing may redistribute traffic from toll roads to freeways, from tolling regions to no-toll regions, and from tolling periods like peak hours to no-toll periods. These redistribution effects may not change land use much. However, little empirical research has been conducted to justify this argument and what exists does not compare the potential development difference along tollways and highways. Instead, most studies have focused only on the effect of highways on real estate development in housing and industry (Ewing, 2008; Pugh & Fairburn, 2008), land rent change (Boarnet & Chalermpong, 2001; Vadali, 2008; Ewing, 2008), and demographic shifts due to relocations of residents and firms (Chi, 2012).

The third argument is that congestion pricing is a necessary condition but not a sufficient condition for triggering compact development if planning failures exist. Even an optimal pricing scheme could not create an optimal land use pattern if land use regulations in the past and present restrict the generation of optimal land use. As argued by Deakin, “zoning regulation would prevent landholders from increasing the density of development” (1994, p. 235). Low-density zoning regulations and separate-use subdivisions are two major sources of planning failure leading to urban sprawl (Levine, 2006), which may damage mobility, accessibility, and social welfare. Langer and Winston (2008) presented an empirical analysis that found an interactional effect of population density and congestion pricing on net social benefits. Improving land use compactness by relaxing low-density zoning would make congestion pricing policies more efficient.

Finally, a self-selection effect may exist in the relationship between congestion pricing and land development. Congestion pricing projects are more welcome in areas with enough facilities to support travel alternatives to driving and less welcome in areas with dispersed land use patterns and no transit service, because “those who find the tolls to be too expensive may not have a viable alternative” (Mahendra, Grant, & Swisher, 2012: 17). In this case, one cannot easily judge whether congestion pricing causes compact development or compact development attracts pricing projects, especially when tolls are imposed on old highway facilities. On the other hand, this potential self-selection effect may emphasize the importance of coordinating congestion pricing and land use policies.

However, both theoretical and empirical evidence of incorporating land use and congestion pricing remains ambiguous. Only limited numbers of empirical studies suggest that combining pricing and land use policies may lead to less driving, more transit use and walking, and lower VMTs than enacting either policy alone (Guo et al., 2011; Lee & Lee, 2003). Some studies of planning practices, such as combining HOT and Bus Rapid Transit (BRT), also report that HOT may be a promising strategy to increase public transit ridership, suggesting a need to connect the compact development with HOT projects (Brinkerhoff, 2009). Therefore, it is important to enrich recent literature by developing theories and empirics to justify incorporating land use and pricing policies.

SUMMARY

This chapter reviewed theories of market failures from the field of economics and planning failures from the area of planning and bridged them into a consistent framework for conceptualizing excessive congestion. The framework illuminates that excessive congestion is shaped by market failures from congestion and agglomeration externalities and planning failures from land use regulations. Excessive congestion occurs when the individually desirable amount of driving exceeds the socially desirable level at which the MSB of travel equals the MSC. It is important to measure optimal and excessive congestion by internalizing not only the external cost of congestion but also the external benefit of crowding activity and traffic from urban agglomeration. While little literature has discussed the benefit of congestion, the following chapter will investigate how economic agglomeration connects with congestion and how anti-congestion policies affect agglomeration economies.

More importantly, this chapter demonstrated that planning failure can play a dominant role leading to excessive congestion and social inefficiency. Planning failures from low-density and exclusionary zoning regulations could increase travel distance and auto dependence, produce excessive driving demand, and cause excessive traffic on the streets and highways. Differing from market failures, planning failures are insensitive to pricing signals and could be better corrected by regulatory reform or innovative land use planning. In most cases, both market and planning failures contribute to excessive congestion. Evaluating congestion-relief policies needs an innovative analytical framework able to internalize both failures. Therefore, Chapter 3 aims to develop such an innovative model to scrutinize how market and planning failures affect excessive congestion, land use patterns, economic performance, and social welfare. Relying on simulations, Chapter 3 will investigate the optimal policies for correcting both failures and reducing excessive congestion. Chapters 4–5 then extend Chapter 3's discussion to tackle more practical land use and pricing remedies for market and planning failures.

Many empirical studies for congestion relief have investigated land use planning and congestion pricing strategies. The land use–travel connection has been widely studied in the planning and transportation fields. This chapter reviewed and summarized three categories of land use planning's mechanism. First, land use planning strategies promoting less auto-oriented design can facilitate nonauto travel modes and reduce driving demand thus mitigating congestion. Second, land use planning strategies such as UGBs and urban densification policies can serve as alternative tools to adjust the cost of travel close to the socially desirable level, especially when pricing policies are not feasible. Third, land use planning for alternative development can meet the unmet demand for non–auto-oriented neighborhoods when existing land use regulations restrict such a development preference.

On the other hand, the reasoning underlying congestion pricing strategies is straightforward. They are primarily used to increase the cost of underpriced travel and adjust traffic volume to the socially desirable level. Most empirical studies have been concerned only with the congestion pricing's efficiency in the transportation market, while an increasing number of studies have turned to the congestion pricing's impact on land use and development.

However, less theoretical and empirical research has fully recognized that inefficient congestion is a consequence resulting from market and planning failures. There are only a limited number of empirical studies that have looked at the interaction of impact of land use and pricing policies. Therefore, Chapter 6 will present an empirical study to justify incorporating land use and pricing policies and substantiate some of the theoretical and simulation findings analyzed in Chapters 3 through 5.

CHAPTER 3: OPTIMAL POLICIES FOR REDUCING EXCESSIVE CONGESTION IN CITIES WITH MARKET AND PLANNING FAILURES

This chapter develops an analytical framework to internalize both market and planning failures, simulate how these failures cause excessive congestion, and examine the theoretical findings articulated in Chapter 2. This research aims to identify and evaluate the optimal policies for alleviating excessive congestion, changing land use patterns, and improving social efficiency.

Specifically, this chapter first develops a new spatial general equilibrium model with endogenously determined congestion and agglomeration externalities. This model examines optimal and excessive levels of congestion and the efficiency of first-best policies and other instruments, like simply Pigouvian congestion toll and simply Pigouvian labor subsidy. For the first-best interventions, this model investigates welfare gains and land use patterns in the social optimum along with the challenges to designing first-best instruments, because these topics are seldom discussed in cities with multiple externalities. The congestion diseconomy and welfare outcomes of the Pigouvian congestion toll and Pigouvian labor subsidy alone policies are compared. A robustness analysis is conducted by changing the congestion and agglomeration parameters to investigate how optimal policies and their welfare and land use outcomes vary with the levels of externalities.

Next the new model is extended to account for planning failures sourced from land use regulations, such as exclusionary zoning and low-density zoning. Extended simulations are thus applied to evaluate how market and planning failures together cause excessive congestion and to show evidence for incorporating land use and pricing policies.

LIMITATIONS OF URBAN ECONOMIC ANALYSES

Cities are full of externalities. The external costs of traffic congestion and the external benefits of firm agglomeration are widely discussed in urban economics literature. Congestion, for example, delays other travelers, adds air pollution and greenhouse gases, and raises a community's energy demands. Firm agglomeration economies can largely explain the geographical centralization of firms, as well as the emergence and evolution of cities. Firms benefit from locating close to each another, via access to intermediate inputs and labor, easier job-worker matching, knowledge spillovers, and other sources (Fujita and Thisse, 2002; Puga, 2010; Rosenthal and Strange, 2004). Such agglomeration externalities rise with the density of economic activities and proximity to other firms. As a result, doubling job density or doubling city size at the aggregate metropolitan level is often associated with a 4%-10% or 3%-8% increase, respectively, in productivity (Combes, Duranton, Gobillon, & Roux, 2010; Rosenthal and Strange, 2004). Some studies at the micro-geographical level (e.g., census tract) find an even larger agglomeration benefit that decay with distance (Arzaghi and Henderson, 2008; Rosenthal and Strange, 2008).

While urban economists have long recognized either negative congestion externalities (e.g., Solow 1972; Arnott, 1979; Pines and Sadka, 1985; Wheaton, 1998; Anas and Xu, 1999; Brueckner, 2007) or positive agglomeration externalities (Fujita and Ogawa, 1982; Lucas and Rossi-Hansberg [LRH], 2002; Berliant et al., 2002; Rossi-Hansberg, 2004; Borck and Wrede, 2009), few have considered their interactions. Incorporating both externalities in urban economic analysis is important, since urban policies for coping with one externality in one distorted market may neglect the spillover effects of this policy on the other distorted market. For example, a Pigouvian congestion tolling strategy charges marginal external costs to travelers who impose such costs, and is regarded as a first-best instrument for correcting distortions from negative congestion externalities. In isolation, this strategy is not first-best for cities, because tolls affect labor costs, land use patterns, and rents, and thereby affect agglomeration economies and firm productivity. By better understanding the interactions between congestion and agglomeration, one can avoid policy distortions informed by partial equilibrium analyses with only one externality, and thereby design more appropriate “first-best” policies while evaluating the benefits and limitations of second-best tolling, labor subsidies, and land use policies.

Few researchers have endogenized multiple urban externalities, and most rely on aspatial settings. For instance, Parry and Bento (1999) explored the interaction of distorted labor and transportation markets and evaluated the welfare effects of a congestion tax in the presence of a labor tax. They found that the congestion tax could reduce labor supply if total toll revenues are equally redistributed to residents, and stimulate labor supply if revenues are used to subsidize labor, with the latter form of revenue recycling generating more welfare improvement. Arnott (2007) developed a two-island model internalized both negative congestion and positive production externalities. In the simplified model, residents locate at an island and firms locate at the other island, with a road of fixed capacity crossing the two islands. He found that a Pigouvian congestion toll only is not the optimal policy since it may harm agglomeration economies and productivity. The optimal congestion tolls should be lower than the Pigouvian level when there is no policy in place to manage agglomeration externalities. He believes that these findings are consistent even though the model is extended to internalize time-varying congestion, heterogeneous individuals and/or firms, residential location and land decisions, and multiple employment centers. These two studies identify the policy importance on incorporating multiple externalities. However, they either neglect the spatial distribution of externalities or assume an exogenously determined urban form (e.g., two islands), failing to fully analyze the interaction between externalities and urban form, which may significantly affect the optimal design of urban policies.

Externalities affect urban form, and urban form affects externalities. Some models rely on discrete spatial settings to track multiple externalities. For example, Anas and Kim (1996) presented a spatial computable general equilibrium (spatial CGE) model integrating congestion and agglomeration externalities for consumers in a linear city) with discrete zones. Here, consumers are assumed to make more shopping trips to larger shopping centers (i.e., those exhibiting retail-job agglomerations). Their simulation results suggest that congestion externalities disperse urban form, while shopping agglomeration favors more compact forms, with fewer and more job-rich centers. Anas (2012) also recently developed a core-periphery model to explore social optima after first recognizing highway congestion's external costs and transit's external benefits, and then allowing for Marshallian agglomeration externalities. His comparative static analysis revealed that the optimal policy in a closed city with two or more externalities (or activities with economies of scale) should satisfy the general Henry George Theorem.

Other studies have internalized multiple spatial externalities by extending the traditional monocentric model. For example, Verhoef and Nijkamp (2004) modeled both agglomeration externalities (of firms) and pollution externalities (from commutes) under monocentric settings. They highlighted the importance of using a spatial equilibrium framework to understand urban externalities since congestion pricing and labor subsidies are not perfect (opposite) substitutes in the presence of spatial interactions. Their simulations show how second-best tolls or subsidies are lower than the Pigouvian levels. Wheaton (2004) combined a congestion externality and center-agglomeration forces into a circular monocentric framework, suggesting that worse congestion is associated with more centralized firm agglomeration. However, such monocentric models often do not internalize land inputs/rents in any production function; they rely on simplified, aspatial measures of agglomeration and thus overlook interactions between agglomeration externalities and urban form.

Therefore, the following section first develops and then applies a spatial general equilibrium model with endogenously determined congestion and agglomeration externalities in a continuous, non-monocentric city space. The agglomeration externality is a Marshallian production externality and defined to be proportional to each site's local jobs density and an integral of inverse-exponential distance-weighted job counts within a pre-existing cluster around the region's center point. This assumption pivots off those in Fujita and Ogawa (1982) and LRH (2002). Fujita and Ogawa (1982) were among the first to explore the economics of non-monocentric urban economies with production externalities, using a linear city form. Production externalities, or location potential (as defined in their paper), is reflected in firm productivity, which varies over space, thanks to clustering of economic activities. LRH (2002) extend the Fujita-Ogawa model to a continuous, circular city setting. Rossi-Hansberg (2004) then applied the LRH model to evaluate labor subsides and zoning restrictions, but without congestion externalities. Thus, the model developed here is among the first to incorporate Fujita-Ogawa- and LRH-type agglomeration economies and congestion externalities in a continuous urban space, enabling more comprehensive policy assessments.

A SPATIAL GENERAL EQUILIBRIUM MODEL WITH ENDOGENOUSLY DETERMINED CONGESTION AND AGGLOMERATION

The model developed here mainly refers to LRH (2002). While the LRH model has well established a nonmonocentic model with agglomeration externalities, the model here extends it to consider traffic congestion and contributes to the discussion of optimum versus equilibrium under congestion and agglomeration externalities.

Also, there are several differences with basic modeling settings. First, the model relaxes the constraint of fixed city boundary in the LRH model, allowing for an endogenously determined boundary under an additional constraint that the city edge land rent equals a fixed agriculture land rent. The latter constraint is often used in monocentric models (e.g., Wheaton, 1998; Brueckner, 2007). This change can internalize city size, which may affect the spatial distribution of land use and commute distance/cost. Second, while the LRH model measures commute time costs determined by travel time and wage, our model's measure is simplified to commute money costs determined only by distance (and traffic volume after considering congestion). In reality, the commute costs consist of the cost of time and money. Third, our model is built in a closed-form city with a fixed population and all revenues (or subsidies) uniformly redistributed to residents (or firms), while the LRH model is built in an open-form city with a fixed utility and without revenue redistribution. These changes increase the complexity of computational simulations, but make this type of nonmonocentric model more flexible for optimal policy analysis.

The model assumes a continuous symmetric circular region of radius $\bar{x}$. The symmetry assumption implies that workers travel only towards or away from the center, along radial street networks. Two homogeneous agent types, households and firms, exist and can reside at the same location inside the region. For any location $x(0 \leq x \leq \bar{x})$, $\theta_f(x)$, $\theta_h(x)$ and θ_t represent the fractions of land area used by firms, households, and transportation infrastructure. $\theta_f(x)$ and $\theta_h(x)$ are endogenously determined, while θ_t is exogenously given.

Household and Congestion Externality

Each household living in location x and working at location x_w consumes a quantity of goods $c(x, x_w)$ (with price p = 1) and enjoys a residential lot size $q(x, x_w)$, resulting in utility level $u\big(c(x, x_w), q(x, x_w)\big)$. Its willingness to pay for land is rental rate $r_h(x)$. Each household has one worker, earning net income $y(x, x_w)$. This net income is comprised of three components: wage income paid by firms at location x_w, $w(x_w)$, minus commuting costs $T(x, x_w)$, plus the return of aggregate rent and toll revenues, $\bar{y}$. Thus, the optimization problem of each household is as follows:

Problem 3.1 *For each household living at location x* $(0 < x \leq \bar{x})$*, choose a job location* x_w $(0 < x_w \leq \bar{x})$ *and evaluate functions* $c(x, x_w)$ *and* $q(x, x_w)$*, so as to maximize utility*

$$u\big(c(x, x_w), q(x, x_w)\big) \tag{3.1}$$

subject to the budget constraint:

$$c(x, x_w) + r_h(x)q(x, x_w) \leq y(x, x_w) = w(x_w) + \bar{y} - T(x, x_w) \tag{3.2}$$

where

$$\bar{y} = \frac{1}{N}\big(y_{rent} + y_{toll} - y_{suby}\big) \tag{3.3}$$

$$T(x, x_w) = \int_x^{x_w} (t(s) + \tau(s))ds \tag{3.4}$$

Eq. (3.3) guarantees that aggregate revenues from land rents y_{rent} and tolls y_{toll}, net of the labor subsidy y_{suby}, are uniformly distributed to households, consistent with a closed-form city of (given) population *N*. This setting allows one to compare more equitably the welfare effects of different policy scenarios. Eq. (3.4) shows that $T(x, x_w)$ is an accumulation of marginal travel costs, from x to x_w. Here, $t(x)$ represents the average travel cost per mile at location *x*, with a negative sign representing inward travel and a positive sign representing outward travel. $\tau(x)$ represents a potential congestion toll on drivers passing location *x*. Consistent with prior works (e.g., Brueckner, 2007; Wheaton, 1998, 2004), $t(x)$ is proportional to a power function of the traffic volume crossing the ring at *x*, $D(x)$, relative to the road supply or width at *x* – plus the free-flow travel-cost component, φ (in dollars per mile). Thus,

$$t(x) = \begin{cases} -\varphi - \rho\left(\frac{-D(x)}{2\pi x\theta_t}\right)^{\sigma} & if\ D(x) < 0 \\ \varphi + \rho\left(\frac{D(x)}{2\pi x\theta_t}\right)^{\sigma} & if\ D(x) > 0 \\ \varphi \quad or - \varphi & if\ D(x) = 0 \end{cases} \tag{3.5}$$

where ρ and σ ($\sigma \geq 1$) are positive parameters designed to reflect network congestibility (very much like the standard Bureau of Public Roads [BPR 1964] formulation for travel times). As with travel costs, traffic volumes, $D(x)$, are negative when flow is inward at location *x*, and positive when flows are outward. When $D(x) = 0$, no traffic crosses location *x*, and the marginal travel cost equals the free-flow cost (which can be either positive or positive).

Proposition 3.1: Suppose $c^*(x, x_w)$and $q^*(x, x_w)$ are the solutions to *Problem 1* and $\bar{u}$ is the maximized utility level; then, the following are true:

(a) For those households living in location *x*, regardless of where they work, they earn an identical *net* income, $y(x)$, so that: □$(x, x_w) \equiv y(x),\ \forall\ x_w > 0$; and they consume the same amount of goods and lot size, $c^*(x)$ and $q^*(x)$, so that: $c^*(x, x_w) \equiv c^*(x)$ and $q^*(x, x_w) \equiv q^*(x), \forall\ x_w > 0$.

(b) $q^*(x) = q^*(y(x), \bar{u})$ and $c^*(x) = c^*(y(x), \bar{u})$ satisfy the equations $c(x) + q(x)u_q/u_c = y(x)$ and $u\big(c(x), q(x)\big) = \bar{u}$;

(c) $y(x) = w(x) + \bar{y}$; and

(d) $y'(x) = w'(x) = t(x) + \tau(x)$.

Proof. See A1 in the Appendix.

From Proposition 3.1a, household attributes at location *x*, including $c(x, x_w), q(x, x_w)$, and $y(x, x_w)$, can be written simply as $c(x), q(x)$, and $y(x)$ in the rest of this article. From Proposition 3.1b, if one assumes a Cobb-Douglas utility function, as follows:

$$u\big(c(x), q(x)\big) = c(x)^{\alpha} q(x)^{1-\alpha}, \ 0 < \alpha < 1 \tag{3.6}$$

then, the solutions to Problem 1 are:

$$q^*(x) = \alpha^{-\alpha/(1-\alpha)} y(x)^{-\alpha/(1-\alpha)} \bar{u}^{1/(1-\alpha)} \tag{3.7}$$

$$c^*(x) = \alpha y(x) \tag{3.8}$$

and maximized bid-rents from households are:

$$r_h^m(x) = (1-\alpha)\alpha^{\alpha/(1-\alpha)} \left(\frac{y(x)}{\bar{u}}\right)^{1/(1-\alpha)} \tag{3.9}$$

Equations (3.7) to (3.9) show that optimal lot size and good consumption and maximum bid-rent at location x are determined by household's net income, $y(x)$, which relates to wages earned and commuting costs, as shown in Eq. (3.1). Proposition 3.1c demonstrates that the net income of households residing at x equals the wage income paid by firms at x plus redistributed revenues. From Proposition 3.1d, the condition that both the wage gradient and the net-income gradient equal the marginal travel cost should be satisfied when maximizing utilities. This condition supports the intuition that no worker can achieve a higher net income (net of commute costs, plus labor subsidies or toll revenue redistributions) by changing his or her job location.

Firms and Agglomeration Externalities

Each firm is a price taker in input and output markets. If a competitive firm located at x operates under constant returns to scale, its total production $P(x)$ depends on the amounts of labor $L(x)$ and land area $H(x)$ used, and its total factor productivity (TFP) $A(x)$, such that:

$$P(x) = A(x)L(x)^{\kappa}H(x)^{1-\kappa} \; (0 < \kappa < 1) \tag{3.10}$$

The production per unit of land, $p(x)$, is therefore as follows:

$$p(x) = \frac{P(x)}{H(x)} = A(x)n(x)^{\kappa} \tag{3.11}$$

where $n(x)$ is labor density along ring x and κ is the production function's elasticity parameter. One can internalize agglomeration economies in the TFP, by assuming that the agglomeration externality $F(x)$ at location x determines the productivity:

$$A(x) = \delta F(x)^{\gamma} \; (\delta > 0, 0 < \gamma < 1) \tag{3.12}$$

Here, δ is the productivity scale parameter, and γ is the elasticity of productivity with respect to agglomeration externalities at location x. Fujita and Ogawa (1982) provided a measure of agglomeration economies for firms based on location potential in a linear city setting: they used job densities and distances to other firms or workers. LRH (2002) extended this measurement to circular space [5]. Similar to LRH's setting, agglomeration externalities are defined here to be proportional to the local employment density (at location *x*) and the integral of an inverse-exponential distance-weighted job

[5] One can set a more general formation of the agglomeration externality function, for example:

$$F(x) = \int_0^{\bar{\bar{x}}} b(r)d(r,x)dr$$

Here, $b(r)$ represents the density of firms or workers at location r. $d(r,x)$ is a distance-based decay function from location *r* to *x*. Two specifications of $d(r,x)$ are widely used. For example, in a linear city, $d(r,x)$ could be a linear form, $1 - \phi|r - x|$ (e.g., Ogawa and Fujita, 1980; Duranton and Puga, 2014), or an inverse-exponential form, $e^{-\phi|r-x|}$ (e.g., Fujita and Ogawa, 1982). These two formations are equivalent when $\phi|r - x|$ is small enough. In simulation experiments in this dissertation, I compared the results using the two types of externality specifications, finding that these two specifications do not bring substantial difference in modeling results (e.g., land use and welfare outcomes). These findings also correspond to those in the linear model (e.g., by comparing Ogawa and Fujita [1980] and Fujita and Ogawa [1982]). Thus, the following discussions only depend on the inverse-exponential specification.

count within the city boundary[6]. Thus, the agglomeration externality at each location along the annulus at radius x is specified as

$$F(x) = \zeta \int_0^{\bar{x}} \int_0^{2\pi} r\theta_f(r)n(r)e^{-\zeta l(x,r,\psi)} d\psi dr \qquad (3.13)$$

where ζ is the production externality scale parameter, and is exogenously determined. ψ is the polar angle around the center (ranging from $\mathbf{0}$ to 2π), and $l(x,r,\psi)$ is the straight-line distance between a firm at a specific location along annulus x and each firm lying within $\bar{x}$ miles of the center (at a counter-clockwise angle of ψ from the first firm). Thus,

$$l(x,r,\psi) = \sqrt{x^2 + r^2 - 2xrcos(\psi)} \qquad (3.14)$$

The firms then maximize the profit function with respect to employment density $n(x)$, with firm output price set at 1 (without loss of generality):

$$\boldsymbol{Max}\, \pi(n(x)) = \delta n(x)^{\kappa} F(x)^{\gamma} - n(x)\big(w(x) - s(x)\big) - r_f(x) \qquad (3.15)$$

where $s(x)$ represents a potential labor subsidy for firms at location x to hire each worker.

From the first-order condition of profit maximization with respect to $n(x)$, one can obtain optimal employment density at location x as follows:

$$n^*(x) = \left(\frac{\kappa\delta F(x)^{\gamma}}{w(x) - s(x)}\right)^{1/(1-\kappa)} \qquad (3.16)$$

[6] LRH's model sets a fixed-boundary assumption while our model estimates an endogenous $\bar{x}$ under the constraint of edge land rent. This change can endogenize city size. Zhang and Kockelman (2014) use a similar measure but assume that production externalities come only from firms within a pre-existing cluster around the region's center point, up to an (exogenously set) boundary distance of $\bar{r}$. This assumption allows for modeling a city system with larger decentralized forces (e.g., congestion diseconomies) than centralized forces (e.g., agglomeration economies). Without such an assumption, an equilibrium city always appears has a larger centripetal force. But this setting may constrain the emergence of polycentricity (see next chapter's discussions).The model here thus relaxes such a constraint.

Given perfectly competitive input and output markets, all firms make zero (excess) profit, with land rents rising to their maximum values to ensure this, as follows:

$$r_f^m(x) = (1-\kappa)\delta^{1/(1-\kappa)}F(x)^{\gamma/(1-\kappa)}\left(\frac{\kappa}{w(x)-s(x)}\right)^{\kappa/(1-\kappa)} \quad (3.17)$$

The Land Market's Equilibrium Conditions

Since both firms and households can exist in the same location, a competitive market requires they bid for the land via their willingness to pay (or maximum bid rents). Given the maximized bid-rents from the partial equilibrium of households and firms at each location x (as shown in Eqs. (3.9) and (3.17)), the land market equilibrium requires that land rents, $r(x)$, satisfy the following two equations:

$$r(x) = max\{r_h^m(x), r_f^m(x), R_a\} \quad (3.18)$$

$$r(\bar{x}) = R_a \quad (3.19)$$

Eq. (3.19) defines the edge land rent $r(\bar{x})$, which equals the agricultural land rent (or opportunity rent) R_a. If both $r_f^m(x)$ and $r_h^m(x)$ are less than i, the equilibrium land use share for firms $\theta_f^*(x)$ and the equilibrium land use share for household $\theta_h^*(x)$ will equal zero. If $r_f^*(x)$ equals $r_h^*(x)$, a mixed land use pattern will emerge at location x, and the equilibrium number of jobs at that location will equal the number of households (or residing workers) at that location (LRH, 2002). Given that both $r_f^m(x)$ and $r_h^m(x)$ will exceed R_a (except at the developed region's edge), $\theta_f^*(x)$ and $\theta_h^*(x)$ at each location *x* are as follows:

$$\theta_f^*(x) = \begin{cases} 1-\theta_t & if\ r_f^m(x) > r_h^m(x) \\ \dfrac{n^*(x)q^*(x)}{n^*(x)q^*(x)+q^*(x)}(1-\theta_t) & if\ r_f^m(x) = r_h^m(x) \\ 0 & if\ r_f^m(x) < r_h^m(x) \end{cases} \quad (3.20)$$

$$\theta_h^*(x) = 1-\theta_t-\theta_f^*(x) \quad (3.21)$$

Eq. (3.21) represents the land market clearing so that all available land or properties are assigned to either firms/jobs, households, or transport infrastructure. Moreover, total city/region land rents (net of the base rent, R_a), y_{rent}, in a spatial equilibrium will satisfy the following equation:

$$y_{rent} = \int_0^{\bar{x}} 2\pi x \left\{ \theta_f^*(x)\left(r_f^m(x) - R_a\right) + \theta_h^*(x)(r_h^m(x) - R_a) \right\} dx \qquad (3.22)$$

The Labor Market's Equilibrium Conditions

Under equilibrium, the commute demand generated in the interval dx from x to $x+dx$ (or absorbed in dx from $x+dx$ to x), $D'(x)dx$ (or $-D'(x)dx$), will equal the number of workers who need to work outside the interval (or the job vacancies in dx)[7]. Thus,

$$D'(x) = 2\pi x \left(\frac{\theta_h^*(x)}{q^*(x)} - \theta_f^*(x) n^*(x) \right) \qquad (3.23)$$

A spatial equilibrium requires that travel demand at the city edge, $D(\bar{x})$, and in the city center point, $D(0)$, equals zero (since there are no jobs or workers beyond this boundary, to attract or generate such trips). Thus, the two boundary conditions for commute demand are:

$$D(0) = 0 \;\; and \;\; D(\bar{x}) = 0 \qquad (3.24)$$

These two boundary constraints also guarantee the second condition for labor market clearing: the total number of workers will equal the number of households, *N*:

$$\int_0^{\bar{x}} 2\pi x \frac{\theta_h^*(x)}{q^*(x)} dx = \int_0^{\bar{x}} 2\pi x \theta_f^*(x) n^*(x) dx = N \qquad (3.25)$$

[7] Here, households living and working at the same location x are assumed to generate no commute. The setting of Eq.(23) refers to Wheaton (2004) and can be comparable with the LRH(2002)'s model. The LRH paper explains *D*(*x*) (labeled as *H*(*x*)) as the stock (work hour) of unhoused workers at *x*. Since the LRH model measures commute costs using travel time and the total time for working and commuting is fixed, the changed stock of unhoused workers from *x* to *x+dx* (or *x-dx*) include two parts. The first part is the net number of unhoused workers in the interval *d*x. Another part is the lost work hours due to passing the interval. The second part is not included in our model, since our model only considers the distance-based commute money costs and no work hours are lost due to commuting change.

Spatial General Equilibrium

One can combine households' and firms' partial equilibria with equilibrium conditions for labor and land markets, thereby creating a spatial general equilibrium model for the region. Given $\bar{u}$ and other parameters, this model has 20 unknowns, including 15 functions of x: $c^*(x), q^*(x), r_h^m(x), y(x), t(x), \tau(x), D(x), w(x), n^*(x)$, $r_f^m(x), s(x), F(x), r(x), \theta_h^*(x), \theta_f^*(x)$, and 5 scalars: $\bar{x}, \bar{y}, y_{rent}, y_{toll}, y_{suby}$. 20 equations are needed to resolve this model, including 16 equations described above (Eqs. (3.2) and (3.4), Proposition 3.1(c) and (d), Eqs. (3.7)-(3.9), (3.13), and (3.16)-(3.23)) plus 4 other equations that define the tolling instrument, $\tau(x)$ and y_{toll}, and the subsidy, $s(x)$ and y_{suby}, which vary across policy scenarios.

Table 3.1 summarizes these four functions, $\tau(x)$, $s(x)$, y_{toll}, and y_{suby}, across six spatial equilibria. In the free-market equilibrium, neither a toll nor a subsidy is imposed, so $\tau(x) = 0, s(x) = 0, y_{toll} = 0, and\ y_{suby} = 0$. Given the simultaneous existence of two externalities in the model, a free-market equilibrium is inefficient; thoughtful policy intervention is needed to cope with market inefficiency. As noted earlier, four types of intervention are considered here: the simultaneous application of two first-best instruments, application of just Pigouvian congestion toll, and application of just Pigouvian labor subsidy.

Table 3.1 Policy Instrument Values $\{\tau(x), s(x), y_{toll}, y_{suby}\}$ for Urban Equilibria under Four Policy Interventions

Policy Interventions	Equations
Free-Market	$\tau(x) = 0; s(x) = 0; y_{toll} = 0; y_{suby} = 0$
First-Best	$\tau(x) = \tau_{pct}(x)$; $s(x) = s_{pls}(x)$; $y_{toll} = \int_0^{\bar{x}} \tau(x)D(x)dx$; $y_{suby} = \int_0^{\bar{x}} 2\pi x\theta_f^*(x)n^*(x)s(x)dx$.
Pigouvian Congestion Toll Alone (PCT-Alone)	$\tau(x) = \tau_{pct}(x)$; $s(x) = 0; y_{toll} = \int_0^{\bar{x}} \tau(x)D(x)dx$; $y_{suby} = 0$.
Pigouvian Labor Subsidy Alone (PLS-Alone)	$\tau(x) = 0; s(x) = s_{pls}(x); y_{toll} = 0$; $y_{suby} = \int_0^{\bar{x}} 2\pi x\theta_f^*(x)n^*(x)s(x)dx$.

Proposition 3.2: First-best instruments to correct congestion and agglomeration externalities satisfy either one of following conditions:

(a) A first-best combination of the *Pigouvian Congestion Toll* $\tau_{pct}(x)$ at each location *x* and the *Pigouvian Labor Subsidy* $s_{plb}(x)$ on every unit of labor supplied at each firm location *x* can be defined as follows:

$$\tau_{pct}(x) = t'\big(D(x)\big)D(x) = \begin{cases} \rho\sigma\left(\dfrac{|D(x)|}{2\pi x\theta_t}\right)^{\sigma}, & if\ D(x) \geq 0 \\ -\rho\sigma\left(\dfrac{|D(x)|}{2\pi x\theta_t}\right)^{\sigma}, & if\ D(x) \leq 0 \end{cases} \tag{3.26}$$

$$\begin{aligned} & s_{pls}(x) \\ &= \int_0^{\bar{x}} \frac{d\left(2\pi r\theta_f^*(r)p(r)\right)}{d\big(2\pi x\theta_f^*(x)n^*(x)dx\big)} dr \\ &= \begin{cases} \gamma\kappa\delta\zeta \displaystyle\int_0^{\bar{x}}\int_0^{2\pi} r\theta_f^*(r)n^*(r)F(r)^{\gamma-1}e^{-\zeta l(x,r,\psi)}\, d\psi dr, & if\ \theta_f^*(x) > 0 \\ 0, & if\ \theta_f^*(x) = 0 \end{cases} \end{aligned} \tag{3.27}$$

(b) *First-best road tolling* for each mile driven at each location *x*, $\tau_{fb}(x)$, is as follows:

$$\tau_{fb}(x) = \begin{cases} \tau_{pct}(x), & if\ \theta_f^*(x) = 0 \\ \tau_{pct}(x) - \dfrac{\partial s_{pls}(x)}{\partial x}, & if\ \theta_f^*(x) = 0 \end{cases} \tag{3.28}$$

and the revenue generated by optimal tolls equals the aggregate congestion externality costs minus the aggregate agglomeration externality benefits.

(c) *First-best labor subsidy* on every worker who lives at x_i and works at x, $s_{fb}(x_i, x)$ will be as follows:

$$s_{fb}(x_i, x) = s_{pls}(x_i, x) - \int_{x_i}^{x} \tau_{pct}(r)\, dr \tag{3.29}$$

and the aggregate optimal subsidy equals the aggregate agglomeration externality benefits minus the aggregate congestion externality costs.

Proof. See A2 in the Appendix.

In the socially optimal city, markets failures from both congestion and agglomeration externalities are needed to be corrected by first-best instruments. As noted in Proposition 3.2, the social optimum can be achieved via three types of first-best instrument. The city can simultaneously impose Pigouvian congestion toll and Pigouvian labor subsidy, both of which equal corresponding marginal externalities, as shown in Eqs. (3.26) and (3.27). The marginal congestion externality (MCE) at each x equals $t'\big(D(x)\big)D(x)$. Intuitionally, the derivative of $t(x)$ of $D(x)$ represents the added marginal travel cost on each individual driver across x when one new driver is added, while $\tau_{pct}(x)$ represents total additional travel costs (imposed on other drivers), as caused by the added driver. The marginal external benefits by hiring additional workers at location x, $d\big(2\pi x\theta_f^*(x)n^*(x)dx\big)$, equals the total gain aggregating marginal output at other locations r $(0 \leq r \leq \bar{x})$, $d\left(2\pi x\theta_f^*(r)p(r)\right)$.

The city can also impose first-best tolls by internalizing external benefits of agglomeration into Pigouvian congestion toll levels. Proposition 3.2b suggests that the first-best tolls largely vary with locations. They should be set at corresponding Pigouvian levels in residential areas but not within-firm clusters. After considering the impact on agglomeration economies, the optimal tolling could be positive or negative (i.e., a subsidy), depending on the locational margin of agglomeration benefits, $s'_{pls}(x)$. In addition, the aggregate optimal toll should lie below the aggregate congestion externality cost. This finding is consistent with Arnott's (2007) result for a relatively straightforward, non-spatial model, where the optimal toll is lower than congestion externality cost and even negative, if the agglomeration externality cannot be subsided. Similarly, when congestion tolls are not feasible (e.g., they may not be politically acceptable), the city can supply first-best subsidies to firms, and the total optimal subsidy will then lie below the total agglomeration benefit. But Proposition 3.2c suggests that such an optimal labor subsidy will be very complicated, since it varies with not only firms' locations but also worker's residence.

Both the Pigouvian congestion toll-alone and Pigouvian labor subsidy-alone policy instruments are not first-best strategies since they only correct one externality. Analytical equilibrium results are very difficult to compute here, for a 20-equation system with several non-linear equations and differential equations. The following section relies on numerical results, to compare the properties of the free-market, first-best and second-best equilibrium settings, by setting function values for $\{\tau(x), s(x), y_{toll}, y_{suby}\}$.

PARAMETER SETTINGS IN NUMERICAL SIMULATION

This chapter simulates an abstract circular, close-form city, where the number of workers N is fixed at 600,000 and the edge agricultural land rent Ra is set to $4,000,000 per square mile per year. This comes from the assumption that farmland at the edge of a city sells for about $50,000 per acre, with amortization of such costs over 40 years at a discount rate of 5% resulting in rural land rents over $4,000,000 per square mile per year.

Table 3.2 shows the parameter values of the base scenario8. Parameters of Cobb-Douglas utility and production functions rely on LRH's (2002) assumptions, where $\alpha = 0.90$ and $\kappa = 0.95$. The agglomeration parameters γ and ζ are set at 0.06 and 2, which are well in line with the empirical estimates ranging from 0.04 to 0.10 (Combes et al., 2010). The constant part of total factor productivity, δ, is set at 30,000, by calibrating Eq. (3.16) under the assumption that per-capita money income is $30,000 (per year) and the city center holds over 100 persons per acre, on average. Following Wheaton's (1998) study, roadways' share of land is assumed to be 30%. The intercept parameter φ in

[8] While calibrating a realistic city using empirical data under the model framework developed here is possible and important, it is not a major focus of this paper. Some calibration examples can refer to several studies relying on monocentric models (e.g., De Lara et al., 2013; Rappaport, 2014) and non-monocentric models (e.g., Brinkman, 2013).

Equation (16)'s average travel cost function represents an average cost of free-flow travel, and is set at $20 dollar per mile per year. This figure is generated from the calculation that marginal free-flow travel cost is about $0.04 per mile when each worker works about 250 days a year. ρ and σ reflect link congestibility, and are set as 0.00001 and 1.5, respectively. In a highly congested location, for example, if there are 50,000 travelers passing a point $x = 1$ mile from the region's center, the marginal congestion cost at $x = 1$ will be $0.17 per vehicle-mile, accounting for about 30% of total marginal costs. In a lightly congested location, say 5,000 travelers per day at a distance $x = 10$ miles away, the marginal congestion cost will account for only 0.4% of total marginal social costs (MSCs) at that point in the network.

Table 3.2 Parameter Value Assumptions in the Base Scenario

N	R_a	α	κ	γ	ζ	δ	θ_t	λ	φ	ρ	σ
600,000	$4M/sq.mi	0.9	0.95	0.06	2	30,000	0.3	0.5	20	0.00001	1.5

COMPUTATIONAL SOLUTIONS: A NESTED FIXED-POINT ALGORITHM

To iteratively solve for location-specific values, one can first divide the circular city into discrete, narrow rings, each of width Δx (e.g., $\Delta x = 0.01\ mile$). Each location x can be labeled as $x_i = i\Delta x\ (with\ i = 1,2, \dots, I)$, with x_1 representing the city center and x_I representing the city's boundary $\bar{x}$. According to the boundary condition in Eq.(3.24), the commute traffic demand for both locations $D(x_1)$ and $D(x_I)$ equals zero.

The spatial equilibria are solved by a nested fixed-point algorithm (three loops) using MATLAB. The inner part of the algorithm refers to LRH's fixed-point algorithm (2002) for finding the fixed points of the agglomeration function $F(x)$. Meanwhile the middle loop of the algorithm is applied to find the fixed points of the redistributed revenue $\bar{y}$. Notice that the boundary conditions in simulations differ from those in LRH's models. While the LRH's simulation assumes a fixed utility level and city boundary, this simulation assumes a fixed population and edge land rent. In addition, the outer part of the algorithm is applied to find fixed points of the land share function $\theta_f(x)$. Detailed algorithms are described in A3 in the appendix.

LRH (2002) provided a strict proof of the existence of a set of equilibrium solutions under certain assumptions of utility and production functions (e.g., when these functions are of Cobb-Douglas form). Rossi-Hansberg (2004) also provided a proof of a set of optimal solutions in his extension of LRH's model to include agglomeration externalities. The substantial difference of our model from LRH's and Rossi-Hansberg's models is the inclusion of congestion externalities and governmental wealth redistribution (i.e., rents, tolls, and subsidies). Especially in the wealth redistribution process, simulations require proofs of whether there are fixed points of y_{rent}, y_{toll}, and y_{suby}. Instead of providing complicated and elusive analytical proofs, this model is solved computationally, so if an equilibrium can be computed, it exists. This is true for all models of this genre, such as Fujita-Ogawa (1982), Anas-Kim (1996), Brueckner (2007), etc. Simulation results suggest that there exists a set of equilibria or optimal solutions if the parameters are appropriately selected.

In addition, for checking the existence of multiple equilibria, simulations in this research use several different initial functions of $\theta_f(x)$, $F(x)$, and $\bar{y}$. If an equilibrium or optimum solution exists, it is the unique one within a family of urban configurations. If multiple families of urban configurations exist under the same set of parameters, the equilibrium solution generating the maximum utility is chosen as the *Pareto*-optimal one. For an example, refer to A4 in the appendix.

PRICING POLICIES FOR ACHIEVING OPTIMAL CONGESTION LEVELS

This section examines anti-congestion, welfare, and land use effects of three policy instruments, comparing to those in the free-market equilibrium. The policies include first-best instruments correcting both congestion and agglomeration externalities, the Pigouvian congestion toll-alone instrument only fully correcting congestion externalities, and the Pigouvian labor subsidy-only instrument only fully correcting production externalities. These policies are first investigated in *the base scenarios* with parameters in Table 3.2, and thus in cities with varying agglomeration scales (by changing γ from 0.04 to 0.08) and congestion levels (by changing ρ from 0.000005 to 0.0001) in the next section.

Table 3.3 Simulated Results of Policy Scenarios

	Free Market	First Best	PCT-Alone	PLS-Alone
Utility Level, $\bar{u}$	5242	5258	5221	5225
Avg. CV (relative to the FM case, $/hh./year)		113.86	-152.26	-118.38
City Boundary, $\bar{x}$ (miles)	15.57	15.26	16.10	14.94
Tolls, $\bar{y}_{toll}$ ($/hh./year)	0	557	190	0
Subsidy, $\bar{y}_{suby}$ ($/wk./year r)	0	2057	0	2081
Rent Revenues Returned, $\bar{y}_{rent}$ ($/hh./year)	1494	1648	1218	1785
Avg. Commute Distance (miles/day)	8.16	6.37	4.79	8.73
Avg. Traffic (1000 vehicles/hr per section dx*dx), T	28.6	18.5	10.5	43.9
Negative Congestion Externalities (million $/year)	583	334	114	1196
Total Congestion Cost Benchmarked by the Free-Flow Cost ($million/year)	971	223	76	1993
Average TFP (compared to the constant)	1.809	1.816	1.743	1.905
Agglomeration Externalities (million $/year)	1235	1235	1225	1249
Avg. Labor Density (workers/sq. mile)	10510	11985	5879	28668
Avg. Residential Density (hhs/sq.mi.)	1260	1299	1246	1305
Avg. Rent for Firms (times Ra)	4.52	5.12	2.50	12.45
Avg. Rent for Housing(times Ra)	1.05	1.09	1.06	1.06
Avg. Labor Wage ($/year)	32660	30382	32260	30912
Avg. Net Income ($/year)	33266	33527	33176	33193

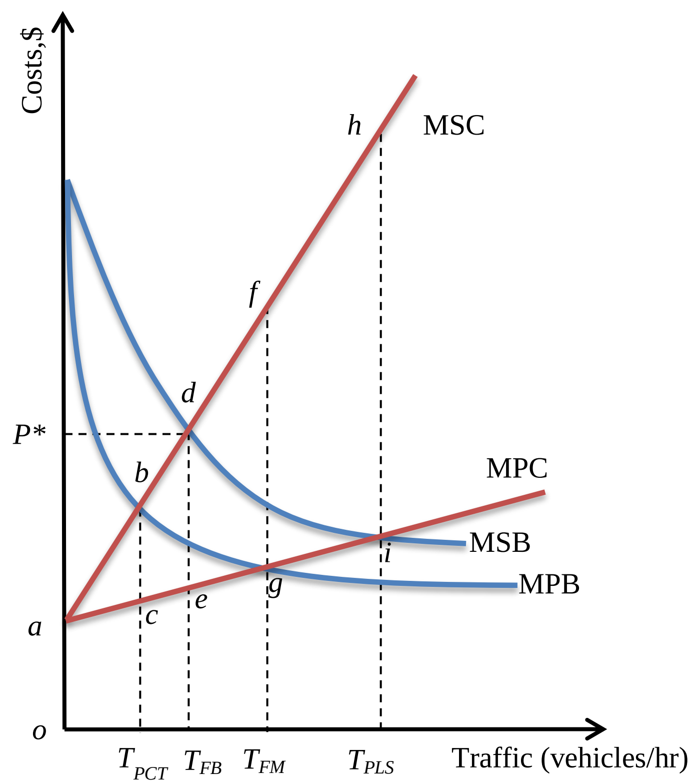
Table 3.3 shows major characteristics of urban equilibria under four policy schemes in the base scenario. In the free-market equilibrium, the utility level is 5242 and

the city's boundary is 15.58 miles away from the city center when the edge land rent equals the agricultural land rent, R_a.

Optimal and Excessive Levels of Congestion in the Base Scenarios

According to Table 3.3, the average daily commute distance in the free-market case is 8.16 miles, and the average traffic volume passing a road section is 28,600 vehicles per hour. These produce about 583 million dollars per year of congestion externalities impairing the society. First-best policies can generate lower total congestion externalities, i.e., 334 million dollars per year, a 43% decrease compared to the free-market level. In the social optimum, all negative externalities to the society are compensated by the tolling revenue. The average commute distance and traffic volume decrease to 6.37 miles per day (a 22% fall) and 18,500 vehicles per hour (a 35% fall). After fully correcting congestion externalities, the Pigouvian congestion toll-alone policy generates a much lower travel demand and lower congestion levels compared to the free market. The average commute distance and traffic volume fall by 41% (4.8 miles per day) and 63% (10,500 vehicles per hour). These traffic conditions are even better than those under first-best interventions. In contrast, the Pigouvian labor subsidy-alone policy fully corrects agglomeration externalities, but causes worse congestion. This includes longer commute distance and larger traffic volume than the free-market levels, and up to double the total congestion externality.

If congestion relief and VMT reduction effects are the primary objectives, the Pigouvian congestion toll-alone policy appears to be the most efficient. However, an economically healthy city probably demands more: not only less congestion, but also more agglomeration. Simulation results here suggest that anti-congestion policies can reduce negative externalities but at the same time, erode agglomeration benefits. The optimal level of congestion should balance congestion diseconomies and agglomeration economies, both of which are caused by the spatial concentration of activities. Figure 3.1 presents a diagram showing schematic curves of marginal benefits and costs of vehicle traffic in the four base scenarios. This diagram shows a schematic framework rather than mimicking the realistic simulation outcomes because the latter vary largely with location. However, this analytical diagram can be regarded as the average consequences of the four base scenarios.



Notes: MSC is marginal social costs of traffic; MPC is marginal private costs. The gaps between MSC and MPC are marginal external costs of congestion that are not priced in the free market. MPB is marginal private benefits of travel; MSB is marginal social benefits. The gaps between MSB and MPB are marginal external benefits of agglomeration that are not priced in the free market. T_{PCT} is the equilibrium traffic volume of the Pigouvian congestion toll-alone case; T_{FB} is the equilibrium traffic of the first-best optimum; T_{FM} is the free-market equilibrium level of traffic; and T_{PLS} is the equilibrium traffic of the PLS-alone case.

Figure 3.1 Excessive Congestion Caused by Free-Market, Pigouvian Congestion Toll Alone, and Pigouvian Labor Subsidy Alone Instruments in the Base Scenarios.

As shown in Figure 3.1, the free-market equilibrium in the base scenarios occurs when the MPC equals the MPB and the equilibrium traffic volume is T_{FM}, i.e., 28,600 vehicles per hour. The social optimum lies at the point of intersection of the MSC and MSB curves, and the optimal traffic level is T_{FM}, i.e., 18,500 vehicles per hour. Therefore, on average, the excessive driving demand in the free market is 10,100 vehicles per hour. The excessive congestion leads to a diseconomy in which each household loses about $114 annually (Table 3.3). However, this diseconomy of congestion imposed on each household is much smaller (only 7%) than $1,618, which is "the total cost of congestion" including internal and external costs. The evaluation of transportation projects or congestion relief policies in practice relies largely on the estimation of "the total cost of congestion" (e.g., Grant-Muller & Laird, 2007; OECD, 2007; Bilbao-Ubillos, 2008; Litman, 2009). Our findings suggest that it is socially inefficient to reduce all of "the total cost of congestion," and policies targeting the free-flow speeds could erode agglomeration economies and cause substantial welfare loss. As stated by Goodwin, "The 'total cost of congestion' is a large number, but it is practically meaningless and by 'devaluing the currency' it distracts attention from more important, achievable, objectives" (2004, p. 3).

This research suggests that anti-congestion policies should aim to reduce the net social cost of excessive congestion, i.e., the congestion diseconomy rather than "the total cost of congestion." The application of "the total cost of congestion" may overestimate the damage of congestion and overrate some anti-congestion policies (OECD, 2007). In the base scenarios, about 93% of the total cost of congestion is necessary for guaranteeing agglomeration economies. Because free-flow traffic is probably never a desirable outcome of social efficiency, the evaluation of anti-congestion policies should not assume that the free-flow level of traffic is the socially optimal level. Policies aimed at the socially optimal level of congestion should balance the reduction effects on both excessive congestion and agglomeration economies.

This research also demonstrates the importance of integrating all potential externalities affecting the transportation market in congestion-relief studies. Much of the literature has recognized the negative congestion externality only as a primary market failure causing excessive congestion (OECD, 2007). After accounting for agglomeration externalities, the Pigouvian congestion toll-alone policy is no longer socially optimal because the congestion toll raises the travel price too much and thus reduces travel demand too much (e.g., 8,000 vehicles fewer than the socially optimal level). This pricing-oriented "restriction" on travel demand could lower the average traffic volume and the level of spatial concentration of activities and encourage dispersal distribution of firms for less crowding, and thus erode agglomeration economies. Finally, the Pigouvian congestion toll-alone policy incurs a greater total cost from the loss of the agglomeration economy than the total benefit it derives from congestion reduction, still leading to a net social loss.

First-Best Policies in the Base Scenarios

According to Proposition 3.2, there are three first-best interventions – a combination of Pigouvian congestion toll and Pigouvian labor subsidy, a first-best congestion toll (that varies by road location), and a first-best labor subsidy (that varies by firm or job location) – and these first-best instruments can each produce the same social optimum. This research uses the combination of Pigouvian congestion toll and Pigouvian labor subsidy to simulate the optimum. Results show that under the social optimum, the city need to impose an average toll of $557 per commuter per year while delivering an annual average labor subsidy of $2057 per job position (Table 3.3). This result does not imply that a combined, equivalent tax of $1500 (i.e., $2057-$557) on each worker will achieve the first-best optimum: spatial variations in tolls and labor subsidies need to be considered.

Figure 3.2a-b shows the corresponding toll and/or subsidy levels across locations in the social optimum. Under this combination instrument, as job densities (or travel flows) increase, the amount of optimal labor subsidy (or optimal tolling) rises. Within the firm cluster area increases from 2.85 mile to 5.55 mile in radius, subsidies increase from about \$1198 to \$2337 per year at the locations of peak labor density, and then fall to \$1064 per year at the other edge of the firm cluster. Congestion tolls peak at the two ends of the firm cluster area, since these two places accumulate of the highest levels of outward and inward commute flows, generating the largest marginal negative externalities. Social optimum can be achieved by levying an optimal toll after internalizing agglomeration externalities. Figure 3.2c shows that the first-best toll equals the Pigouvian congestion toll in the residential areas, but varies quite a bit within the annulus of jobs, consistent with Proposition 3.2. The optimal toll levels across locations in the firm cluster area lie below the Pigouvian congestion toll and even become negative (thereby incentivizing such travel). These findings extend Arnott's (2007) aspatial analytical discussion, underscoring the importance of enabling spatial variation in policy interventions, in order to optimally address urban externalities.

Welfare improvement is visible under the first-best instruments. The utility level increases from 5242 to 5258, so it appears to be just 0.3% higher than that of the free-market equilibrium (Table 3.3). However, utils are only ordinal in nature; the average worker's willingness to pay to live in this optimally managed city, versus the free-market setting, is \$114 per year (as a compensating variation[9]). This welfare gain comes from the

[9] Given the utility levels are u_0 in the free-market case and u under a specific policy scheme, the average compensating variation is simply calculated as $\mathrm{CV} = \frac{1}{N}\int_0^{\bar{x}} 2\pi \frac{\theta_h(x)}{q(x)}(u-u_0)\frac{dy(u,q(x))}{du}dx$. $\frac{u}{c}\frac{dy}{du}$ represents the

benefit of excessive congestion reduction. When congestion externalities are internalized, the average commute costs rise from \$0.27 to \$0.57 per mile per day, leading to a decrease in travel demand and commute distance (which falls by 22%). With the Pigouvian labor subsidy, firms can hire workers by lower wage (the average wage drops by 7%), equaling social marginal costs that varies with locations (Figure 3.3a). The TFPs, however, in most job locations significantly improve (Figure 3.3b), with average TFP rising 0.42%. These findings suggest that first-best instruments simultaneously reduce congestion and enhance agglomeration benefits.

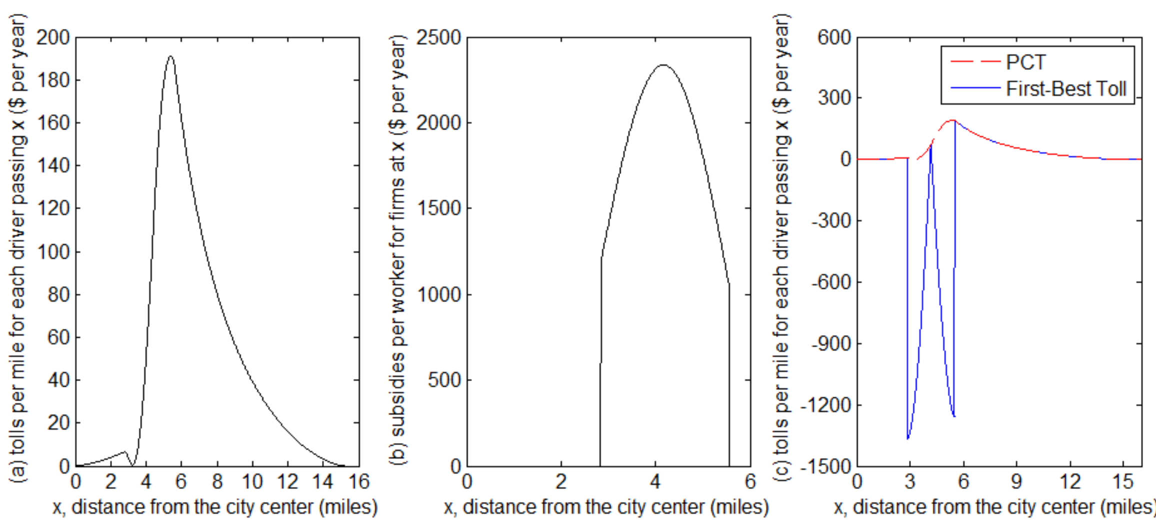


Figure 3.2 Levels of Toll (a) and Subsidy (b) under the First-Best Instrument Combining Both Pigouvian Congestion Toll and Pigouvian Labor Subsidy and Levels of Toll (Compared to Pigouvian Congestion Toll-Only) under the First-Best Tolling Instrument after Internalizing Agglomeration Externalities (c). (φ=20, ρ=0.00001, σ=1.5, γ=0.06, N=600,000)

---

point elasticity of net income with respect to utility level at location x, and $(u-u_0)\frac{dy}{du}$ represents the income change due to the utility level changes from u to u_0.

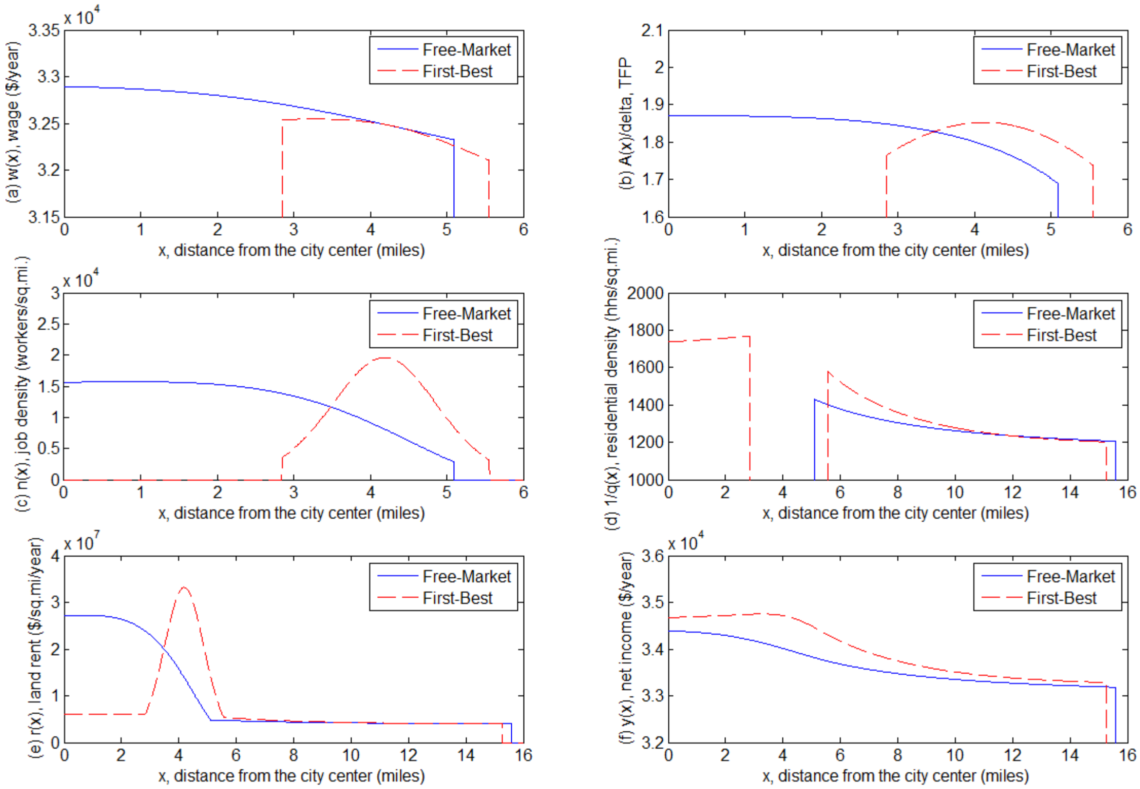


Figure 3.3 Spatial Distribution of Wages (*w*), TFP (*A/δ*), job (*n*) and Residential Densities (1/*q*), Land Rents (*r*) and Net Income (*y*) in the First-Best Optimum versus the Fee-Market Equilibrium (φ=20, ρ=0.00001, σ=1.5, γ=0.06, N=600,000)

Land use patterns are also affected, as shown in Figure 3.3. The first-best instrument causes firms to decentralize, away from the city center, and agglomerate in a smaller cluster, as an annulus, with average labor density rising by 14% (in that ring, versus the original jobs zone). This is a combined consequence of the imposition of Pigouvian congestion toll and Pigouvian labor subsidy. First, the Pigouvian labor subsidy encourages firms at locations of higher productivity to hire more workers, thereby reinforcing agglomeration externalities of their locations. Since labor supply is assumed fixed, firms at locations with lower productivity will lose labor and thus productivity. These shifts stimulate firms to locate closer to each other, clustering in a smaller area, raising job densities and total agglomeration economies (Figure 3.3c). Second, the Pigouvian congestion toll increases the per-mile commuting costs, thereby encouraging firms and workers to co-locate closer together, to reduce travel costs. While road tolls are paid by workers, firms need to provide an attractive wage that internalizes much of the toll to remain competitive. Firm decentralization (and some inward migration of households) can bring them closer to their workers while reducing inward traffic flows.

First-best instruments also centralize households, resulting in a shrinking city boundary from 15.57 to 15.26 miles and higher residential densities over most areas of the city, especially at locations closer to the firm cluster (Figure 3.3d). If comparing Figure 3.3d and 3.3e, we find that higher residential densities raise household bid-rents. The average land rents for firms and houses in the socially optimal setting are 4.52 and 1.05 times the opportunity rent (i.e., the rent at the city edge, Ra), and 13% and 3.8% higher than those in the free-market equilibrium (Table 3.3). Given that all congestion tolls and rent revenues (net of labor subsidies) are uniformly returned to each household, net incomes rise in all locations (Figure 3.3f), with average net income rising by 0.8% (Table 3.3). Notice that utility values rise with net income levels and fall with residential rents, everything else equation (as evident in Eq. 3.8). Even though housing's rent growth is about five times the net income growth, households still experience higher utility, since the elasticity of utility with respect to residential rent is much lower than that with respect to net income (0.1 versus 1).

Pigouvian Congestion Toll Alone (PCT-Alone) Policy in the Base Scenarios

The Pigouvian congestion toll-alone policy imposes a congestion toll that equals the MCE, but does not correct the agglomeration externality. Simulations shows that each worker driver needs to pay an average toll of $190 per year, about one-third of the average first-best toll (Table 3.3). The Pigouvian congestion toll-alone policy, however, generates a significant welfare loss: the utility decrease to 5221 and the CV value relative to the free-market case is -152. This utility loss results from the negative side effect of Pigouvian congestion toll on the agglomeration economy. Compared to the free-market equilibrium, the Pigouvian congestion toll-alone equilibrium leads to a 92% decrease in congestion diseconomies and a 7.4% decline in the agglomeration benefits. While the average commute distance decreases by 41%, the average productivity drops by 3.6% (Table 3.3).

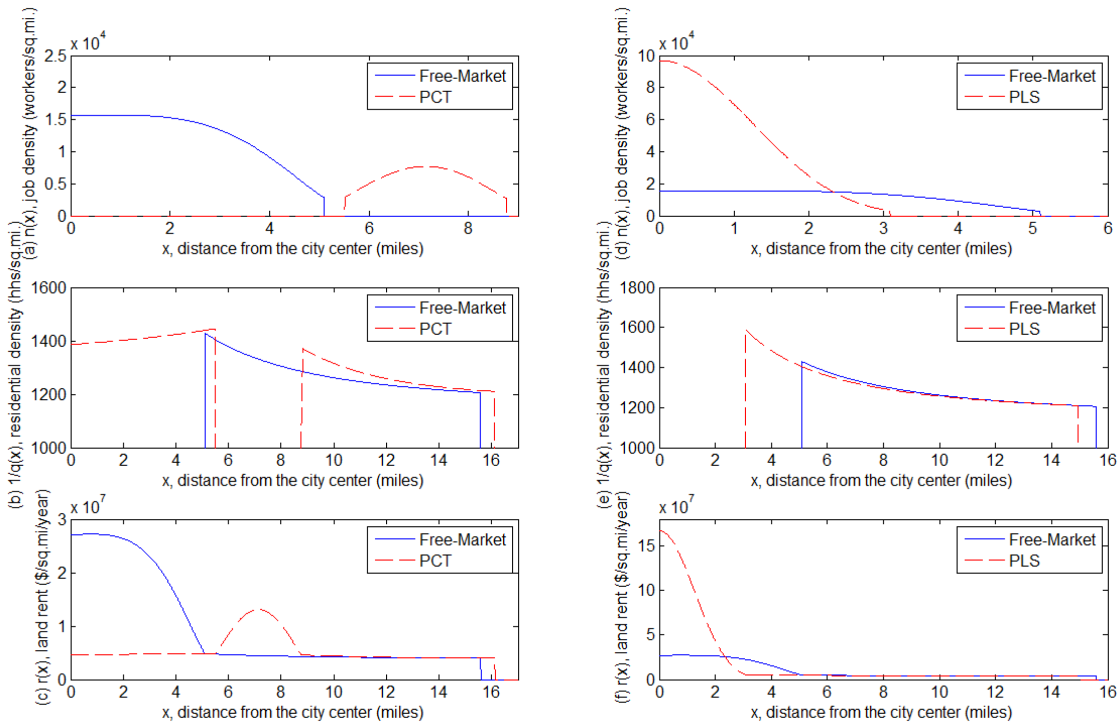


Figure 3.4 Spatial Distribution of Job Density (n), Residential Densities ($1/q$), and Land Rents (r) in the Pigouvian Congestion Toll-Alone versus the Free-Market Equilibria (Left) and in the Pigouvian Labor Subsidy-Alone Equilibrium versus the Free-Market Equilibria (Right) (φ=20, ρ=0.00001, σ=1.5, γ=0.06, N=600,000)

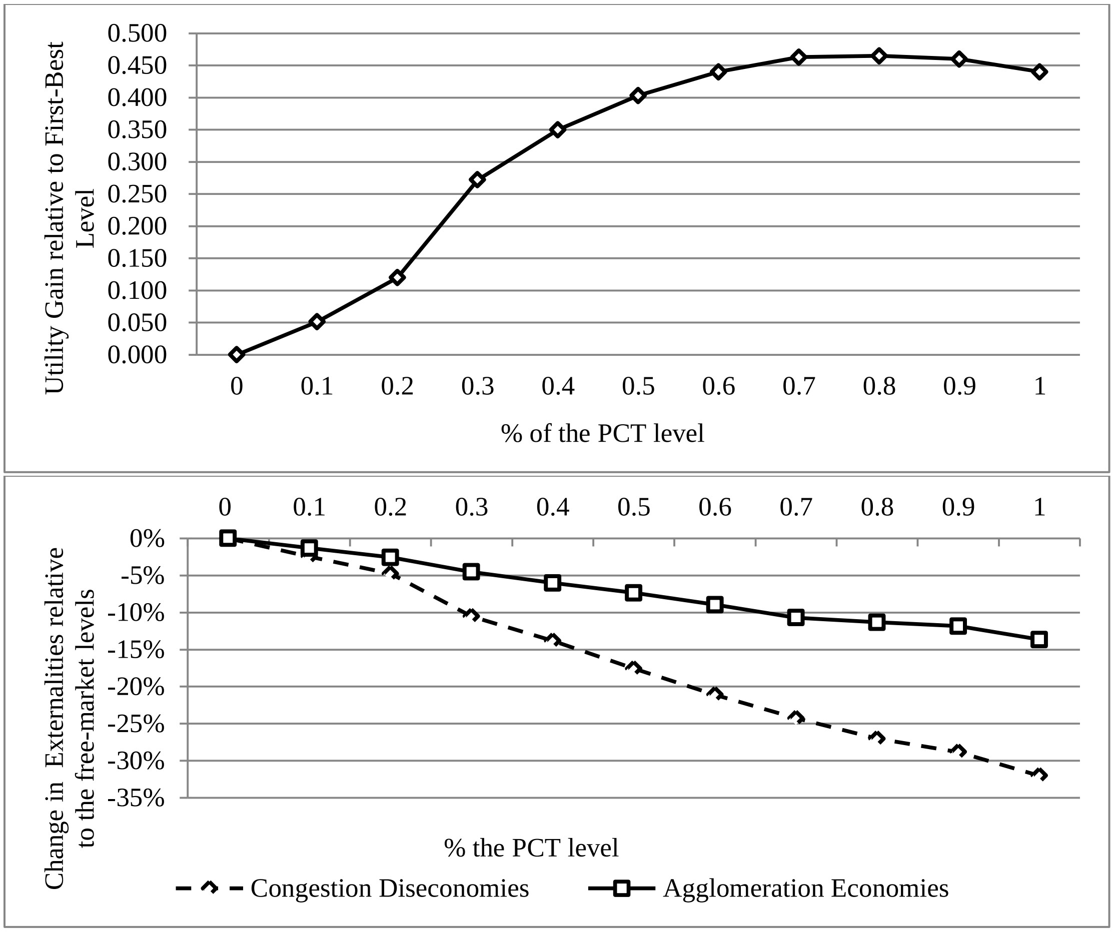


Figure 3.5 Utility Gains Relative to the First-Best Level (Left) and the Percent Changes in Aggregate Congestion Diseconomies and Agglomeration Economies (Right) under 10 Tolling Schemes.

The Pigouvian congestion toll-alone policy leads to a more sprawling urban form and more job decentralization than the free-market equilibrium (Figure 3.4a-c) and first-best optimum (Figure 3.3). The city boundary increases to 16.1 miles, creating 6.9% more land areas than the free-market case. Without the incentive of a Pigouvian labor subsidy to guarantee labor supply, the Pigouvian congestion toll-alone policy incentivizes firms and workers to locate closer to each other, to reduce commuting costs and better match labor supply and demand. For example, a Pigouvian congestion toll levied in location x will lessen the level of commute volume passing *x* to a socially optimal level, making some workers relocate to avoid paying the toll at x. Some workers will change their workplace to the location outside *x*, while some workers will move inside to live near the city centerpoint for outward commuting. These demand-side adjustments will decentralize firms to relatively low-productivity locations, since the lower-productivity locations are closer to the edge of the firm cluster and thus households. Compared to the free-market equilibrium, the average residential and labor densities decrease by 1.1% and 44%, and the residential and firm's rents drop by 1.5% and 45% (Table 3.3).

For seeking the impact of anti-congestion policies on agglomeration economies, Figure 3.5 tracked the change in utility and externalities under ten additional tolling schemes, which impose a fixed share (ranging from 0 to 0.9) of the Pigouvian congestion toll level on each mile driven. The figure presents the percent of utility gains relative to that in the first-best optimum and the percent change of two aggregate externalities relative to the free-market case. The second-best utility gains peak at about 47% of the first-best utility gains (compared to the based equilibrium), when the toll level is set as about 72% of the Pigouvian congestion toll level. As the share increases, the congestion and agglomeration externalities decline, indicating that anti-congestion policies may erode agglomeration economies. These findings also suggest that an efficient toll level should lie below the Pigouvian congestion toll level, as agglomeration economies are internalized.

Pigouvian Labor Subsidy Alone (PLS-Alone) Policy in the Base Scenarios

The Pigouvian labor subsidy-alone policy offers labor subsidies to firms in the amount of agglomeration's marginal externality benefits, but does not correct the congestion externality. In the Pigouvian labor subsidy-alone equilibrium, the city needs to deliver an average subsidy of $2081 per job per year, so that agglomeration external benefits can be redistributed back to firms. Similar to the Pigouvian congestion toll-alone policy, the Pigouvian labor subsidy could cause significant welfare loss, with an average $118 per year loss of CV value relative to the free-market equilibrium. But in contrast to the Pigouvian congestion toll-alone instrument, the Pigouvian labor subsidy-alone policy increases total agglomeration benefits and congestion costs by 26% and 105%, respectively. The growing agglomeration economies links with a 5.3% increase in average productivity and a 5.4% decrease in average wage cost, while the rising congestion diseconomies are resulted from a 6.9% increase in average commute distance (Table 3.3).

The Pigouvian congestion toll-alone policy leads to a more compact urban form than the free-market equilibrium (Figure 3.4d-f) and first-best optimum (Figure 3.3). The city boundary decreases to 14.94 miles, causing a 7.9% decrease in land areas in comparison with the free-market case. Without the Pigouvian congestion toll's congestion correction, the Pigouvian labor subsidy-alone intervention could encourage firms to locate closer to each other. After levying a Pigouvian labor subsidy policy, firms at locations with relative low productivity (often at the edges of firm clusters) will move to locations with higher productivity. This tendency would agglomerate firms in a smaller area (Figure 3.4d) and job densities increase near the centerpoint and drop at the edge of the firm cluster. The traffic volumes will thus rise within the firm cluster, triggering a rise in congestion. While job centralization accompanies with housing centralization, Pigouvian labor subsidy appears to have trivial direct impact on household's spatial decision, as shown in Figure 3.4e.

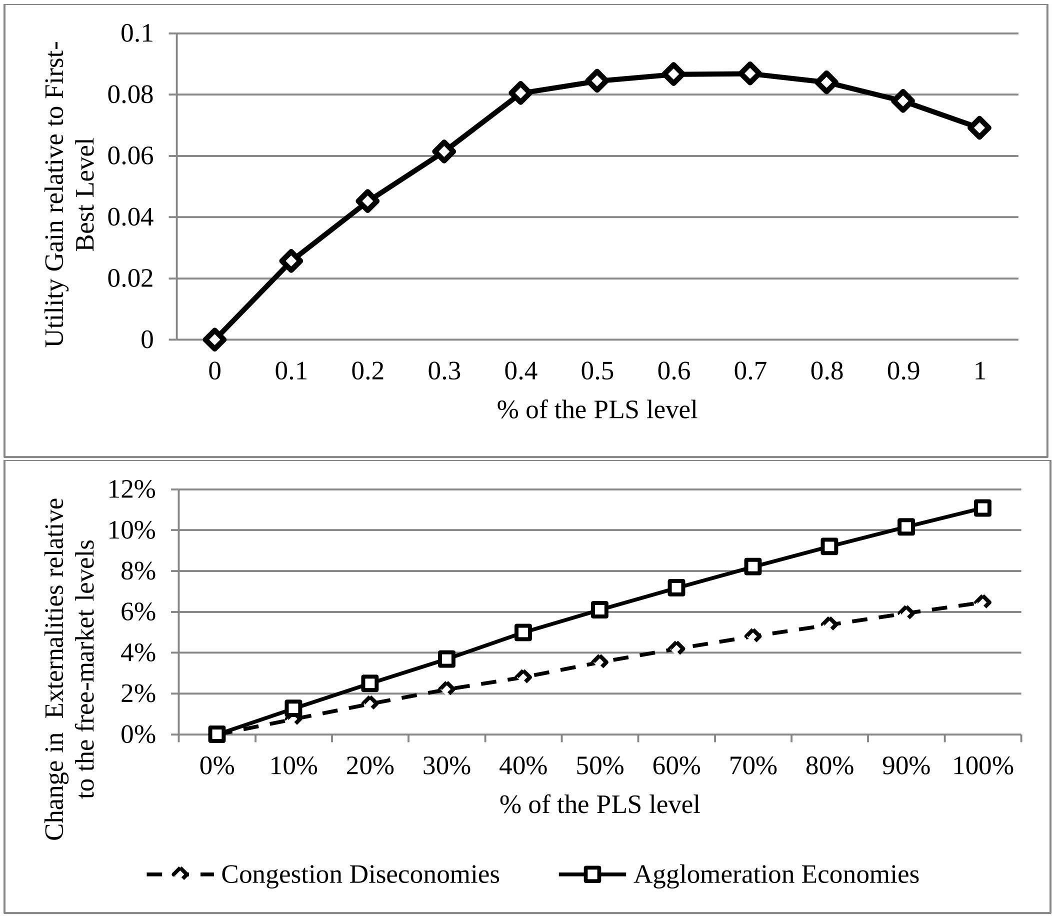


Figure 3.6 Utility Gains Relative to the First-Best Level (Left) and the Percent Changes in Aggregate Congestion Diseconomies and Agglomeration Economies (Right) under 10 Subsidy Schemes.

Similar to those tolling-only schemes, the policies with the labor subsidy level setting below the Pigouvian labor subsidy level may generate more welfare gains than those at the exact Pigouvian level. Figure 3.6 shows that the utility gains relative to the first-best level peak at 8.7%, when the labor subsidy is set at about 68% of the Pigouvian labor subsidy level. As the subsidy levels increase, both the aggregate congestion externality cost and agglomeration externality benefit rise. This finding suggests that the implementation of policies for promoting agglomeration economies should recognize their potential negative impact on worsening congestion.

Policy Evaluations Varying with Congestion and Agglomeration Scales

Table 3.4 Policy Scenario Results under Varying Congestion Levels

	Network Congestibility Parameter ρ				
	5e-6	1e-5	1.5e-5	3e-5	1e-4
Types of Urban Form at FM equilibria	FH	FH	HFH	HFH	HFH
CV of First-Best Policies (relative to FM cases, $/hh./year)	171	114	209	190	103
CV of PCT-Alone Policies (relative to FM cases, $/hh./year)	-255	-152	-59	36	46
CV of PLS-Alone Policies (relative to FM cases, $/hh./year)	13	-118	-251	-379	-652
Share of Congestion Diseconomy in Total Congestion Cost (%)	17.72	7.04	20.91	23.37	9.30
Total Externalities in the FM case (million $ /year)	886	652	932	890	620
Percent Change in Congestion Costs (from FM to FB, %)	-38.40	-77.06	-76.92	-67.38	-56.48
Percent Change in Agglomeration Benefits (from FM to FB, %)	22.12	14.56	17.54	18.67	20.20
Percent Change in City Boundary (from FM to FB, %)	-2.95	-1.99	-3.32	-5.23	-7.97
Percent Change in Avg. Labor Density (from FM to FB, %)	101.44	14.04	37.53	50.50	63.37
Percent Change in Avg. Residential Density (from FM to FB, %)	2.86	3.37	4.11	7.26	12.89
Percent Change in Avg. Rent for Firms (from FM to FB, %)	103.09	13.26	36.94	49.66	63.34
Percent Change in Avg. Rent for Housing (from FM to FB, %)	2.33	3.84	3.86	6.30	12.35

This section conducts a robustness analysis on pricing policies in cities with varying agglomeration or congestion levels, including two parts. The first part discusses the welfare gain/loss, measured by CV relative to corresponding free-market equilibria, of first-best optimum and Pigouvian congestion toll-alone and Pigouvian labor subsidy-alone policies, as major findings summarized in Table 3.4 and 3.5. Meanwhile, by comparing the socially optimal level of traffic volume, one can identify whether congestion is excessive or insufficient in the other three equilibria. The diseconomies of congestion equal the difference in utilities or CV. The robustness analysis thus can examine how inefficient levels of congestion change with agglomeration and congestion parameters.

Table 3.5 Policy Scenario Results under Varying Agglomeration Levels

	Agglomeration Parameter γ				
	0.08	0.07	0.06	0.05	0.04
Types of Urban Form at FM equilibria	FH	FH	FH	HFH	HFH
CV of First-Best Policies (relative to FM cases, $/hh./year)	196	129	114	122	65
CV of PCT-Alone Policies (relative to FM cases, $/hh./year)	-492	-301	-152	-67	18
CV of PLS-Alone Policies (relative to FM cases, $/hh./year)	5	-55	-118	-159	-196
Share of Congestion Diseconomy in Total Congestion Cost (%)	5.88	5.24	7.04	22.66	20.87
Total Externalities in the FM case (million $ /year)	885	742	652	737	566
Percent Change in Congestion Diseconomies (from FM to FB, %)	-47.76	-48.61	-77.06	-58.42	-59.39
Percent Change in Agglomeration Economies (from FM to FB, %)	20.65	20.24	14.56	18.98	17.33
Percent Change in City Boundary (from FM to FB, %)	-2.44	-3.11	-1.99	-4.82	-4.60
Percent Change in Avg. Labor Density (from FM to FB, %)	76.17	75.01	14.04	57.35	42.32
Percent Change in Avg. Residential Density (from FM to FB, %)	3.91	4.66	3.37	5.52	5.26
Percent Change in Avg. Rent for Firms (from FM to FB, %)	78.38	76.46	13.26	56.74	41.21
Percent Change in Avg. Rent for Housing (from FM to FB, %)	4.33	4.59	3.84	3.79	3.20

The second part provides a sensitivity analysis on the land use impact of optimal policies. This helps to explore socially optimal land use patterns. There are two families of equilibrium urban configurations in our simulations, including “FH” and “HFH” (Table 3.4 & 3.5). “FH” represents the traditional monocentric urban structure, with firms/business surrounding the city center and housing locating at the annulus outside the firm cluster. "HFH" represents a non-monocentric structure, in which housing occupies the areas near the city center and edge and firms locate at the middle annulus[10].

Welfare Impact of First-Best Policies

Results suggest that first-best policies are able to lower congestion diseconomies and enhance agglomeration economies, leading to welfare improvement (Table 3.4 & 3.5). The magnitude of welfare gains is determined by types of urban form and total externalities, i.e., agglomeration economies minus congestion diseconomies. An improvement of network congestibility (i.e., ρ drops) or an enhancement of agglomeration (i.e., γ increases) can generate similar effects on urban form, changing from the HFH to FH type. In the same family of urban configuration, first-best policies can achieve higher welfare gains in cities with lower congestion levels (or larger agglomeration scales). This appears to contradict the partial equilibrium findings relying on traditional monocentric models with congestion externalities internalized only, which suggest that the welfare gains in the optimum are larger in higher-congestion cities (e.g., Brueckner, 2007).

In general, the larger the total externalities, the more welfare gain created by first-best policies. Our findings suggest that higher congestion levels may indeed create more negative external costs to the society but, meanwhile, discourage agglomeration and lower positive production externalities, leading to a decrease in total externalities (Table 3.4 & 3.5). Similarly, higher agglomeration scales may bring increases in both congestion and production externalities. However, the increased amount of positive externalities is larger than the increased amount of negative externalities, leading to an increase in total

[10] In addition, mixed land use patterns could be an equilibrium solution but this equilibrium allocation is never *Pareto-optimal* under the modeling framework in this paper. Since the policy scenarios only compare the *Pareto-optimal* equilibria or optimum, the family of mixed urban forms is thus not this chapter's focus. A detailed theoretical and simulation discussion can refer to Appendix A4.

externalities. Thus, it may be inappropriate for policy makers to apply optimal policies found in partial equilibrium models to improve market efficiency in cities with multiple externalities.

Optimal Levels of Congestion

Two key questions are surveyed here: (1) Do the diseconomies of congestion increase with the levels of network congestibility (i.e., ρ)? (2) How much of "the total cost of congestion" does the congestion diseconomy contribute? The answer to the first question is "yes" in the traditional monocentric models where only congestion externalities are internalized (Brueckner, 2007). After accounting for agglomeration externalities, the simulation results suggest that the diseconomies of congestion in free-market cases vary with urban forms. Among the same type of urban form (e.g., FH or HFH), the congestion diseconomies appear to decrease with an increase in congestibility. In addition, the percentage of the congestion diseconomy in the total congestion cost ranges from 5% to 23%, varying with the two parameters.

This section also examines when Pigouvian congestion toll-alone or Pigouvian labor subsidy-alone policies are somewhat effective in generating less inefficient congestion than the free-market case. According to Tables 3.4 and 3.5, Pigouvian congestion toll-alone policies generate positive welfare gains when the congestion levels are high enough (e.g., ρ = 0.00003 and 0.0001) or the agglomeration levels are low enough (e.g., γ=0.04). In some extreme situations where agglomeration externalities do not exist, i.e., γ=0, the Pigouvian congestion toll-alone policies are first-best. Similarly, the Pigouvian labor subsidy-alone policies can be effective when the agglomeration levels are high (e.g., γ=0.08) or the congestion level is low (e.g., ρ = 0.000005). Many simulations, however, remind policymakers that correcting only for one externality in cities with multiple externalities may achieve very low, or even negative, welfare gains.

Land Use Effects

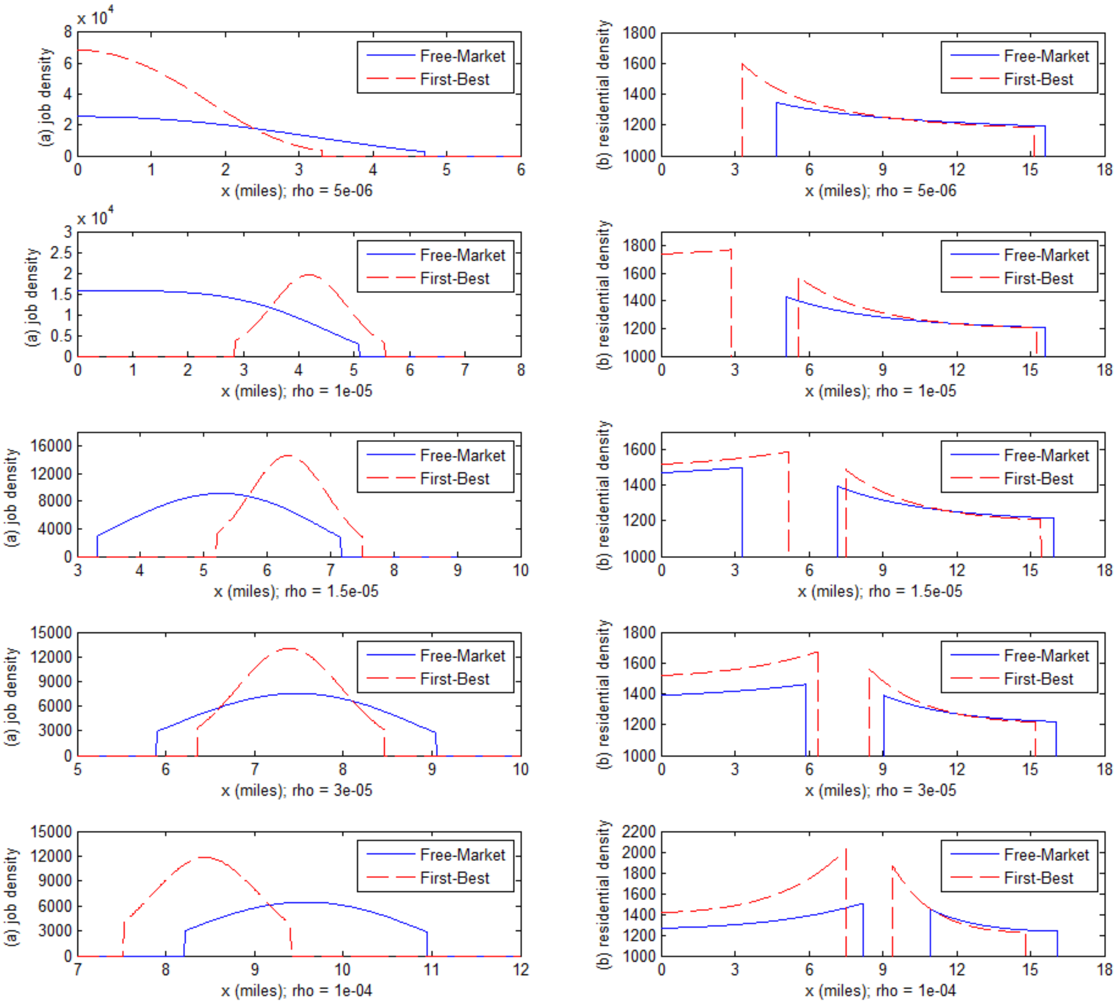


Figure 3.7 Spatial Distribution of Job Density (n) and Residential Densities ($1/q$) in the First-Best Optimum versus the Free-Market Equilibrium, When the Congestion Parameter ρ Increases from 0.000005 to 0.0001. (φ=20, σ=1.5, γ=0.06, N=600,000)

Network congestibility is an important determinant of equilibrium and optimum city size and land use patterns. Figure 3.7 shows land use densities and urban forms of the free-market equilibrium and the first-best optimum when road congestibility levels rise (i.e., ρ increases from 0.000005 to 0.0001). As network congestiblities increase, firms/jobs in the free-market equilibria increasingly decentralize, leading to a change in urban forms from monocentric (i.e., FH) to annular (i.e., HFH) structures (Figure 3.7). These findings are generally consistent with those of Fujita and Ogawa (1982), Berliant et al. (2002), and LRH (2002), although those other models do not allow for congestion and wealth redistribution. Such connections support the notion that job decentralization are at least in part driven by worsening congestion levels in existing, evolving regions, and in turn help relieve traffic congestion, as suggested by commute times and costs in Giuliano and Small's (1991) and Crane and Chatman's (2003) empirical work.

Several findings are comparable to those found in monocentric settings. For example, under free-market equilibria, higher congestion levels may make the city/region more compact (as noted in Brueckner [2007]). The boundary of 15.59 mile defines the region's radius under the less congestion case ($\rho = 0.000005$), falling to 15.57 mile in the higher congestion case ($\rho = 0.00001$). However, several findings differ from, or not easily detected in, a monocentric model. When the congestion levels increase to relatively high levels, the urban form changes to HFH structure and higher levels may make the city more sprawling, rather than more compact. As ρ increases from 0.000015 to 0.0001, the equilibrium city boundaries rise from 15.94 to 16.06 miles. In the monocentric models without internalizing firm's spatial decisions, an increase in congestion levels can only affect households' behaviors and make them live closer to the city center, leading to a compact city size. However, when firm's spatial decisions are internalized in our model, an increase in congestion levels not only encourage job-housing proximity, but also make firms decentralized and distribute within a larger area. The combination of these spatial impacts could lead to a sprawling city size.

Figure 3.8 compares land use densities and urban forms of the free-market equilibrium and the first-best optimum. As the agglomeration parameter γ increases from 0.04 to 0.08, firms become increasingly centralized and the urban structure changes from the HFH to FH form. No matter in which urban forms, a higher γ is associated with a smaller firm cluster area but a larger city size, with the boundaries ranging from 15.23 at $\gamma = 0.04$ to 16.83 at $\gamma = 0.08$. As γ increases, firms are more willing to locate closer to other firms for earning external benefits. This cause firms to agglomerate in a smaller area (Figure 3.8), raising job densities, locational productivity, bid-rents, and wage. The increase in workers' wage income thus allows them to live in larger house and pay for farther commute, leading to lower residential densities and a larger city size (Figure 3.8).

First-best policies lead to more compact urban forms, regardless of congestion and agglomeration levels. The optimal city boundaries are always smaller than the equilibrium boundaries, and residential densities significantly raise at most locations (Figure 3.7). The percentage changes in average residential density after levying the optimal policies increase from 2.9% to 12.9% when ρ increases (Table 3.4). These residential densification effects of first-best policies are similar to monocentric studies (e.g., Brueckner, 2007). In addition, optimal policies can largely raise job densities and land rents for firm use, while bringing relatively smaller increases in residential densities and housing land rents (Table 3.4).

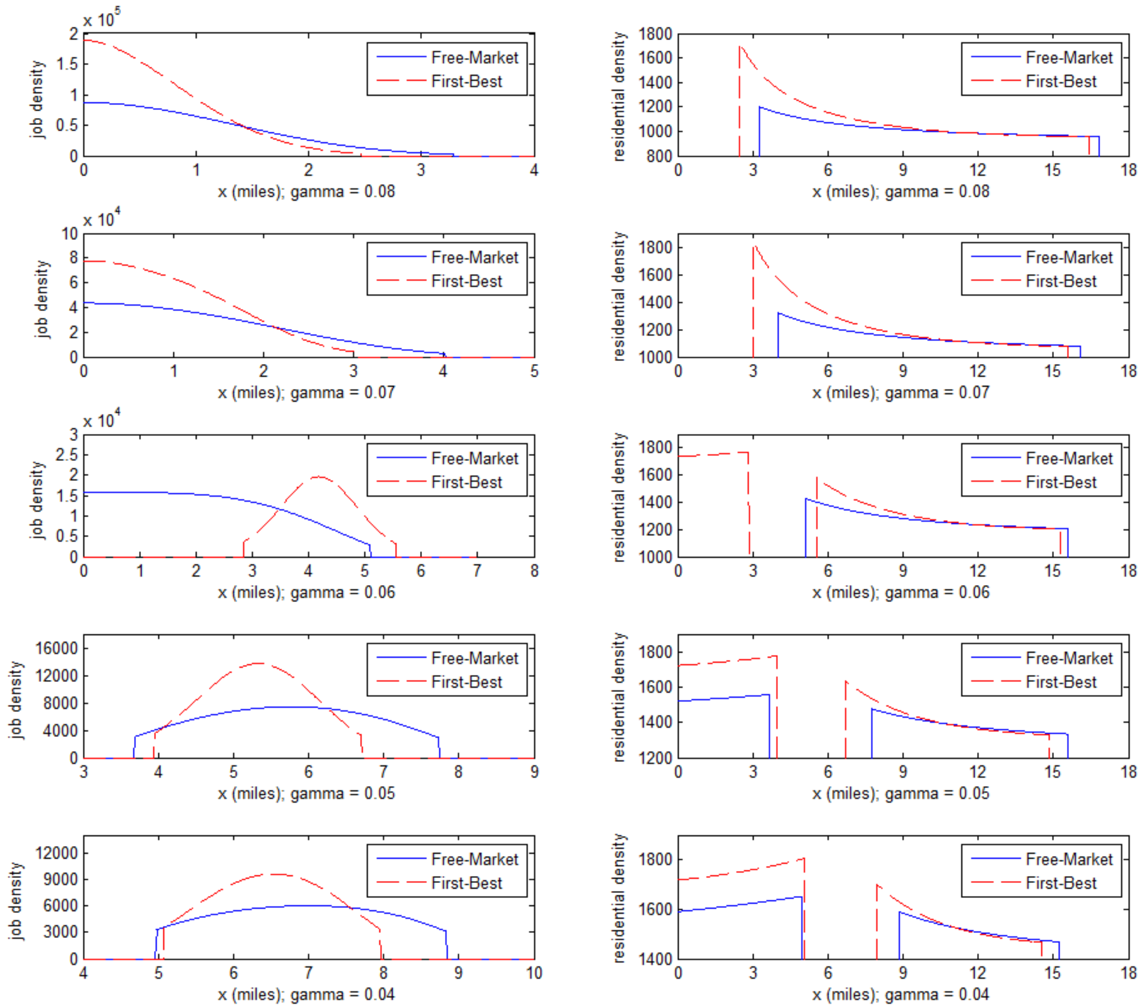


Figure 3.8 Spatial Distribution of Job Density (n) and Residential Densities ($1/q$) in the First-Best Optimum versus the Free-Market Equilibrium, When the Agglomeration Parameter γ Increases from 0.04 to 0.08. (φ=20, ρ=0.00001, σ=1.5, N=600,000)

On the other hand, depending on ρ and γ, first-best policies may differently affect the spatial distribution of firms and jobs (Figure 3.7 and Figure 3.8). In the low-congestion ($\rho = 0.000005$) or high-agglomeration case (γ=0.07 and 0.08), the optimal policy leads jobs and firms more centralized in a smaller firm cluster area near the centerpoint, causing signficiantly increase in average job densities (Table 3.4 & 3.5). In the median-congestion (ρ=0.00001 and 0.000015) or median-agglomeration (γ=0.06) cases, the optimum generates significant job decentralization and still an increase in job densities. In the high-congestion (ρ=0.00003 and 0.0001) or low-agglomeration (γ=0.04 and 0.05) cases, firms in the free-market equilibria appear over-decentralized and first-best policies thus cause a more centralized firm cluster. Thus, while first-best policies will always cluster firms in smaller areas with higher job densities, they can lead to either job centralization (in very low or very high congestion or agglomeration levels) or decentralization (in median congestion or agglomeration levels).

EXCESSIVE CONGESTION IN CITIES WITH PLANNING AND MARKET FAILURES

While the above simulation findings suggest that nonmonocentric urban forms are market desirable or socially optimal outcomes, especially for those with high-level congestion, this type of urban form is observed less often in our living cities. There appear to be some regulations – either from land use zoning law, infrastructure shortage, or other historical constraints – that deter the emergence of such nonmonocentric forms. In fact, our living cities are not only filled with externalities, but also zoning regulations, especially in the US. Two types of land use regulations are discussed in this chapter: exclusionary zoning and low-density zoning. As discussed in Chapter 2, such exclusionary and low-density regulation can cause more auto-travel demand than desired by the market and the society, leading to excessive congestion. This section extends the model framework developed in previous sections to evaluate the optimal and excessive levels of congestion in cities with both market and planning failures. The modeling results will be compared with those that account only for market failures.

Modeling Land Use Regulations

Exclusionary zoning is probably the most popular land use regulation that excludes certain types of land use from a particular area such as a subdivision. The direct consequence of exclusionary zoning is the separation of land use and the single-use landscape. Because the analysis here relies on an abstract model with only two types of land use, residential and commercial, we simply define the exclusionary zoning as the regulation that excludes firms from suburban residential areas. Thus, the share of commercial land use for firms under exclusionary zoning $\theta_f^r(x)$ is defined below:

$$\theta_f^r(x) = \begin{cases} \theta_f^*(x), & x \in [x_0, x_1] \\ 0, & x \notin [x_0, x_1] \end{cases} \tag{3.30}$$

Density regulation is another widely used zoning ordinance that restricts the maximum density that may be constructed within an area. Different municipalities have variable approaches to establishing a low-density zoning regulation, including restrictions on maximum numbers of houses per acre, the minimum lot size, the maximum floor area ratio (FAR), and the maximum height of buildings. This modeling analysis provides a general and straightforward way to internalize low-density regulations. Under low-density zoning regulations, no residential densities can exceed a preset density cap *M*. When the market-desirable density at location *x*, i.e., the inverse of $q^*(x)$ in Eq.(3.7), is below the density cap, the residential density under regulation $\frac{1}{q_r(x)}$ equals the density the market desired. $q_r(x)$ is the equilibrium outcome of residential density at location *x* under low-density zoning. Once the market-desirable density is above *M*, the derived density under regulation equals the cap level *M*. Thus, the equilibrium residential density at location *x* under low-density regulation is defined as follows:

$$\frac{1}{q_r(x)} = \begin{cases} \frac{1}{q^*(x)}, & if \ \frac{1}{q^*(x)} < M \\ M, & if \ \frac{1}{q^*(x)} \geq M \end{cases} \tag{3.31}$$

Low-density zoning regulation will also affect residential rents. When the city reaches equilibrium, residents' utility levels will be maximized at $\bar{u}$. The equilibrium rent under low-density regulations $r_h^r(x)$, which differs from the level in the free market equilibrium $r_h^m(x)$ in Eq.(3.8), is defined as follows:

$$r_h^r(x) = \frac{y(x) - \bar{u}^{\frac{1}{a}} q_r(x)}{q_r(x)} \tag{3.32}$$

This section continues to use computational simulations to evaluate the impacts of zoning regulations on travel demand, congestion, land use, and social welfare. All experiments were conducted in the context of base scenarios (with γ =0.06 and ρ =0.00001). The simulations will examine two packages of regulation including exclusionary zoning and LUZ. For robustness analysis, we increase the restrictions on firms' limits from 4 to 3 to 2 miles; that is, firms can locate only at the planning areas from (0, 4] to (0, 3] to (0, 2]. In addition, the density caps *M* decrease from 1,250 to 1,000 to 800 households per square mile. Under these regulations, simulations solve for the equilibrium in the free market and social optimum with a combination of Pigouvian congestion toll and Pigouvian labor subsidy. The results are compared to the free-market equilibrium and social optimum in cities without land use regulations, as described in previous sections.

Excessive Congestion under Land Use Regulations: Market Failures versus Planning Failures

The total amount of diseconomy (net social costs) of excessive congestion is calculated by the utility difference (monetized by the values of compensating variation CV) of the free-market equilibrium under regulations and the social optimum without any regulations. Among them the utility differences between the free-market and first-best equilibria under regulations represent the diseconomy from market failures, while the utility differences between first-best optimum with and without regulations are the diseconomy from planning failures (for more discussion of the calculation principles, refer to Chapter 2).

Table 3.6 summarizes simulation results from the net costs of excessive congestion from six scenarios of land use regulations and identifies which costs come from planning versus market failures. In the base scenarios without any regulations (Table 3.3), firms are agglomerated in the urban center from 0 to 5.09 miles; i.e., in the circular interval [0, 5.09 miles]. As shown in Table 3.6, after an exclusionary zoning regulation is imposed to exclude firms from the area outside the circular interval [0, 4 miles], the net social costs of excessive congestion for each household are about $52 per year, in which 47% are from market failure and 53% are from planning failure. This suggests that first-best pricing policies (e.g., a combination of Pigouvian congestion toll and Pigouvian labor subsidy) can correct only 47% of excessive congestion, while first-best land use policies removing all regulations can correct 53% of excessive congestion. When exclusionary zoning regulations become increasingly restrictive with the zoning area from the intervals outside 4, 3, and 2 miles, the total diseconomies of excessive congestion gradually increase from $52 to $154 to $609 per household per year. Meanwhile, the shares of costs from planning failures increase from 53% to 85% to 96%. When firms are allowed to locate only at the urban core 2 miles from the center point, the planning failure from the exclusionary zoning regulation is the dominant cause of excessive congestion and causes about $584 in losses per year for each household.

Similar results are found in the scenarios with low-density zoning regulations. This research sets three density caps under the average residential density in the free market without regulations (i.e., 1,260 households per square mile as shown in Table 3.3): 1,250, 1,000, and 800 households per square mile. Simulations show that the lower the maximum levels of residential density, the more total diseconomies of excessive congestion occur (Table 3.6). When the density cap is limited to 1,250 households per square mile, each household incurs a $142 loss, only 16% from regulation and 84% from market failure. Above 85% of excessive congestion results from planning failure when residential densities are constrained to under 800 households per square mile. Thus planning failure plays a more important role in affecting congestion and social welfare when more restrictive density regulations are imposed.

Table 3.6 Components of Total Diseconomy of Excessive Congestion: From Planning versus Market Failures in the Base Scenarios (γ=0.06 and ρ=0.00001)

Land Use Regulation	Net Social Cost of Excessive Congestion ($/hh/year)	Costs from Planning Failures ($/hh/year)		Costs from Market Failures ($/hh/year)	
Exclusionary Zoning, Firm Cannot Locate outside [0,4]	51.67	27.43	53.08%	24.24	46.92%
Exclusionary Zoning, Firm Cannot Locate outside [0,3]	154.08	131.62	85.42%	22.46	14.58%
Exclusionary Zoning, Firm Cannot Locate outside [0,2]	609.29	583.98	95.85%	25.31	4.15%
Low-Density Zoning, Residential Density Cap is 1250 hhs/sq.mi.	141.90	22.75	16.03%	119.16	83.97%
Low-Density Zoning, Residential Density Cap is 1000 hhs/sq.mi.	283.78	172.93	60.94%	110.85	39.06%
Low-Density Zoning, Residential Density Cap is 800 hhs/sq.mi.	661.88	563.75	85.17%	98.13	14.83%

These findings suggest that planning failures can be more serious, causing more excessive congestion and welfare loss than market failures from congestion and agglomeration externalities. This demonstrates the importance of combining both land use and pricing policies. Neither single policy can fully and feasibly correct both failures. Even the first-best pricing policies cannot correct planning failure from exclusionary zoning and low-density zoning. While economists often believe pricing policies are superior to land use policies for mitigating traffic congestion (Brueckner, 2000), this research suggests that land use planning could be superior to economic policies. In cases when planning failures dominate, the first-best pricing policies can reduce only a very low percentage of excessive congestion, and they produce low welfare improvement. Instead, many planners often embrace only land use policies such as regulation reform and development of alternatives to low-density sprawl (Levine, 2006). This research suggests that land use planning-promoting alternative developments are important when land use regulations largely constrain development desired by the market, but could be trivial when planning failures are negligible.

Table 3.7 Anti-Congestion Effects of Combined Congestion Pricing and Land Use Planning strategies

		Comparisons to NO Policies (%)		
		Optimal Congestion Pricing(CP)	Optimal Land Use Planning(LUP)	CP+LUP
Exclusionary Zoning, Firms Cannot Locate outside [0,2]	Avg. Commute Distance (miles/day)	-4.95	-3.86	-24.96
	Avg. Traffic (vehs/hr)	0.81	-37.53	-59.60
Exclusionary Zoning, Firms Cannot Locate outside [0,3]	Avg. Commute Distance (miles/day)	-3.86	-2.24	-23.69
	Avg. Traffic (vehs/hr)	-0.01	-24.99	-51.50
Exclusionary Zoning, Firms Cannot Locate outside [0,4]	Avg. Commute Distance (miles/day)	-4.16	-0.94	-22.68
	Avg. Traffic (vehs/hr)	-4.18	-13.10	-43.81
Low-Density Zoning, Residential Density Cap is 1250 hhs/sq.mi.	Avg. Commute Distance (miles/day)	-18.32	-1.41	-23.05
	Avg. Traffic (vehs/hr)	-32.54	3.43	-33.12
Low-Density Zoning, Residential Density Cap is 1000 hhs/sq.mi.	Avg. Commute Distance (miles/day)	-13.44	-12.14	-31.42
	Avg. Traffic (vehs/hr)	-26.33	7.46	-30.51
Low-Density Zoning, Residential Density Cap is 800 hhs/sq.mi.	Avg. Commute Distance (miles/day)	-12.79	-22.43	-39.45
	Avg. Traffic (vehs/hr)	-27.76	9.82	-28.99

Table 3.7 compares simulated anti-congestion effects of optimal congestion pricing, optimal land use planning, and a combination of congestion pricing and land use planning policies. Findings indicate the effectiveness of incorporating land use and economic policies as a strategy to reduce auto travel and relieve congestion. An interesting finding is that either a congestion pricing-alone or land use planning-alone policy may worsen congestion with more commutes or traffic on the roads. For example, in cities with very restricted exclusionary zoning regulation (i.e., firms are regulated at [0,2] only) optimal pricing can reduce commute distance by about 5%, but increase average traffic by 0.8%. Under low-density zoning regulations, the average levels of traffic volume will increase and traffic congestion will become worse. On the other hand, the policy incorporating congestion pricing and land use planning can serve as the most effective strategy to reduce excessive commute demand and congestion, and the combination policy may perform much better than either policy alone.

Land Use Impacts of Regulations

Apart from land use regulations' travel impact, this section focuses on their influences on urban form, land use distribution, and land rent. Figure 3.9 shows simulated urban forms and densities of a metropolitan area with and without exclusionary zoning regulation. After imposing exclusionary zoning regulations, the city has a more compact urban size (Table 3.8) because firm decentralization is restricted. While the socially optimal location of the firm cluster will decentralize to the surrounding area from 3 to 5.5 miles, the exclusionary zoning regulation will restrict such firm and job decentralization and cause welfare loss. According to Table 3.8, exclusionary zoning regulations may trigger a large increase in the densities of firms and jobs (by 62% to 546%) but bring a relatively low increase in average residential density (by 0.3% to 3%).

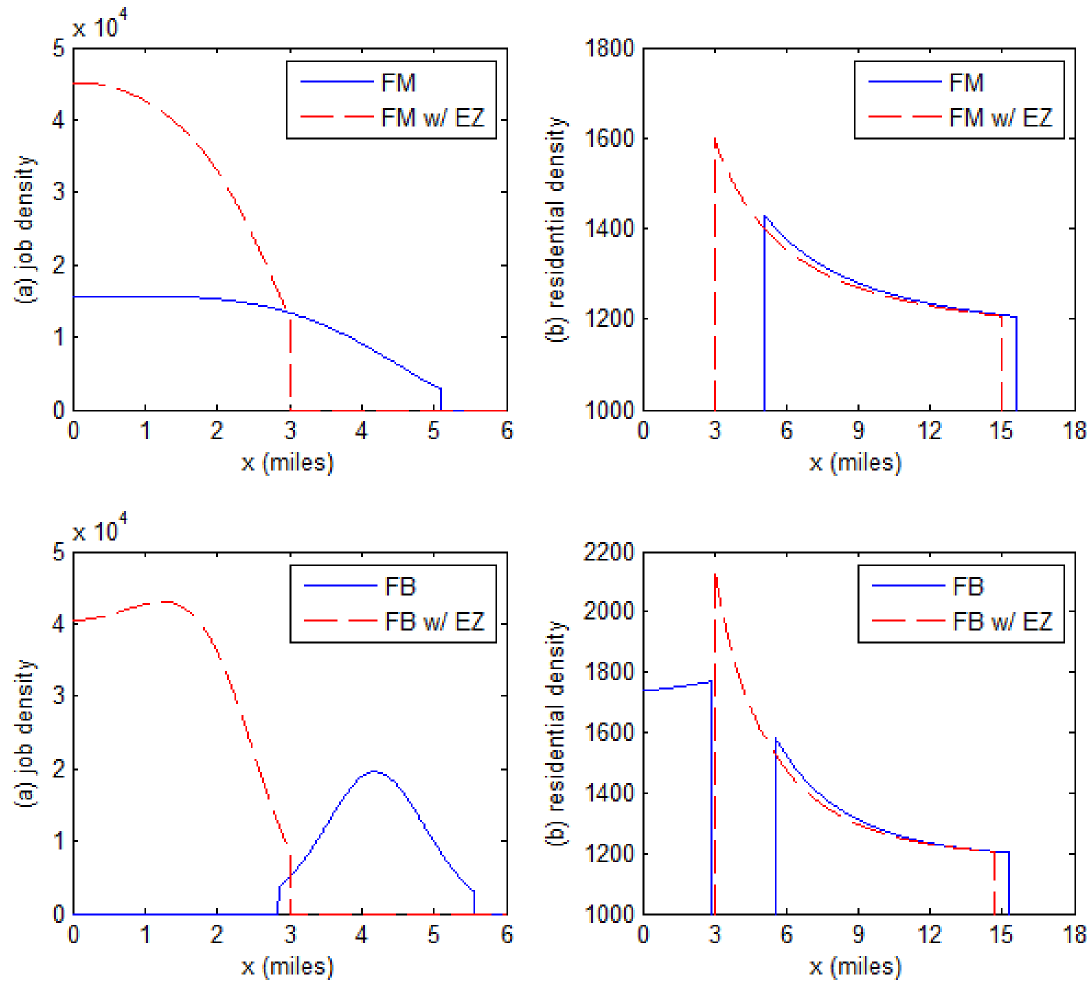


Figure 3.9 Residential and Job Densities of Free-Market (upper) and First-Best (below) Equilibria with and without Exclusionary Zoning that Excludes Firms in the Area outside 3 Miles from the Center

Low-density zoning regulations differ from exclusionary zoning in their land use impacts. Low-density zoning does not place a constraint on firm decentralization but does restrict the maximum density of housing or residents. Figure 3.10 shows density distribution and urban form before and after imposing a particular low-density zoning. Simulations show that cities with low-density zoning may have a relatively sprawling urban form. The city will spread out with low, flat density and a larger city size. This corresponds to many planning studies that believe low-density zoning regulation is a major factor in creating urban sprawl. As estimated in Table 3.8, low-density zoning can lower densities of both firms/jobs and housing/residents.

Table 3.8 Percentage Change of Land Use Characteristics before and after Regulations in the Free-Market Cases

Percentage Change After Regulations in the Free Market	City Area (sq. mi.)	Avg. Labor Density (wks./sq. mi.)	Avg. Residential Density (hhs /sq. mi.)	Avg. Rent for Firms ($/sq. ft.)	Avg. Rent for Housing ($/sq. ft.)	Avg. Labor Wage ($/yr.)
Exclusionary Zoning, Firm Cannot Locate outside [0,2]	-11.59	545.75	2.93	547.41	1.80	0.26
Exclusionary Zoning, Firm Cannot Locate outside [0,3]	-7.93	187.48	1.07	188.61	1.09	0.39
Exclusionary Zoning, Firm Cannot Locate outside [0,4]	-4.32	61.84	0.26	62.23	0.52	0.24
Low-Density Zoning, Residential Density Cap is 1250 hhs/sq.mi.	2.85	-8.45	-2.04	-8.40	-0.13	0.06
Low-Density Zoning, Residential Density Cap is 1000 hhs/sq.mi.	24.17	-8.45	-20.61	-8.50	-0.63	-0.05
Low-Density Zoning, Residential Density Cap is 800 hhs/sq.mi.	51.59	-3.81	-36.43	-3.84	-1.10	-0.03

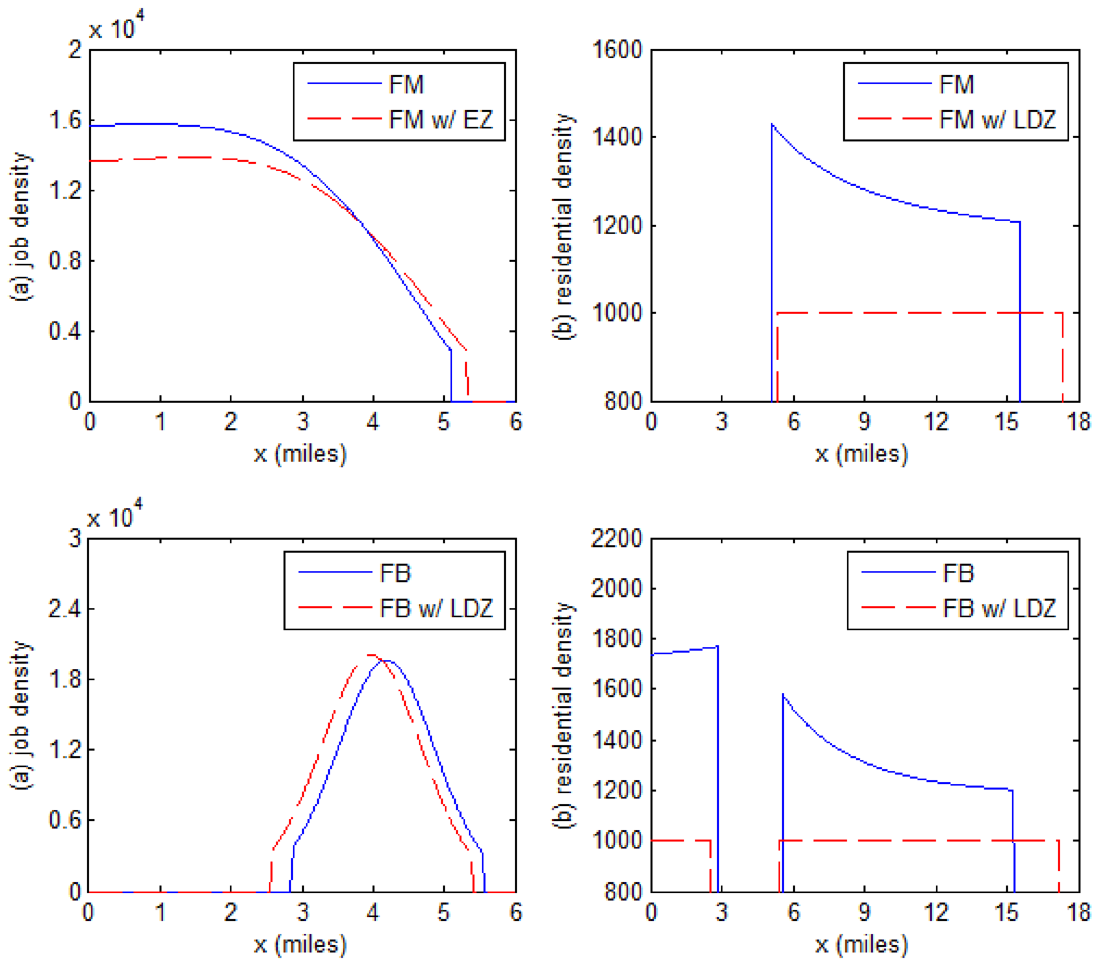


Figure 3.10 Residential and Job Densities of Free-Market (upper) and First-Best (below) Equilibria with and without low-density zoning that Restrict the Maximum Residential Density as 1000 Households per Square Mile

A much-discussed aspect of land use regulations in academics is their impacts on land rents and housing price. Many empirical studies have reported that low-density regulations restrict the supply of housing and land, thus producing escalating housing prices or rent and decreasing social surplus in land markets (Mayer & Sommerville, 2000; Glaeser & Gyourko, 2003; Glaeser, Gyourko, & Saks, 2005; Gyourko, Mayer, & Sinai, 2006; Quigley & Rafael, 2005; Turner et al., 2014). However, our simulation results suggest that exclusionary zoning regulations may have very slight rising impacts on residential land rents (Figure 3.11), ranging from 0.5% to 1.8%, while low-density zoning may actually lead to slight decreases in residential land rents (–0.1% to –1.1%). This is mainly because low-density zoning may limit the supply of land and housing in the short term but may have no significant impact in the long term if the city is allowed to sprawl out as assumed in the model here. In reality, physical limitations or urban growth boundaries (UGBs) restrict such spatial extension, and these factors could limit land supply in either the short or long term.

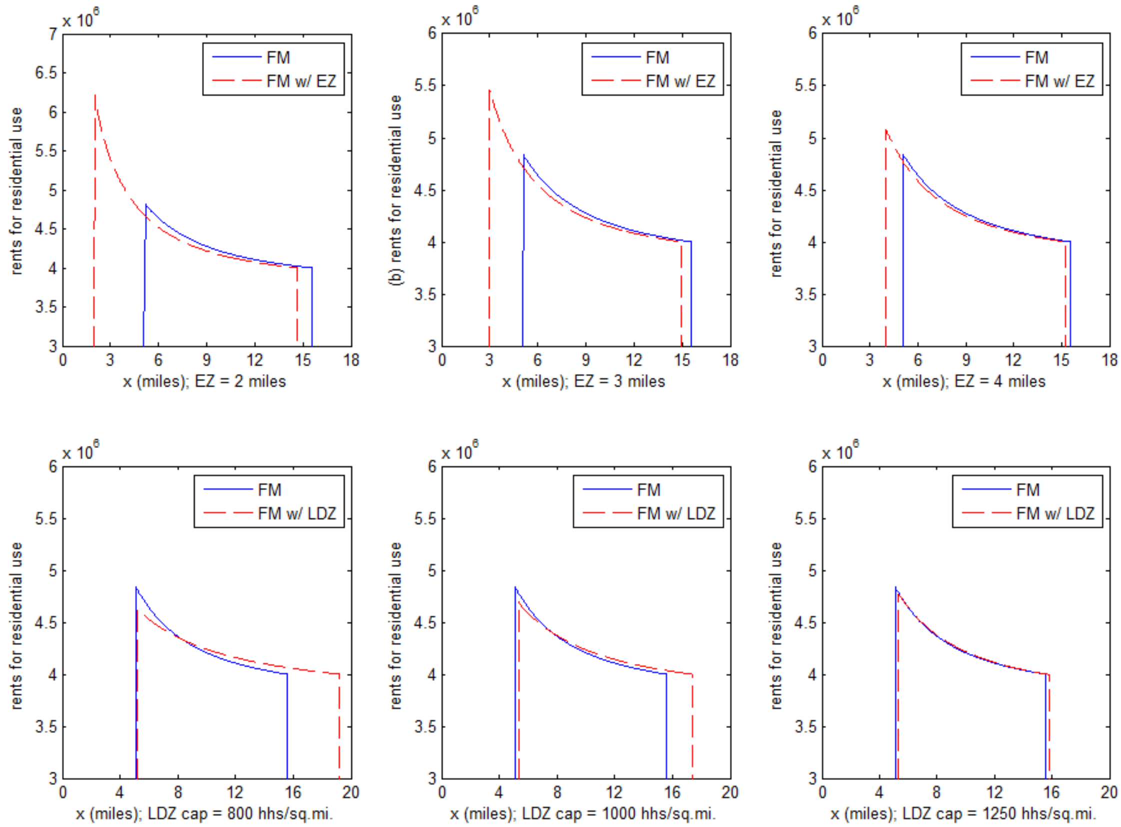


Figure 3.11 Land Rents for Residential Use in Free-Market Equilibria with and without Land Use Regulations

SUMMARY

This chapter developed and then applied a new spatial general equilibrium model to explore congestion relief, welfare, and land use effects of optimal policies in cities with market and planning failures. This new model differs from many existing studies (e.g., Fujita & Ogwa, 1982; Anas & Kim, 1996; Lucas & Rossi-Hansberg, 2002; Wheaton, 2004; Verhoef & Nijkamp, 2004; Arnott, 2007; Anas, 2012) by recognizing both congestion externalities and agglomeration externalities on production while allowing endogenous land use decisions by households and firms under land use regulations, such as exclusionary zoning and low-density zoning.

Simulation findings in this research demonstrate that congestion is born with agglomeration; increasing congestion diseconomy is associated with increasing agglomeration economy. Anti-congestion policies can reduce the congestion diseconomy but at the same time erode agglomeration economies. For example, the Pigouvian congestion toll-alone policy is no longer socially optimal once we consider agglomeration externalities, because this policy could reduce agglomeration economy more than congestion diseconomy, leading to a net loss. In some cases an imposition of congestion pricing could even bring a greater welfare loss than continuing without such pricing policies, i.e., in the free market. Thus these findings validate the importance of integrating congestion and agglomeration to assess congestion relief projects in practice.

This research also demonstrates that the socially optimal level of congestion would probably never occur under the free-flow status. Those congestion indices benchmarked at free-flow speeds, such as travel time index and annual congestion costs such as from TTI (2012), are widely used to indicate the social costs of congestion. However, these indices that assume free-flow speed is the objective could exaggerate the real costs of congestion because they overlook the potential “benefits” of congestion.

Simulations in this research show that the net social cost (i.e., diseconomy) of congestion is about 5% to 23% of the total congestion costs, varying with the levels of congestion and agglomeration.

More importantly, the research in this chapter is among the first to present an economic analysis of how planning failures cause excessive congestion. Simulation results suggest that excessive congestion and social inefficiency could be largely increased by land use regulations such as zoning excluding firms from areas outside the urban core or restricting the maximum density in residential areas. Both exclusionary zoning and low-density zoning regulations could lead to longer travel distances and more traffic volume on the roads, incurring substantial diseconomy of congestion. These negative impacts are mainly rooted in that fact that regulations constrain the occurrence of the market-desirable urban form. For example, exclusionary zoning regulations could largely restrict the decentralization of firms and jobs, while low-density zoning regulations could restrict denser development and lead to urban sprawl.

Facing both markets and planning failure, neither congestion pricing nor land use planning alone could fully reduce excessive congestion. Even the first-best pricing policy (e.g., a combination of Pigouvian congestion toll and Pigouvian labor subsidy) may not be effective for congestion relief. For example, in a simulated city with 600,000 workers and jobs, if all firms are restricted to the urban core of a 2-mile radius, even the first-best pricing policy can reduce only 4% of excessive congestion. Planning failure dominates in such a city; land use planning strategies via regulatory reform or promoting alternative development could reduce up to about 96% of excessive congestion and social inefficiency. This example is somewhat extreme but demonstrates the importance of acknowledging the role of planning failure. This study also suggests that pricing strategies are not always superior to land use planning strategies. In most cases, effective policies need to incorporate both land use and pricing policies.

However, both optimal pricing and land use policies are infeasible in practice due to corresponding technical, political, or financial issues. Allowing for more diverse and realistic policies like flat-rate tolls on freeways, cordon area congestion pricing, UGBs, densification in particular areas, and suburban centers would be meaningful. Therefore, the following two chapters focus on investigations of these practical policies.

CHAPTER 4: PRACTICAL PRICING POLICIES FOR REDUCING EXCESSIVE CONGESTION: LAND USE AND SOCIAL WELFARE IMPACTS

Optimal pricing policies for alleviating excessive congestion are seemingly never found in reality, though they are widely discussed in theory, as addressed in Chapter 3. Most applications of congestion pricing strategies are more practically feasible and second-best, rather than first-best, in the form of strategies such as cordon charges, area-wide pricing, variable-rate highway tolling, and VMT tax. Pioneering examples include Singapore's Area Licensing Scheme in the early 1970s, its Electronic Congestion Pricing policy in 1998, and London's 2003 introduction of an area-wide toll (Santos, 2005). By 2011, 10 U.S. metropolitan areas had introduced 12 high-occupancy toll (HOT) facilities on freeways, and 13 new HOT lanes were under construction or extension (GAO, 2012). Congestion pricing schemes in these regions are expected to reduce congestion, moderate negative congestion externalities (like traffic delays, air pollution, and greenhouse gas emissions), and offer revenues to fund transport system improvements, including public transit.

This chapter extends Chapter 3's model to explore the congestion relief and land use effects of two practical pricing policies, distance-based VMT taxes and cordon tolls, after controlling for these policies' effects on firms' agglomeration economies. Because pricing policies are focused on here and are insensitive to planning failures, the model discussed will account for market failure from only two externalities. The practical congestion pricing policies are compared with the first-best policies. The resulting model endogenously determines monocentric and polycentric structures, where the latter is a duocentric urban form (i.e., a center plus an annulus). In this way, the work compares the effectiveness of second-best pricing policies in monocentric versus polycentric settings. While the anti-congestion effects of congestion pricing strategies are straightforward, this chapter primarily focuses on land use effects of these congestion pricing policies. They can also help to design second-best land use policies (see more discussion in Chapter 5).

LAND USE MODELS EVALUATING CONGESTION PRICING

Congestion pricing strategies differ from many other sources of transport funding (e.g., fuel, sales, and property taxes) and can influence land use decisions rather directly, since trip charges affect travel routes, destinations, timing, and ultimately home and business location decisions. Tolls can affect firms' labor costs, productivity, and customer access. Many experts believe that a tax on vehicle-miles traveled (VMT) may accelerate new development of compact, mixed-use, walkable neighborhoods, and may modestly affect commercial land uses, especially retail (ULI, 2013). Gupta et al.'s (2006) simulations of Austin, Texas suggest that congestion pricing may catalyze land development around tolled roads, while London's area-based charge has had a somewhat negative effect on the city center's economy, particularly in retail (Santos and Shaffer 2004). Associations between congestion tolls and land use patterns in Singapore and Stockholm remain ambiguous (Bhatt, 2011; Litman, 2011).

This chapter develops modeling improvements for analyzing congestion pricing's anti-congestion and land use effects. Many studies (e.g., Brueckner, 2007; Kono and Joshi, 2012; Pines and Sadka, 1985; Wheaton, 1998) provide theoretically rigorous frameworks to explore land use patterns under Pigouvian congestion toll strategies in monocentric settings, with firms' location decisions exogenously given (i.e., all jobs are placed in the central business district, or CBD). In a city or region with only congestion externalities, Pigouvian congestion toll is a first-best policy to reflect the gap between marginal social and marginal private costs of each trip. In a closed-form monocentric model, Pigouvian congestion toll raises residential densities near the CBD, while slightly lowered edge densities (Pines and Sadka, 1985; Wheaton, 1998; Kono and Joshi, 2012).

A well-executed lot-size zoning policy can replace such Pigouvian congestion toll policies and still reach the first-best optimum, including an upward adjustment of central densities and downward adjustment of edge densities. However, these findings largely rely on the monocentric assumption and hardly reflect most regions' polycentric reality, with firm location decisions endogenous and dependent, to some extent, on household choices.

Several studies have explored the effects of first-best congestion pricing strategies in polycentric cities and their land use effects on both firm and household location choices. For example, Anas and Xu (1999) developed a spatial general equilibrium model without predetermined firm locations to explore the locational effects of Pigouvian congestion toll in a linear city with discrete zones. They found that the addition of Pigouvian congestion toll policies could disperse producers away from the regional center while centralizing households, thus bringing jobs and workers closer together. However, their model did not treat the Marshallian agglomeration economies that can cause firms to locate close to one another, arising from nonmarket interactions, and thus can somewhat misestimate congestion pricing's effects on job dispersion[11].

Several other studies have built models for continuous space, allowing more direct comparison of results to those of the traditional monocentric setting. For example, Wheaton (2004) extended a monocentric model to involve both congestion and center-agglomeration externalities, and found that higher congestion levels may cause greater job decentralization12. Though his model did not test the toll policy's efficiency, his

[11] Anas and Xu's (1999) model endogenizes firm locations and so can treat the agglomeration benefits from firms locating near their workers.

[12] Based on discrete spatial structure, an early model developed by Anas and Kim (1996) already reflects both congestion externalities and agglomeration externalities (on the producer and consumer sides). Firms are allowed to exchange inputs with each other and thus benefit from locating close to one another. Consumers are assumed to make more shopping trips to larger shopping centers, leading to retail-job agglomeration. They found similar results to these from monocentric models in that higher congestion levels may lead to larger numbers of job sub-centers.

results suggest that land use-congestion studies of this sort should not overlook interactions between congestion and agglomeration externalities[13].

Several theoretical papers have investigated the land use effects of second-best pricing in a monocentric framework (see, e.g., Mun, Konoshi, & Yoshikawa, 2003; Verhoef, 2005; De Lara, de Palma, Kilani, & Piperno, 2013). Some have sought to extend the monocentric model by involving non-monocentric features, like allowing flexible commute-trip destinations, instead of requiring that all such trips head to the CBD (Mun, Konoshi, & Yoshikawa, 2005), or positing two CBDs, instead of one (De Lara et al., 2013). Such improvements still heavily rely on the assumption that firms' location choices are exogenously given, so they cannot anticipate congestion pricing's effects on job location patterns. Recent studies relying on discrete non-monocentric settings have examined the spatial redistribution of population and employment after levying a cordon toll or instituting area pricing. For example, Fujishima (2011) extended Anas and Xu's (1999) model to compare the cordon toll and area pricing impacts and found both schemes can lead to population centralization and job dispersion in Osaka, Japan. Anas and Hiramatsu (2013) applied the RELU-TRAN model to the Chicago region, to offer a more comprehensive evaluation of cordon tolling's land use and welfare effects. Their findings suggest that restrictive cordons around Chicago's CBD may decentralize jobs, while cordons around inner suburbs may centralize jobs. Related research is less common when using urban economic models with continuous space.

[13] Other researchers tend to focus on second-best land use policies, instead of second-best pricing schemes. These include urban growth boundaries in monocentric regions (Kanemoto, 1977; Pines and Sadka,1985; Brueckner, 2007) and polycentric regions (Anas and Rhee, 2006) , and building size/floor-area-ratio regulations in monocentric regions (Pines and Kono, 2012; Kono et al., 2012).

EXTENSIONS OF CHAPTER 3'S MODEL

While the extended model developed here has the same settings of geographical context, household behaviors, and congestion as the Chapter 3's model, the major differences are the setting of agglomeration externalities and equilibrium conditions under different congestion pricing policies.

Agglomeration Externalities

A larger market may benefit more from the sharing of facilities and suppliers, a better matching between firms and workers, and the facilitation of social learning through knowledge transmission (Rosenthal and Strange, 2004; Puga, 2010). The setup used here mainly considers the agglomeration effects that come from sharing of facilities and social learning, by assuming that clustered firms benefit more from their workers' knowledge spillovers. Although the model is designed to deliver in a static, long-term spatial equilibrium, it is based on a dynamic agglomeration economy, which assumes that both current and historical economic activities at a given location affect agglomeration economies in production (Henderson, 2003; Rosenthal and Strange, 2004). Thus, $F(x)$ consists of two components:

$$F(x) = F^0(x) + F^1(x) \tag{4.1}$$

where $F^0(x)$ represents a given historical agglomeration economy that reflects the natural advantage and long-term benefits from the sharing of facilities at location *x*, and $F^1(x)$ is the current agglomeration effect at location *x*. When $F^1(x) = 0$ for any locations, $F(x)$ becomes pre-determined/exogenous, and the model collapses to a traditional monocentric model. In this paper, $F^1(x)$ is defined as the integral of exponentially distance-weighted job counts within a given boundary[14], $\bar{r}$:

[14] Fujita and Ogawa (1982) first provided a measure of agglomeration economies for firms based on job densities and distances to other firms or workers in a linear city setting (termed locational potential or communication externalities, in Fujita and Thisse [2002]). In their "LRH" model, Lucas and Rossi-

$$F^1(x) = \zeta \int_0^{\bar{r}} \int_0^{2\pi} r\theta_f(r)n(r)e^{-\zeta l(x,r,\psi)} d\psi dr \tag{4.2}$$

where ζ is exogenously determined to describe the strength/level of production externalities that exist, ψ is the polar angle around the center (ranging from 0 to 2π), and $l(x, r, \psi)$ is the straight-line distance between a firm at location *x* and any firm lying within $\bar{r}$ miles of the center (at a counter-clockwise angle of ψ from the first firm).

If $F^0(x) = 0$ for any locations and Eq. (4.2) holds, this model setting is basically equivalent to Zhang and Kockelman's (2013) model, which can achieve either monocentric or single-ring structure[15] but not polycentric urban forms. Based on Eq. (4.2), once the firm cluster shifts away from the city center, agglomeration benefits to firms near the city center fall; firms leave the centerpoint CBD and form an annulus. This annulus structure is rarely (if ever) observed in practice, mainly because of the presence of a historical agglomeration economy $F^0(x)$ at the CBD. In other words, cities evolve from small towns, so the centerpoint generally retains a long-term advantage.

Based on Eqs. (4.1) and (4.2), one can calculate the marginal production benefit to firms at location *x* of hiring an additional worker, $s(x)$,. One more worker employed in location *x* will affect the productivity of firms not only at location *x* but nearby (e.g., *r* distance away), through $F(r)$'s labor effects. As shown in Zhang and Kockelman (2014), $s(x)$ thus equals:

$$s(x) = \zeta \int_0^{\bar{r}} r\theta_f(r)p_F(r) \int_0^{2\pi} e^{-\zeta l(x,r,\psi)} d\psi \, dr \tag{4.3}$$

Hansberg (2002) extended this idea to circular space. The only difference in the current formulation (provided here) is that LRH's model considers production externalities from all firms in the entire city (inversely weighted by distance), and assumes a fixed city boundary. Our model assumes that production externalities come only from firms within a pre-set area, and the city's boundary/limit is endogenously determined.

[15] Here, single-ring structure occurs when households occupy the urban core and firms are clustered in an annulus outside this core area.

where $p_F(r)$ is the marginal product (per unit of land) $p(r)$ of $F(r)$, i.e., $\partial p(r)/\partial F(r)$. The aggregate agglomeration benefit, *S*, of firms in the city is thus as follows:

$$S = \int_0^{\bar{x}} 2\pi x \theta_f(x) n(x) s(x) dx \tag{4.4}$$

The price of firm output is set to 1.0 (as the numeraire) without loss of generality; thus, a firm's profit per unit of land at location *x*, $\Pi(x)$, can be given by the following:

$$\Pi(x) = f\big(n(x)\big)A\big(F(x)\big) - w(x)n(x) - r_f(x) \tag{4.5}$$

where $w(x)$ is the wage paid to each laborer and $r_f(x)$ is the rent firms are willing to pay (per unit of land) at location *x*.

Solving for the General Spatial Equilibria

Given the transportation parameters described above, one can combine the households' and firms' partial equilibria with equilibrium conditions for labor and land markets, thereby creating a general spatial equilibrium model for the region. Four types of spatial equilibrium are discussed here, including the no-toll (i.e., free-market) city, the Pigouvian congestion toll equilibrium, and the VMT tax and cordon toll equilibria. The existence of both congestion and agglomeration externalities increases the difficulty of comparing congestion pricing policies, since the pricing instruments can affect agglomeration economies (Verhoef and Nijkamp, 2004; Zhang and Kockelman, 2014). This paper focuses on the efficiency of tolling policies for correcting negative congestion externalities and their spatial consequences, after agglomeration externalities are corrected via a uniform labor subsidy to firms (per hired worker). The equilibrium population under the three pricing policies is set to equal those in the no-toll (base case) equilibrium.

The No-Toll Equilibrium

The no-toll equilibrium is an efficient market solution if both congestion and production externalities do not exist. Thus, given $t(x)$ and $F(x)$, the solution to a no-toll equilibrium is achieved by determining five factors, $\{n(x), q(x), c(x), \theta_f(x), D(x)\}$, at each location x, so as to maximize household utility levels under the five constraints (4.6)–(4.10), as defined in Problem 1.

Problem 4.1. Choose functions $n(x), q(x), \mathrm{c}(x), \theta_f(x), D(x)$ so as to maximize

$$u\big(c(x), q(x)\big)$$

subject to the following conditions:

$$c(x) + r_h(x)q(x) = y(x) = w(x) + \bar{y} \tag{4.6}$$

$$f\big(n(x)\big)A\big(F(x)\big) - (w(x) - \bar{s})n(x) - r_f(x) \geq 0 \tag{4.7}$$

$$\theta_h(x) + \theta_f(x) + \theta_t = 1 \tag{4.8}$$

$$D'(x) \leq 2\pi x\left(\frac{\theta_h(x)}{q(x)} - \theta_f(x)n(x)\right) \tag{4.9}$$

$$\int_0^{\bar{x}}\left\{2\pi x\left(\theta_f(x)f\big(n(x)\big)A\big(F(x)\big) - \frac{\theta_h(x)}{q(x)}c(x) - (1-\theta_t)R_A\right) - t(x)D(x)\right\}dx \geq 0 \tag{4.10}$$

for all $x \in [0, \bar{x}]$, with boundary conditions:

$$r(\bar{x}) = R_A \tag{4.11}$$

$$D(0) = 0 \; and \; D(\bar{x}) = 0 \tag{4.12}$$

$$\int_0^{\bar{x}} 2\pi x \frac{\theta_h(x)}{q(x)} dx = N \tag{4.13}$$

where R_A is the opportunity cost of land inside a city, which is assumed to equal the exogenous rent of agriculture use outside the city (as done by Pines and Sadka [1986] and

Bruckner [2007]). $r(x)$ is the highest bid-rent at location *x*, so $r(x) = max\{r_h(x), r_f(x), R_A\}$, and *N* is the exogenously given regional population total.

Constraint (4.6) is the household budget constraint. Since no toll revenue is earned, $\bar{y}_{toll} = 0$ and $\bar{y} = \bar{y}_{rent}$, where $\bar{y}_{rent}$ is set as follows:

$$\bar{y}_{rent} = \frac{1}{N}\int_0^{\bar{x}} 2\pi x(1-\theta_t)(r(x) - R_A)dx \tag{4.14}$$

Constraint (4.7) guarantees non-negative profits for each firm. A uniform/constant labor subsidy, $\bar{s}$, is paid to all firms per worker hired and the aggregate labor subsidy expended equals the equilibrium agglomeration benefit, *S*, as defined in Eq.(4.3). This per-capita subsidy, $\bar{s}$, is thus calculated as follows:

$$\bar{s} = \frac{S}{N} = \frac{1}{N}\int_0^{\bar{x}} 2\pi x\theta_f(x)n(x)s(x)dx \tag{4.15}$$

Constraint (4.8) represents land market clearance, so that all available land or properties are assigned to agents, while the city's edge rent equals the agricultural land rent, as defined in boundary condition (4.11). Constraint (4.8) guarantees that an additional number of travelers passing the infinitesimal interval *dx* (from *x*+*dx* to *x* or from *x*-*dx* to *x*), $D'(x)dx$, will not exceed the maximum travel demand generated in the interval *dx*: $2\pi x dx\left(\frac{\theta_h(x)}{q(x)} - \theta_f(x)n(x)\right)$. This constraint relates to boundary condition (4.12), in which no travel demand exists at the regional center point or at the city's edge. This ensures a city-wide jobs-housing balance. Finally, Constraint (4.9) is the output market's clearing condition. Given that aggregate land rents (net of the opportunity costs) will be returned uniformly to each household (due to the closed-city formulation, which facilitates welfare comparisons across settings, and as done in Solow [1973], Pines and Sadka [1986], Anas and Xu [1999] and Brueckner [2007], for example), the net surplus is equivalent to aggregate production minus consumption of goods produced by the firms, plus land opportunity costs, minus commuting costs. In order to arrive at a closed-form solution, the equilibrium population equals an exogenous value, N, as shown in boundary condition (4.13). The resulting solution will satisfy the following proposition:

Proposition 4.1. In a closed city with $\bar{u}$ as the equilibrium utility level, the equilibrium solution set $\{n^*(x), q^*(x), c^*(x), \theta_f^*(x)\}$ satisfies the following equations:

(a) $n^*(x) = n^*(w(x))$, and $n^*(x)$ satisfies $f_n(n^*(x)) = w(x)/A(F(x))$;

(b) $q^*(x) = q^*(w(x), \bar{u})$ and $c^*(x) = c^*(w(x), \bar{u})$, and $q^*(x)$ and $c^*(x)$ satisfy the equation set: $\begin{cases} c(x) + q(x)u_q/u_c = y(x) \\ u(c(x), q(x)) = \bar{u} \end{cases}$

(c) $\theta_f^*(x) = \begin{cases} 1-\theta_t & if\ r_f(x) > r_h(x) \\ (0,1-\theta_t) & if\ r_f(x) = r_h \\ 0 & if\ r_f(x) < r_h(x) \end{cases}$

(d) $y'(x) = w'(x) = t(x)$

Proof. Appendix A5 provides this proof.

In equilibrium, households pursue optimal good consumption, $c^*(x)$, and housing lot sizes, $q^*(x)$, by *minimizing* expenditures given the target utility level (Proposition 4.1[b]). Firms pursue optimal employment densities, $n^*(x)$, in order to *maximize* their profits (Proposition 4.1[a]). At the same time, available land and property are assigned to agents offering the highest bid rents, while city edge rents equal the background (agricultural) land rent and jobs and housing are in balance, consistent with Proposition 4.1(c). Proposition 4.1(d)'s differential equation suggests that the net-income gradient and the wage gradient both equal $t(x)$ only, since no congestion toll is levied on workers/travelers. This condition guarantees that all workers are equivalent in the eyes of each firm owner, and all firms are equivalent in the eyes of each worker.

Propositions 4.1(a)-(c) show how equilibrium values $n^*(x), q^*(x), c^*(x), and\ \theta_f^*(x)$ are only determined by $w(x)$, when given $\bar{u}$, $F(x)$, and $\bar{y}$. If the wage function is derived first, all other solution values for this no-toll equilibrium can then be generated. Moreover, if one knows $w(1)$ or $w(\bar{x})$, one can derive $w(x)$ at any other location *x*, and so derive all other solution values. This suggests that the urban equilibrium problem here can be resolved using a recursive algorithm, which searches for a unique $w(1)$ and $\bar{u}$ until the boundary conditions (4.11)-(4.13) are entirely satisfied. Following Eqs. (4.3) and Proposition 4.1's equilibrium solutions, one can derive the agglomeration economies, S_{nt}, and congestion diseconomies, Γ_{nt}, under the no-toll equilibrium.

The Pigouvian Congestion Toll Equilibrium

The Pigouvian congestion toll case represents the spatial equilibrium under a "perfect" congestion pricing policy. Here, negative congestion externalities are fully internalized in the Pigouvian congestion toll equilibrium, while the aggregate agglomeration benefit is endogenously adjusted to equal that arising in the no-toll equilibrium (i.e., S_{nt}), in order to equitably compare each policy's results. The optimization problem setup of the Pigouvian congestion toll case thus matches that of Problem 4.1 (defined above, for the no-toll case), but with an additional constraint on travel costs. By resolving this optimization problem, one can prove that the equilibrium solutions in Proposition 4.1(a)-(c) still hold, while the wage gradient in Proposition 4.1d becomes the following:

$$y'(x) = w'(x) = t(x) + \tau_{mce}(x) \tag{4.16}$$

This condition shows that the net-income and wage gradients need to cover the marginal social costs of travel, which reflect both marginal private costs (MPBs) and marginal external (delay) costs imposed on other travelers, $\tau_{mce}(x)$. A Pigouvian congestion toll, $\tau_{PCT}(x)$, equaling the marginal congestion externality $\tau_{mce}(x)$, needs to be levied on each worker/each traveler passing location *x*:

$$\tau_{PCT}(x) = \tau_{mce}(x) = \begin{cases} -\rho\sigma\left(\frac{|D(x)|}{2\pi x\theta_t}\right)^{\sigma}, if\ D(x) \leq 0 \\ \rho\sigma\left(\frac{|D(x)|}{2\pi x\theta_t}\right)^{\sigma}, if\ D(x) > 0 \end{cases} \quad (4.17)$$

Thus, this Pigouvian congestion toll instrument is an optimal policy for correcting the system's negative congestion externalities.

In a closed-form city, a lump-sum amount of congestion toll revenues, $\bar{y}_{toll}$, may be returned to each worker, such that:

$$\bar{y}_{toll} = \frac{1}{N}\int_0^{\bar{x}} \tau_{mcp}(x)D(x)dx \quad (4.18)$$

The VMT-Tax and Cordon-Toll Equilibria

In practice, second-best congestion pricing policies typically involve a cordon or area-based toll ($\bar{\tau}_{ct}$, levied at location $\bar{x}_{ct}$) or a (flat-rate) distance-based (VMT) tax (of $\bar{\tau}_{vmt}$). If $\tau(x)$ represents the congestion toll levied on each worker crossing ring *x* (positive for outward travel and negative for inward travel), the magnitudes of these two distinctive tolls can be presented as follows:

$$|\tau(x)| = \begin{cases} \bar{\tau}_{vmt} & VMT\ tax \\ \begin{cases} \bar{\tau}_{ct}, & if\ x = \bar{x}_{ct} \\ 0 & otherwise \end{cases} & Cordon\ toll \end{cases} \quad (4.19)$$

Proposition 4.2. In both the cordon-toll and VMT-tax equilibria (agglomeration externalities are corrected), if the aggregate tolling revenues cover $\frac{1}{1+\sigma}$ of the overall social costs of the congestion externality, Γ (as defined in Eq. (12)):

$$\int_0^{\bar{x}} \tau(x)D(x)\,dx = \frac{1}{1+\sigma}\Gamma \quad (4.20)$$

then, the corresponding tolling level, $\tau(x)$, is second-best optimal.
Proof. See A6 in the Appendix.

An imperfect cordon toll or VMT tax (with revenues lying below or above $\frac{1}{1+\sigma}$ of the overall congestion diseconomies) will lead to labor market distortions, where workers are overpaid or underpaid by firms, to help cover travel costs and/or tolls. Only when the toll equals the optimal level defined in Proposition 4.2 will it not distort the labor market. Proposition 4.2 also illuminates the "second-best" nature of a second-best congestion pricing, which demonstrates that such toll policies cannot (fully) correct the market failure of negative congestion externalities. The second-best optimum only corrects $\frac{1}{1+\sigma}$ (less than 1.0) of overall congestion externalities.

Based on Proposition 4.2, one can calculate the optimal VMT tax as follows:

$$\bar{\tau}_{vmt}^{*} = \frac{\Gamma}{(1+\sigma)\int_0^{\bar{x}} D(x)\,dx} \tag{4.21}$$

and, the optimal cordon toll at (exogenously given) location $\bar{x}_{ct}$ will be:

$$\bar{\tau}_{ct}^{*} = \frac{\Gamma}{(1+\sigma)D(\bar{x}_{ct})} \tag{4.22}$$

Finding optimal prices is almost any urban economic model is a challenge. In traditional monocentric models, the basic strategy uses a heuristic search method to identify $\bar{\tau}_{vmt}^{*}$ or $\bar{x}_{ct}$ and $\bar{\tau}_{ct}^{*}$, by seeking maximum utility or social surplus (Mun et al., 2003; Verhoef, 2005; De Lara et al., 2013). Proposition 4.2 provides an alternative, effective approach for non-monocentric simulations, by increasing $\bar{\tau}_{vmt}$ or $\bar{\tau}_{ct}$ until Eq. (20) is satisfied.

SYSTEM SIMULATIONS

The model system and its parameters are specified so as to yield both monocentric and polycentric structures. The general urban form is largely determined by parameters that reflect past and present contexts, such as $F^0(x)$ and $\bar{r}$, while specific land use details (like densities and distribution of firms and households) are determined mostly by other parameters. This paper emphasizes the monocentric versus polycentric urban forms in a series of policy scenario evaluations, rather than exploring how specific parameter value choices lead to different urban structures. It seeks to show how the three styles of pricing policy are likely to affect land use patterns under the monocentric and polycentric settings.

In order to achieve a monocentric urban equilibrium, one can set $F^0(x) = 0$ for any x, $\bar{r} = 1$, and $\bar{u} = 3500$ utils in the no-toll case. These settings may reflect an emerging city, where the CBD is relatively new and firm agglomeration exists in a relatively small area. For a polycentric case, we set $F^0(x)$ to be the equilibrium production externality function $F(x)$, as solved for in the monocentric no-toll equilibrium. The agglomeration limit extends to $\bar{r} = 6$, while $\bar{u}$ increases to 4000 utils. These settings will generate a sub-center ring of development/density in the suburbs. This two-center equilibrium can be understood as an evolution from the initially monocentric city, after population, jobs and utility levels grow.

Using these two city settings (mono- and poly-centric cases), four policy scenarios (a no-toll base case, an Pigouvian congestion toll case, a VMT-tax case, and a cordon-toll case) were simulated. The spatial equilibria were solved using MATLAB, following a fixed-point algorithm, as described in Chapter 3. Using Proposition 4.1, given pre-set values of $F(x)$ and $\bar{y}$, the process of finding an equilibrium corresponds to seeking an equilibrium initial wage $w(1)$ to clear all land and labor markets and to satisfy the boundary conditions defined in Eqs. (11)-(13). New $F(x)$ and $\bar{y}$ can be derived, along with a new equilibrium initial wage at the region's centerpoint, $w(1)$. The equilibrium solutions process achieves convergence when the iterations find fixed-point $F(x)$ and $\bar{y}$ values.

PRICING POLICIES IN MONOCENTRIC CITIES

Table 4.1 summarizes simulation results under different pricing regimes in a monocentric city. The four monocentric solutions rely on the same final population, of N = 711,000 workers. This baseline population was derived from the no-toll equilibrium solution, when worker or household utility levels were set to 3500 utils. Under this setup, the optimal Pigouvian congestion toll toll rates range from $0 per mile of travel at the city edge to a peak of $3.38 per mile at the fringe of the monocentric city's firm cluster (assuming 250 workdays per year). The average toll in the Pigouvian congestion toll equilibrium is $0.94 per trip-mile (since the average toll payment per worker is $2,556 per year and the average commute distance [one-way] is 5.44 miles per day). The (flat) VMT tax is computed to be $0.46 per mile (each way). The optimal cordon location and toll is calculated to be about 3.5 miles away from the city center and $1,500 per year per commuter, $120 per month, or about $6 per workday (assuming 250 workdays per year, Table 4.1).

Table 4.1 Simulation Results under Different Pricing Regimes in a Monocentric City

	No Toll	PCT	VMT Tax	Cordon Toll
Utility level, $\bar{u}$ (utils per household)	3500.00	3500.78	3500.35	3500.34
CV, ($/year/household)		129	48	41
Average commute distance per worker (miles/day)	6.06	5.44	4.83	5.45
Average travel costs ($/year/worker)	3,305	2,790	2,822	3,053
Rent revenues returned, y_{rent} ($/year/worker)	2,040	2,302	2,522	2,185
Toll revenues returned, y_{toll} ($/year/worker)	0	2,556	1,113	1,182
Labor subsidy, y_{lt} ($/year/worker)	761	637	715	722
City boundary, $\bar{x}$ (miles)	11.28	10.88	9.82	10.83
Central wage, $w(1)$ ($ per year per worker)	25,094	25,235	25,220	25,092
Central rent, $r(1)$ (million $/sq.mi.)	252	173	212	256
Jobs density, $n(1)$ (workers/sq.mi.)	190,510	130,207	159,576	193,795
Residential density at edge, $1/q(\bar{x})$ (hhs/sq.mi.)	1,806	1,805	1,807	1,805

Since both land rents and toll revenues are assumed to be uniformly redistributed across workers or households, the welfare gains per household under different pricing regimes can be calculated using the average CV change in a household's income minus any changes in the labor subsidy[16]. The welfare gain, when moving from the no-toll to Pigouvian congestion toll equilibrium case, is estimated to be $129 per household per year. Welfare gains for the VMT-tax and cordon-toll policies are $48 and $41 per household per year, and thus about 37% and 32% of the gains under Pigouvian congestion toll policies (Table 4.1).

In addition, the total diseconomy caused by excessive congestion is $129 per household per year in the free market, reducing to $81 in the VMT-tax equilibrium and $88 in the cordon-toll equilibrium (Table 4.1). These findings suggest that the two second-best pricing policies can partially reduce excessive congestion. Meanwhile, the VMT tax can lead to lower average commuting distance but higher average travel cost than the Pigouvian congestion toll policy. Similarly, the average commute distance in the cordon-toll scheme basically equals to that in the Pigouvian congestion toll scheme, while the average travel costs increase by 9.4% (from $2,790 to $3053 per household per year). These findings indicate that the VMT-tax and cordon-toll policies will generate higher levels of congestion than the Pigouvian congestion toll policy.

All three pricing strategies lead to a more compact city sizes than the no-toll equilibrium case (of city radius 11.28 miles). The Pigouvian congestion toll narrows the city boundary to 10.88 miles, a net decrease of 0.4 mile, causing an area reduction of about 28 square miles (a 7% drop in city area). The VMT tax is associated with a 1.45-mile decrease in boundary and a 24% reduction in city area, while the optimal cordon toll generates a 0.45-mile decrease of boundary and an 8% reduction of city area (Table 4.1).

[16] The welfare change calculation refers to Anas and Hiramatsu (2013), which suggested that the citywide welfare change under a cordon-toll regime consists of the utility gain of consumers (measured by compensating variation values), the gains of real estate investors (i.e., change in property values), government gains (in tolls and taxes collected), and the gain of firms (i.e., zero profits in a competitive product market). Since our model assumes a government-distributed labor subsidy, government gains equal toll revenues minus expenditure on labor subsidies.

Congestion pricing's effects on compactness are also reflected in the three policies' travel distance impacts. The average commute distance per day falls from 6.06 miles in a no-toll equilibrium to 5.44 miles in the Pigouvian congestion toll equilibrium (a 10% drop), 4.83 miles in a VMT-tax case (a 20% drop), and 5.45 miles in a cordon-toll case (a 10% drop) (Table 4.1). These are practically very significant changes in residents' travel patterns, and are reflected in the land use patterns.

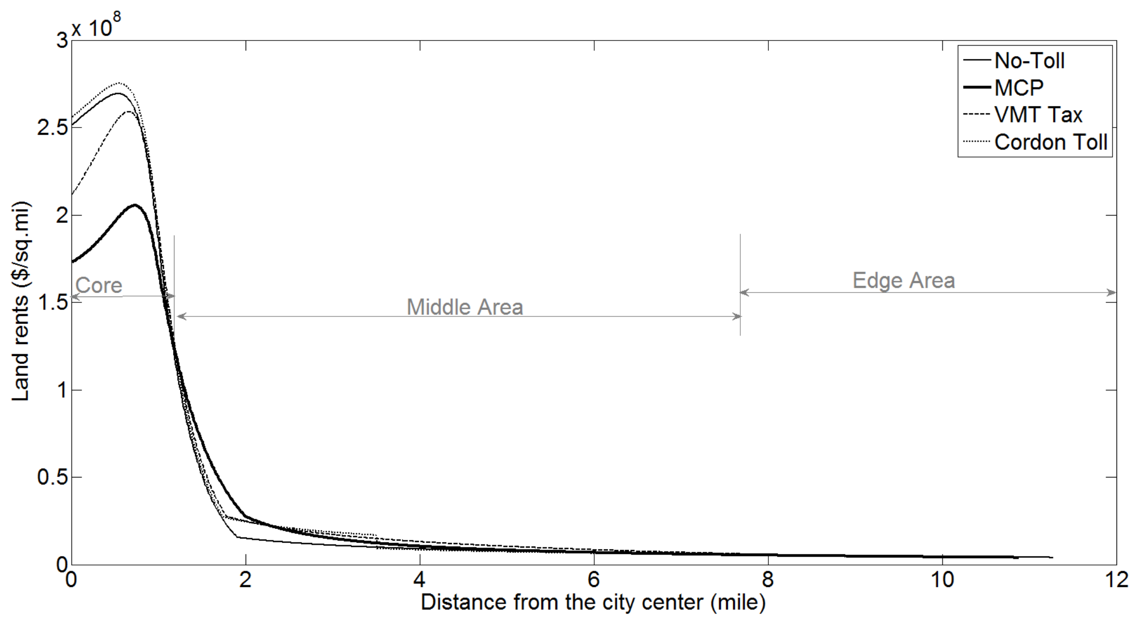


Figure 4.1 Land Rent Distribution under Different Pricing Policies in a Monocentric Setting (N=711,000, $\bar{r} = 1$)

Equilibrium land rents vary across policies and locations (Figure 4.1). To facilitate this discussion, we separate the monocentric city into three areas: an urban core area near the city center, an edge area near the city boundary, and a middle area between these two. In the urban core, both the Pigouvian congestion toll and the VMT tax bring a sharp decrease in land rents. The center-point rent, which is about $252 million per square mile (or $6M per acre) in the no-toll equilibrium, falls to $273 million (a 31%

drop) in the Pigouvian congestion toll equilibrium and $212 million (a 16% decrease) in the VMT-tax equilibrium. Peak land rents (almost one mile from the regional center) also decrease: from $269 million to $206 million (a 23% drop) after the Pigouvian congestion toll, and to $259 million (a 3.7% drop) after an optimal VMT tax. In contrast, a cordon toll causes a modest increase in land rents in the urban core: about 1.7% higher center-point and peak values. In most of the central area, the Pigouvian congestion toll and the VMT tax significantly elevate land rents: The Pigouvian congestion toll equilibrium is estimated to raise land rents by up to 91%, and the VMT tax raises rents by up to 64%. Since the optimal cordon locates at the middle area, land rents within the cordon are up to 62% higher, and those outside the cordon drop sharply, falling below no-toll rents in various locations. In the edge area, all three pricing scenarios cause a slight rent decrease. These findings in the middle and the edge areas are consistent with those found in traditional monocentric models (Verhoef, 2005; De Lara et al., 2013). But the spatial distribution of land rents in the urban core area is less commonly observed, since most monocentric models regard the urban core as an exogenously determined CBD.

Third, congestion pricing policies have significant impacts on firms' equilibrium distributions. Figure 4.2 compares job densities or firm distributions under different pricing scenarios [17]. The Pigouvian congestion toll policy increases the per-mile commuting costs and thereby encourages firms to decentralize, to locate closer to their workers. Thus, the Pigouvian congestion toll policy largely decreases central job densities but raises job densities near the edge of the firm cluster, indicating that firms are less agglomerated and jobs are more decentralized. Similar shifts emerge under the VMT-tax

[17] Regardless of pricing policies, land rents and job densities increase with distance near the center (at about 0.8 miles from the centerpoint). These spatial consequences relate closely to trends in technology and wage levels. A location with better technology and/or wage levels will attract more jobs/firms, thus raising job densities and bid rents. Near the city center, wage levels fall with radial distance and technology levels rise and peak at a location about 0.8 miles from the center. These factors cause a rising trend of rents and job densities near the city center.

equilibrium, though the drop in central job density is weaker. The Pigouvian congestion toll policy gives the market a clear signal that commuters need to pay the social costs of congestion if the city is over-concentrated, so a dispersal force emerges against agglomeration economies. As compared to the no-toll equilibrium, the Pigouvian congestion toll equilibrium encourages firms and jobs to decentralize toward the edge of the monocentric region’s firm cluster. The VMT tax generates a similar job decentralization, since some firms may desire to pay fewer VMT taxes for their workers (in the form of higher wages), and decide to relocate away from the central area for proximity to their workers.

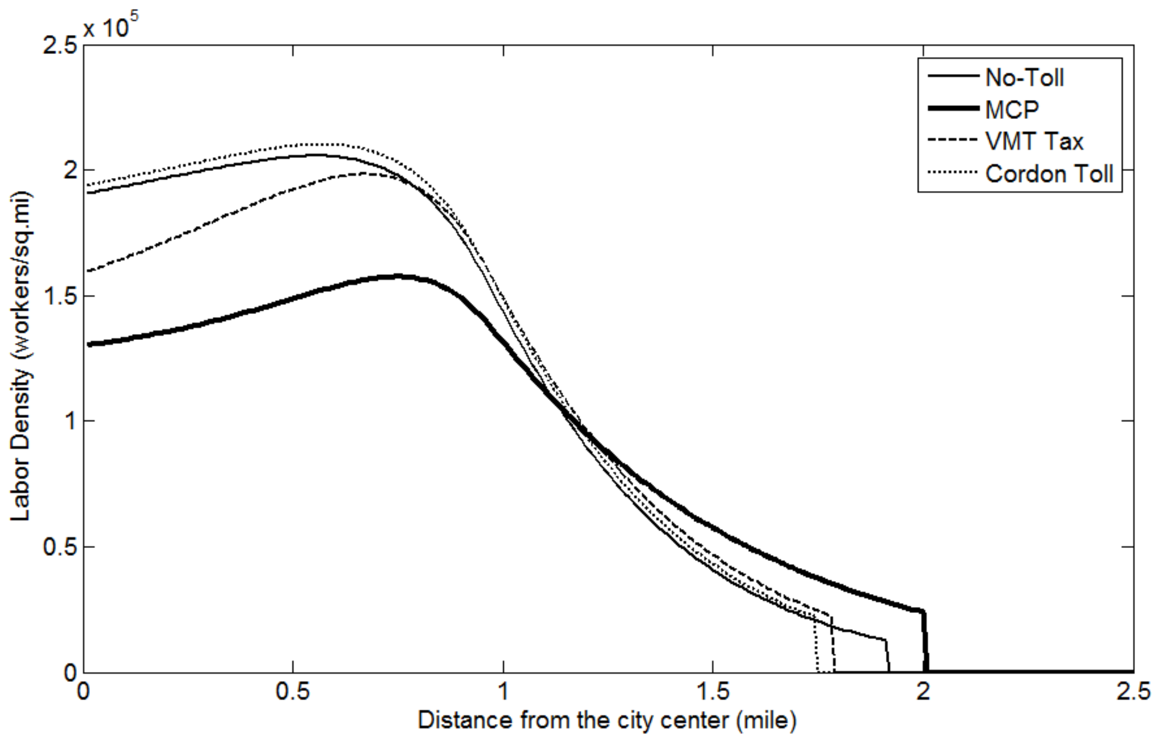


Figure 4.2 Firm Distribution and Job Densities under Different Pricing Scenarios in a Monocentric Setting

While approximations of Pigouvian congestion toll and VMT-tax policies exist in some corridors and cities, area or cordon charge schemes are simpler to apply and more popular, especially in cities with very strong CBDs, like Singapore and London. One interesting question for practice is how a cordon toll affects a CBD's economy (Santos and Shaffer 2004; Bhatt, 2011; ULI, 2013). Our simulation suggests that, in a monocentric setting, the optimal cordon toll tends to create more firm agglomeration, with slightly higher densities in a smaller area, but these effects appear practically insignificant (Figure 4.2). These results reveal that firm and labor markets are probably more sensitive to Pigouvian congestion toll and the VMT-tax policies than to area or cordon charges. However, if the cordon is not set at or near the theoretically optimal location (at 3.5 miles) and the toll is far from optimal, such tolls may encourage firms near the CBD's edge to move just outside the cordoned area (as shown in Figure 5.3, when the cordon is placed at 2.2 miles). In such settings, households may provide higher bid rents (to avoid regular commute charges) than firms can just inside the cordon line. In such cases, a cordon toll creates an "edge" effect for firm and household location choices.

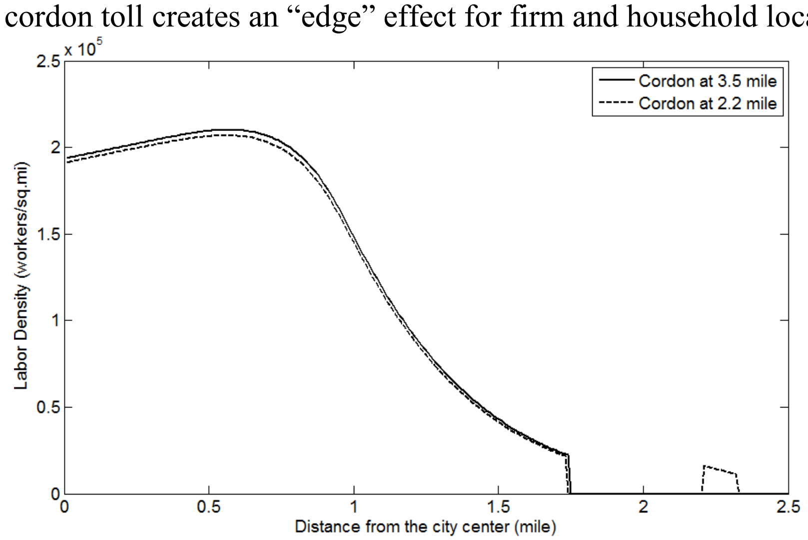


Figure 4.3 Effects of Different Cordon Locations on Firm Distribution in a Monocentric Setting.

The travel pricing effects on household locations and residential densities found here are similar to those found in most monocentric modesl (Wheaton, 1998; Verhoef, 2005; Kono and Joshi, 2012; De Lara et al., 2013). According to Figure 4.4, Pigouvian congestion toll tolling shifts market population densities down near the city edge, and upward near the central firm cluster. The VMT tax causes similar effects, with a relatively flat population density gradient near the center and a relatively sharp density gradient near the edge. In addition, the cordon charge generates a dramatic drop or "plummet": residential densities in the area inside the cordon area are quite high, while those just outside the cordon fall off sharply

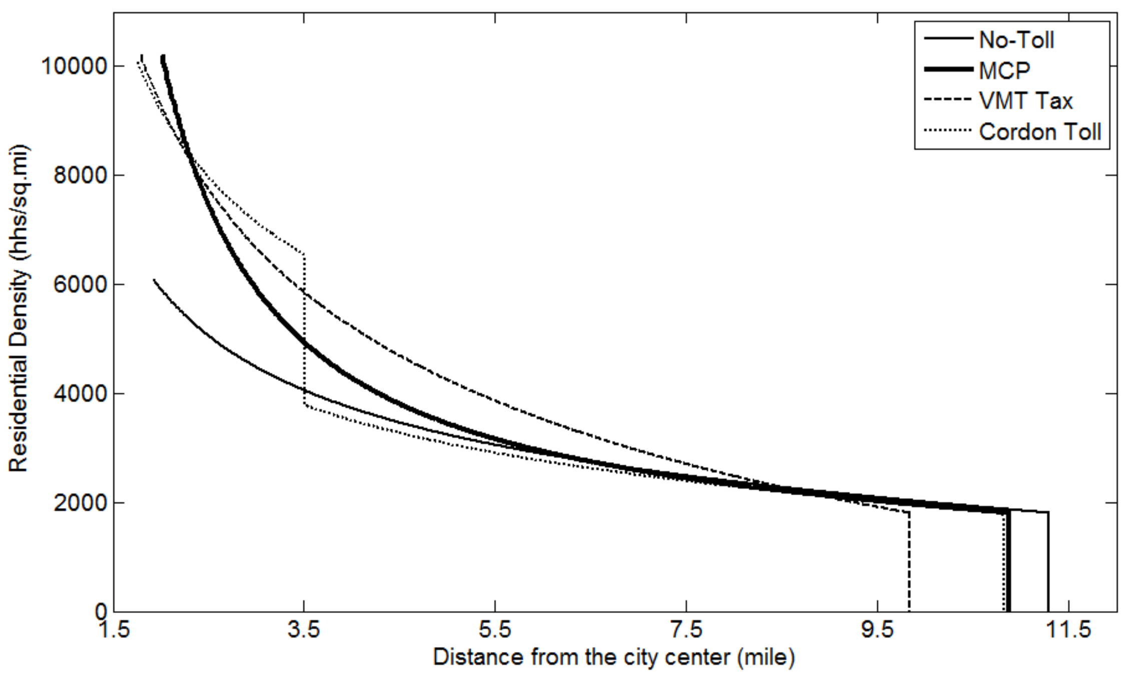


Figure 4.4 Household Distribution and Residential Densities under Different Pricing Scenarios in a Monocentric Setting

PRICING POLICIES IN POLYCENTRIC CITIES

The no-toll polycentric setting's urban form was endogenously determined after assuming household levels to be 4000 utils, yielding a population of N = 1,048,000 workers (Table 4.2). The no-toll equilibrium yields two city "centers" or densely developed rings of clustered firms. The first firm cluster, at the city center, is referred to here as the "traditional CBD", while the second, in the suburban area (about 4.5 to 6.5 miles away from the center), is called the region's sub-center. Simulation results suggest that the optimal Pigouvian congestion toll tolls rise as high as $3 per mile of travel, while the average Pigouvian congestion toll across locations is $0.71 per mile. In addition, the optimal VMT tax is computed to be $0.40 per mile of travel. The cordon toll's optimal location is found to be about 2 miles away from the city center[18], with an optimal cordon fee of $1,210 per year per worker – roughly $5 per workday, or $100 per month.

The utility values rise about 0.86% in the Pigouvian congestion toll equilibrium, 0.59% in the VMT-tax equilibrium, and up to 0.58% in the cordon-toll equilibrium (Table 4.2), while the corresponding welfare gains are estimated to be $156, $59, and $41 per household per year, respectively. These amount to 0.54%, 0.21%, and 0.14% of the average net income. The VMT-tax and cordon-toll policies can reduce the diseconomy of excessive congestion by about 38% and 26%, respectively.

Similar to the monocentric setting, the polycentric city solutions becomes more compact after Pigouvian congestion toll and VMT taxes are imposed (Table 4.2). The city boundary distance falls from 14.76 miles in the no-toll equilibrium to 14.40 miles in the Pigouvian congestion toll case (a 4.8% drop in total city area) and 13.88 miles in the VMT-tax case (a 6.7% drop in area). The cordon toll policy appears to slightly expand

[18] Cordon locations between 2 and 2.5 miles generate nearly constant maximized utility levels, based on the solution routine's simulation accuracy. Thus, without loss of generality, we chose 2 miles for the optimal cordon location.

the city, rather than restrict it, with a 1.8% increase in city area. Such pricing policies also reduce average travel distances in the polycentric region, by 20%, 18%, and 11% under the Pigouvian congestion toll, VMT tax, and cordon toll cases, respectively, relative to the no-toll base case.

Table 4.2 Simulation Results under Different Pricing Regimes in a Polycentric City

	No-Toll	PCT	VMT Tax	Cordon Toll
Utility level, $\bar{u}$ (utils per household)	4000.00	4034.58	4023.56	4023.28
Average CV, relative to No Toll case ($ per worker per year)		156	59	41
Average travel distance per worker (miles/day)	6.27	4.97	5.16	5.57
Average travel costs ($/year/worker)	2,446	1,582	1,898	1,801
Rent revenues returned, y_{rent} ($/year/worker)	2,256	2,241	2,487	2,053
Toll revenues returned, y_{toll} ($/year/worker)	0	883	518	418
Labor subsidies, y_{ls} ($/year/worker)	527	602	625	642
City boundary, $\bar{x}$ (miles)	14.76	14.4	13.88	14.89
Central wage, $w(1)$ ($ per year per worker)	28,504	29,070	28,663	28,844
Central rent, $r(1)$ (million $/sq.mi.)	254	173	228	201
Jobs density, $n(1)$ (workers/sq.mi.)	169,510	113,081	151,183	132,685
Residential density at edge, $1/q(\bar{x})$ (hhs/sq.mi.)	1588	1566	1571	1571
Average rent in the CBD ($M/sq.mi.)	115.73	98.95	83.58	111.89
Average rent in the sub-center ($M/sq.mi.)	18.99	24.26	22.42	22.94
Jobs in the sub-center (1,000)	470	648	530	659
Percentage of jobs in the sub-center (%)	44.83	61.8	50.53	62.89
Job density in the CBD (workers/sq.mi.)	54,373	45,704	55,791	51,537
Job density in the sub-center (wrkrs/sq.mi.)	9,888	12,716	11,704	11,926

In this two-center city, tolling policies cause interesting effects on land rent distributions. A major tendency is for central-area/CBD land rents to fall significantly, while sub-center land rise (Figure 4.5). The average CBD rent falls by 15%, 28%, and 3.3% under the Pigouvian congestion toll, VMT-tax, and cordon-toll equilibria, respectively (Table 4.2). Meanwhile, the average rent in the sub-center increases by 28%, 18%, and 21% in the Pigouvian congestion toll, VMT-tax, and cordon-toll schemes. All available land outside the CBD and the sub-center goes to housing. The land rent effects for housing in the polycentric setup are similar to those discussed above, for the monocentric cases. Under the Pigouvian congestion toll and VMT-tax schemes, residential land rents rise either in the area between the CBD and the sub-center or in the area near the sub-center, dropping near the city edge. Under the cordon-toll equilibrium, residential land rents inside the cordon area mostly rise, while those outside the cordon line fall.

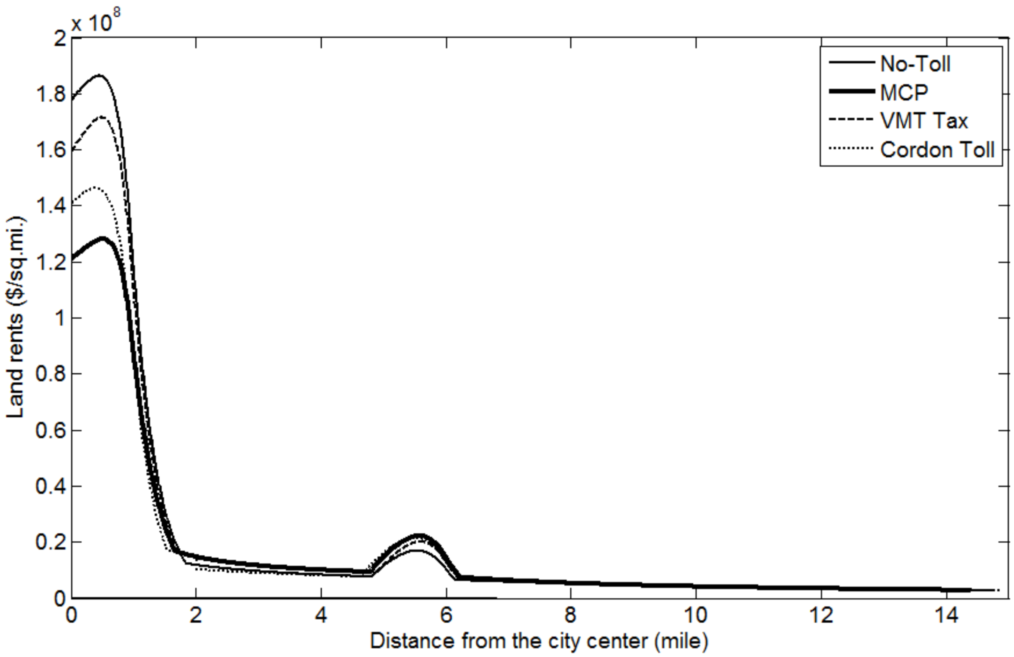


Figure 4.5 Land Rent Distribution under Different Pricing Policies in a Polycentric Setting (N = 1,048,000, $\bar{r} = 6$)

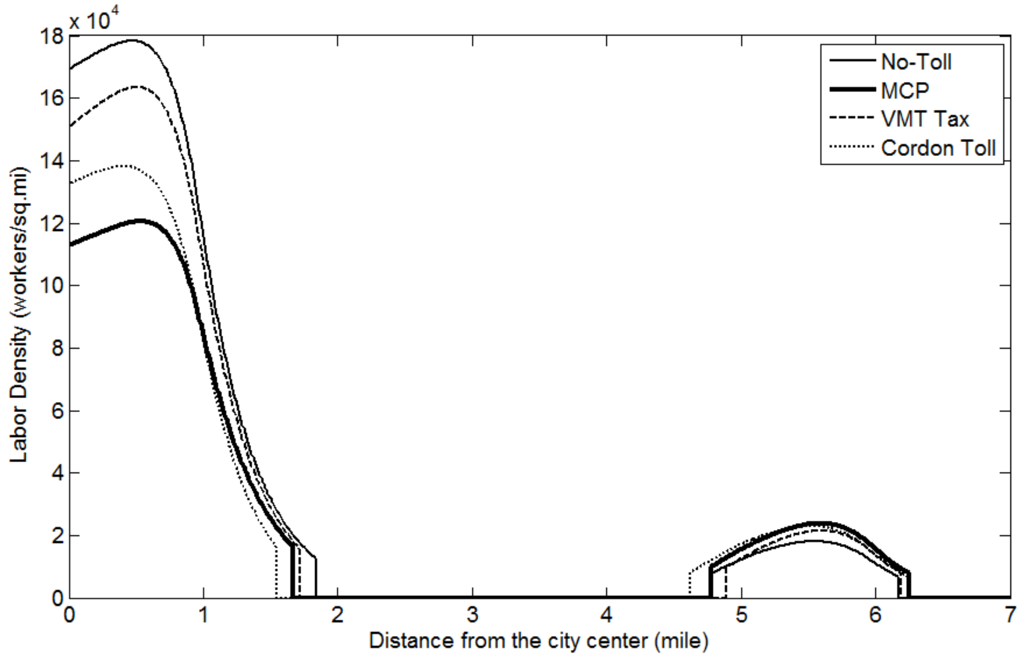


Figure 4.6 Firm Distribution and Job Densities under Different Pricing Policies in a Polycentric Setting (N = 1,048,000, $\bar{r} = 6$)

Figure 4.6 shows the distinct tendency toward job decentralization after the implementation of pricing policies. In the no-toll equilibrium, about 55% of jobs locate in the CBD and 45% in the sub-center. The Pigouvian congestion toll scheme causes about 17% of jobs to move outside the CBD and relocate at the sub-center. Levying a VMT tax is associated with a 5% increase in sub-center jobs, while the cordon toll is associated with an 18% increase in sub-center jobs. These job-decentralization effects are similar to those found in Fujishima (2011) and Anas and Hiramatsu (2013), though those two studies rely on a rather different modeling framework. Pricing also tends to significantly lower CBD job densities, while raising sub-center job densities (Figure 4.5): average CBD's job densities are computed to fall 16% and 5.2% under the Pigouvian congestion toll and cordon-toll equilibria (versus the no-toll base case), but rise 2.6% in the VMT-tax case (Table 4.2). This VMT-tax result emerges because, while a number of firms depart the center, those remaining in the CBD become more agglomerated (so the CBD's area becomes smaller). In addition, the average *sub-center* job densities rise 29%, 18%, and 21% in the Pigouvian congestion toll, VMT-tax, and cordon-toll equilibria (versus the base case). Firms leaving the CBD will enhance agglomeration economies in the sub-center areas.

Pricing's effects on residential densities are similar to those discussed earlier, for the monocentric setting. Policymakers' and planners' residential density targets in a polycentric city will presumably need upward adjustment near the city center, and downward adjustment near the city boundary (Figure 4.7). According to Table 4.2, the average residential density slightly decreases after an imposition of one of these three pricing policies (around 1%).

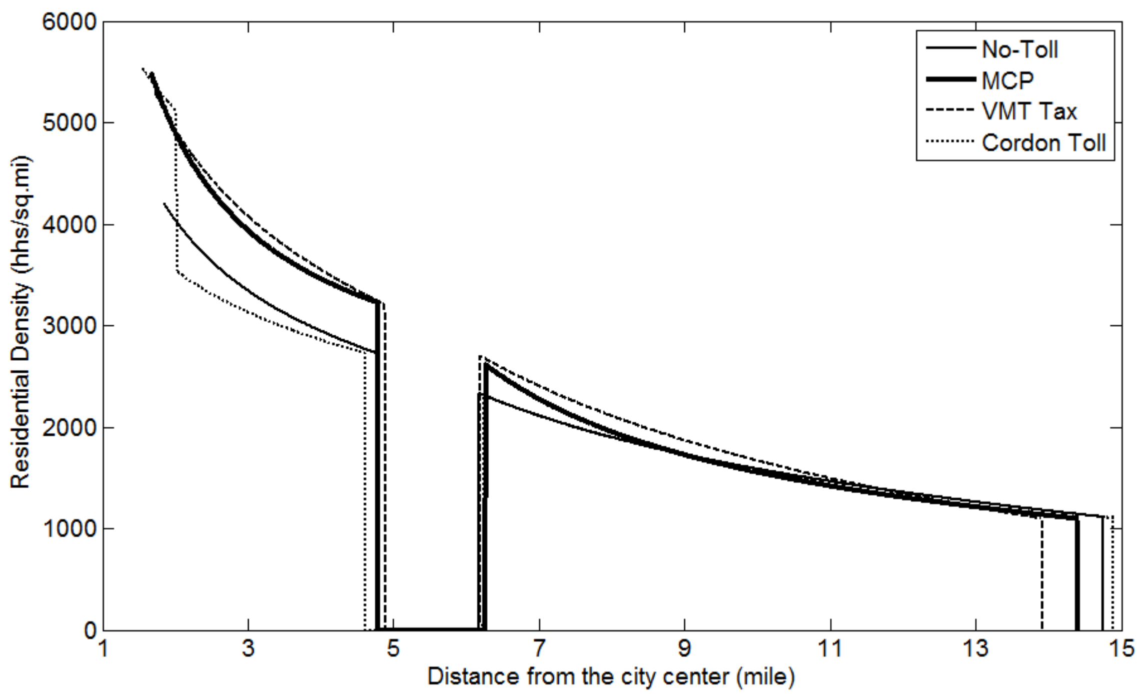


Figure 4.7 Household Distribution and Residential Densities under Different Policies of Congestion Pricing in a Polycentric Setting (N = 1,048,000, $\bar{r} = 6$)

SUMMARY

Relying on internalized congestion and agglomeration externalities, this chapter examined three pricing policies—Pigouvian congestion toll, VMT tax, and a cordon toll, alongside a no-toll base case—and compared their land use, travel, and rent impacts under both monocentric and polycentric settings. The practical pricing policies like VMT tax and cordon toll partially reduce excessive congestion: around 30% of total excessive congestion from simulations in this research. They can also produce significant decreases in average commute distance and travel costs.

The simulation results reveal that all pricing policies deliver more compact city/regional forms. Both the VMT tax and the cordon toll can generate somewhat higher household utility, although their welfare improvement is less than that of the Pigouvian congestion toll policy, as expected. The VMT tax is predicted to generate a more compact urban form than the Pigouvian congestion toll policy by incentivizing firms and households to locate closer together to reduce commuting distance, while the Pigouvian congestion toll may allow firms and/or households to trade a longer travel distance for less congestion. The compactness effects are also reflected in the findings that all three congestion pricing policies can reduce daily travel distance by more than 10% (with results ranging from 10% to 20%, varying across settings and policies).

The Pigouvian congestion toll scheme's land use patterns are more efficient than those in a free (no-toll but congestible) market. In the closed-form polycentric-city setting, efficient land use regulation may promote some job decentralization from the CBD to subcenter locations (because simulations showed more than 17% of the CBD-ring jobs moving to the suburban jobs ring). Regulation recommendations for residential densities in a polycentric city are similar to those for the monocentric setting: raise central-area population densities and reduce edge densities. The VMT tax results are not too far from those of the Pigouvian congestion toll and should be much easier to achieve in practice; unfortunately, no pricing policy is easy to get right, especially in the context of heterogeneous regions and travel plans that regularly shift (from day to day and year to year). Cordon or area tolls are more popular in practice and estimated here to have different impacts on firms when moving from a monocentric to polycentric setting. In the monocentric case, the cordon toll raised most CBD-area job densities, while a cordon line near the edge of a polycentric city's central ring may cause significant CBD-area job loss (18% simulated here). While both first-best and second-best congestion pricing strategies can lead to a significant change in city land use patterns, the following chapter will focus on practical land use planning strategies for reducing excessive congestion.

CHAPTER 5: LAND USE PLANNING FOR REDUCING EXCESSIVE CONGESTION: REMEDIES FOR MARKET AND PLANNING FAILURES

This chapter relies on the urban economic model developed in Chapter 3 to evaluate the efficiency of second-best land use policies for reducing excessive congestion. This chapter includes two parts. The first part assesses land use planning strategies in cities with market failures only and examines the welfare and anti-congestion effects of second-best land use policies such as urban growth boundaries (UGBs) and firm cluster zoning. The effectiveness of UGB policies on congestion relief remains ambiguous. Some studies have suggested that imposition of a UGB may be an effective second-best policy to reduce excessive congestion because a UGB increases densities and reduces travel distances, much like optimal pricing will do (Pines & Sadka, 1985), while others have argued that UGBs have a lower, or even negative, welfare impact than Pigouvian congestion toll strategies (Anas & Rhee, 2006; Brueckner, 2007; Kono et al., 2012). Another debate concerning UGB regulation is whether such boundaries facilitate central-city revitalization via raising productivity and attracting new development activities (Nelson et al., 2004). For comparison, this research also discusses another coarse land use policy by designating a cluster zone exclusively for firm/business use, i.e., firm cluster zoning. Our simulation results suggest that optimal firm cluster zoning policies cause a much greater welfare improvement than the optimal UGB policy. Such questions and comparisons relate closely to planning practices and planning debates and so merit exploration here.

The second part of this chapter extends the research to explore the second-best land use policies for congestion relief when land use regulations cannot be totally removed (due to political or property rights issues). Two policies are discussed here: the residential densification policy in a particular suburban area, which is against low-density zoning regulations, and job decentralization policies, such as planning a new employment center in the suburbs to decentralize firms and jobs and reduce congestion in urban areas.

LAND USE PLANNING FOR CORRECTING MARKET FAILURES

Although first-best interventions presumably are the best choice for a city authority wishing to pursue welfare improvements (Chapter 3), they may be associated major construction and operations costs (for variable toll collection, for example) that are generally not internalized in theoretical models. And a combination of Pigouvian congestion toll and Pigouvian labor subsidy may require much coordination between transportation agencies and departments of labor, which presents added transaction costs and political difficulties. A first-best tolling/subsidy-only may reduce the need for coordination, but optimal toll/subsidy levels for each location are difficult to set.

This section turns to the welfare and land use effects of second-best land use policies that are easier performed in planning practice in including UGBs and a novel firm cluster zoning policy. The UGB policy is a land-use regulation without any pricing adjustments, where the fixed-land-rent assumption at the city edge is replaced by fixing a city boundary, $\bar{x}_{ugb}$. The firm cluster zoning policy imposes a relatively idealistic land use zoning regulation by designating one or more cluster areas for firm use only and the rest areas for residential use. While the UGB policies have been applied in several cities, the firm cluster zoning policies appear less discussed in theory and practice. In fact, many cities have implemented zoning policies close to firm cluster zoning, such as planning for industrial parks and/or high-tech development zones.

For modeling specification, after imposing a UGB in a free-market city, the condition of edge rent in Eq. (3.18) is replaced by the fixed-boundary condition:

$$\bar{x} = \bar{x}_{ugb} \tag{5.1}$$

Similarly, after imposing firm cluster zoning boundaries which begin at location x_0 and end at location x_1, the equilibrium share of firm's land use at location x is defined as follows:

$$\theta_f^z(x) = \begin{cases} 1-\theta_t, x \in [x_0, x_1] \\ 0, others \end{cases} \quad (5.2)$$

Urban Growth Boundary Policies

Table 5.1 Welfare and Land Use Effects of Optimal UGB Policies in Cities with Varying Congestion and Agglomeration Levels

	Congestibility Parameter ρ			Agglomeration Parameter γ		
	5e-6	1e-5	3e-5	0.08	0.06	0.04
Types of Urban Form at FM equilibria	FH	FH	HFH	FH	FH	HFH
CV of UGB Policies (relative to FM cases, \$/hh./year)	22.11	9.13	15.33	4.71	9.13	23.76
% UGB CV relative to the First-Best CV	12.95	8.02	8.06	2.4	8.02	36.29
Percent Change in Avg. Labor Density (from FM to UGB)	7.16	0.79	9.46	1.23	0.79	9.42
Percent Change in Avg. Residential Density (from FM to UGB)	6.26	4.38	10.79	5.17	4.38	9.23
Percent Change in Avg. Rent for Firms (from FM to UGB)	7.16	0.8	9.50	1.23	0.8	9.26
% UGB Business Rent Rise relative to the First-Best Rent Rise	6.95	6.03	19.13	1.57	6.03	22.47
Percent Change in Avg. Rent for Housing (from FM to UGB)	7.2	5.17	12.35	5.7	5.17	10.42
% UGB Housing Rent Rise relative to the First-Best Rent Rise	309.0	134.6	196.0	131.6	134.6	198.1

Although there are no analytical solutions to the optimal location of UGBs, our simulation results demonstrate that the optimal UGBs should be located at the equilibrium boundary of the first-best optimum[19]. The optimal UGB policies can improve citywide welfare and the welfare gains range from 2.4% to 36% of the firm-best optimum. The UGB policies appear more effective in the cities with relatively lower agglomeration scale or higher congestion levels. Under the base scenario ($\rho = 0.00001$

[19] To find the optimal UGBs in simulations, we applied a bisection algorithm to search an optimal location for UGBs in the interval $[2\bar{x}_{fb} - \bar{x}_{fm}, \bar{x}_{fm}]$. Here, $\bar{x}_{fb}$ is the optimal boundary in the first-best case and $\bar{x}_{fm}$ is the equilibrium boundary in the free-market case.

and γ=0.06), the CV of the UGB policy relative to the free-market case is 8% of the first-best CV level. When γ decreases to 0.04, the CV gain increases to 36% of the corresponding first-best level, though the CV value is still low, at about $24 per household per year.

In addition, the UGB equilibrium could produce worse land market distortion than the free-market equilibrium. Figure 5.1 compares the spatial patterns of job and residential densities and land rents for firm and residential use in the UGB, first-best, and free-market equilibria under varying agglomeration parameters. The UGB policies could largely raise residential densities and escalate residential rents over the optimum levels at most locations, regardless of urban forms. The average residential rents under the optimal UGB policies are more than 30% larger than those under first-best instruments (Table 5.1). For $\rho = 0.000005$, the average residential rent in the UGB equilibrium is even three times that in the first-best optimum. In addition, UGBs can slightly centralize firms, leading to a trivial increase in productivities. But the increases in job densities and firms' rents caused by the optimal UGB policies are much smaller than those by the first-best instruments. Thus, restrictive UGBs appear have less significant impact on firms' spatial distribution and rents but excessively raise residential densities and rents. This may explain why even the optimal UGB regulation gains a relatively low welfare improvement.

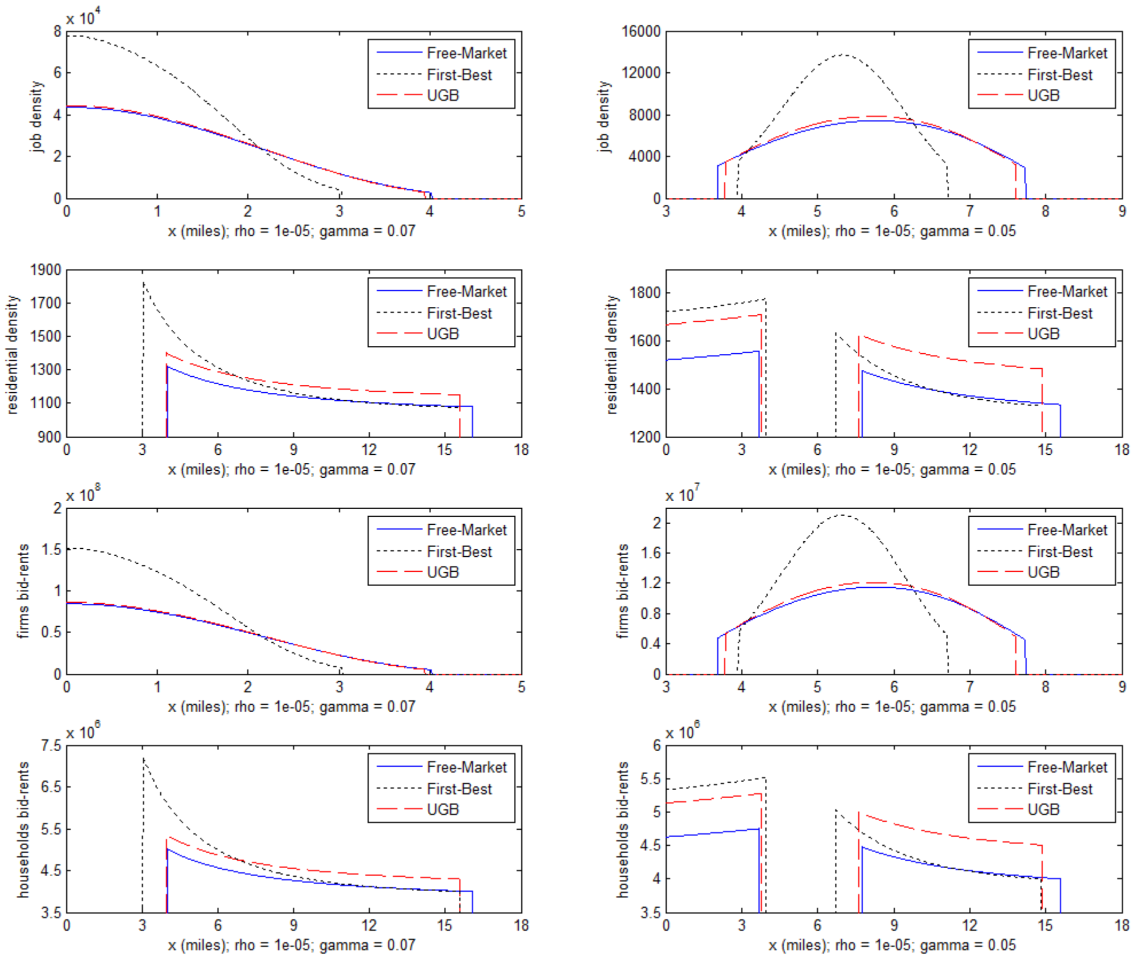


Figure 5.1 Spatial Distributions of Job and Residential Densities and Land Rents for Firm and Residential Use in the UGB, First-Best, and Free-Market Equilibria Varying between the FH(i.e., monocentric, Left) and HFH (i.e., Nonmonocentric, Right) Urban Forms (φ=20, ρ=0.00001, σ=1.5, γ=0.05 or 0.07, N=600,000)

Firm Cluster Zoning Policies

The optimal firm cluster zoning policies need to delimit a zone exclusively regulated for firm use (commercial and industrial), with the zonal boundaries setting at the locations of the firm cluster in the first-best optimum20. Simulations suggest that the

[20] There is no analytical solution to the optimal firm cluster zoning setting. The optimal firm cluster zoning defined here is derived from simulations.

optimal firm cluster zoning policies are more effective than the optimal UGB policies. Taken the base scenario ($\rho = 0.00001$ and γ=0.06) as an example (Table 5.2), the CV of the optimal firm cluster zoning instrument relative to the free-market equilibrium is 73% of the first-best CV level, above eight times the UGB level. The larger the agglomeration parameter γ, the larger welfare gain the firm cluster zoning policies can obtain. In the highly agglomeration case (γ=0.08), the firm cluster zoning welfare gain is about 48% of the first-best level, but about 20% of the UGB level. These findings suggest that the firm cluster zoning policy is more likely to be a second-best policy for correcting both agglomeration and congestion externalities than the UGB policy. However, such an effective policy appears less discussed in the literature. The major reason is probably related to the fact that urban economic analysis remains heavily relying on monocentric models, which are unable to explore firms' spatial behaviors and land use regulations on firms.

Differing from the UGB policy, the firm cluster zoning policy can to some extend remedy for land market distortion by raising both residential and commercial rents closer to the first-best levels. In particular, the average commercial rents in the firm cluster zoning equilibrium are very close to the optimum level, regardless of the values of ρ and γ. Though the average residential rents in the firm cluster zoning equilibrium are much smaller than the first-best levels, they remain larger than the free-market levels. More importantly, the firm cluster zoning policies will not lead to an excessive escalation in residential rents as the UGB policies do, and thus will not worsen the land market distortion due to the existence of urban externalities. Figure 5.2 shows the spatial patterns of densities and rents in the first-best, free-market, and firm cluster zoning equilibria. The distributions of job densities and commercial rents in the firm cluster zoning equilibrium are closer to the optimum while the allocation of residential densities and rents are closer to the free-market equilibrium. These findings suggest that the firm cluster zoning policies have significant effects on firms' spatial behaviors but fewer effects on household's residential decision. While the UGB policies appear benefits firms more, the firm cluster zoning policies may benefit residents and commuters more.

Table 5.2 Welfare and Land Use Effects of Optimal Firm Cluster Zoning Policies in Cities with Varying Agglomeration Levels

	Congestibility Parameter ρ			Agglomeration Parameter γ		
	5e-6	1e-5	3e-5	0.08	0.06	0.04
FCZ Policies						
CV of FCZ Policies (relative to FM cases, $/hh./year)	138.58	83.12	161.68	93.51	83.12	52.69
% FCZ CV relative to the First-Best CV	81.14	73	85.03	47.73	73	82.49
Percent Change in Avg. Labor Density (from FM to FCZ)	101.44	14.04	49.89	76.17	14.04	41.71
Percent Change in Avg. Residential Density (from FM to FCZ)	1.33	0.39	-9.93	-8.94	0.39	-9.13
Percent Change in Avg. Rent for Firms (from FM to FCZ)	102.05	13.13	49.36	76.89	13.13	40.83
% FCZ Business Rents relative to the First-Best Rents	98.99	99.02	99.40	98.10	99.02	99.08
Percent Change in Avg. Rent for Housing (from FM to FCZ)	0.41	0.28	0.09	0.32	0.28	0.28
% FCZ Housing Rents relative to the First-Best Rents	17.60	7.29	1.43	7.39	7.29	5.32

By recognizing this advantage, those planner and policy makers who search for land use policies for reducing congestion and enhance agglomeration should consider the of firm cluster zoning policies. One difficulty for such firm cluster zoning policies should be the determination of the optimal cluster zone. Our simulations suggest that such an optimal firm cluster zoning area could be more centered (e.g., γ=0.05) or decentered (e.g., γ=0.06). If a city prefers to enhance agglomeration economies, the optimal firm cluster zoning area should be more compact than the free-market firm cluster area. Instead, if a city prefers to reduce congestion diseconomies, the optimal firm cluster zoning should allow for job decentralization.

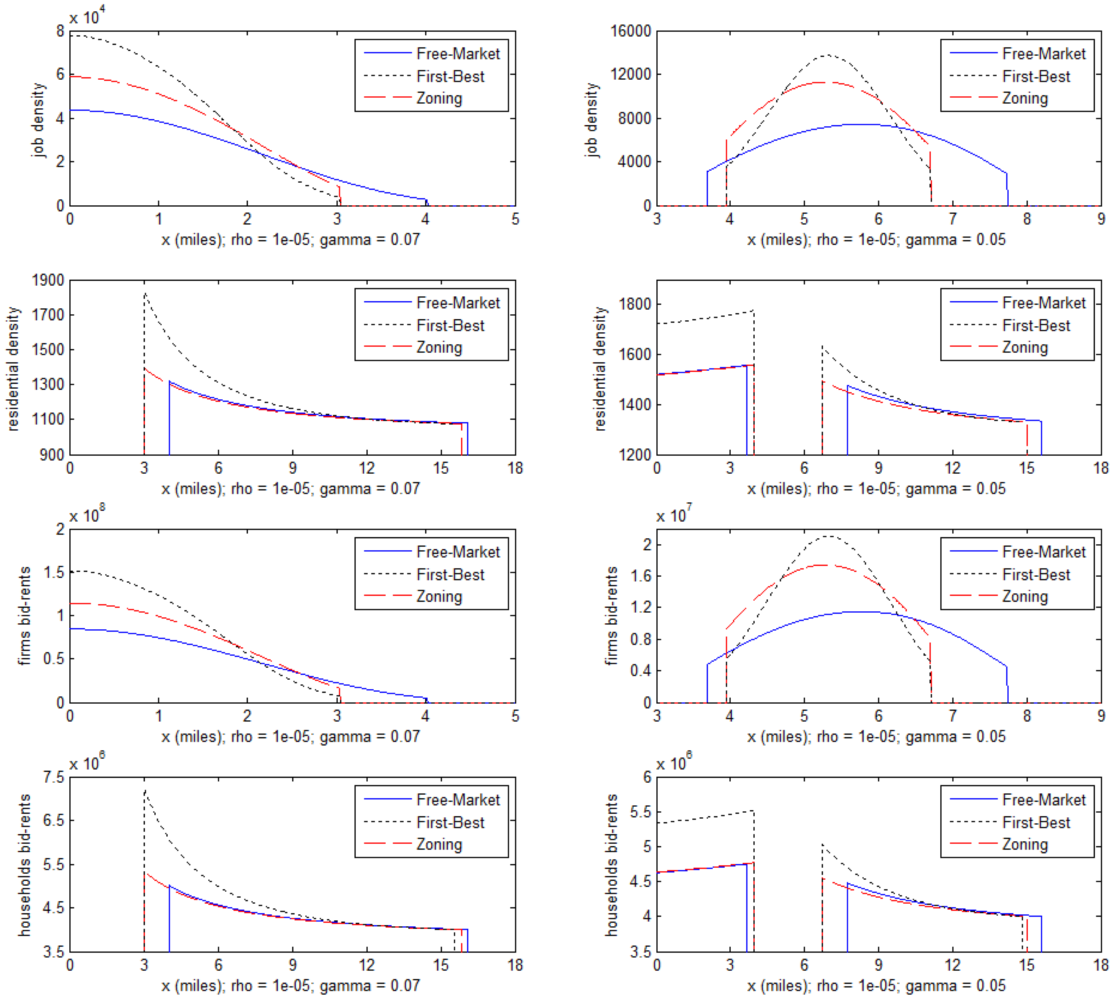


Figure 5.2 Spatial Distributions of Job and Residential Densities and Land Rents for Firm and Residential Use in the Firm-Cluster-Zoning, First-Best, and Free-Market Equilibria Varying with the FH(i.e., Monocentric, Left) and HFH (i.e., Nonmonocentric, Right) Urban Forms. (φ=20, ρ=0.00001, σ=1.5, γ=0.05 or 0.07, N=600,000)

LAND USE PLANNING FOR CORRECTING PLANNING FAILURES

Similar to first-best pricing instruments, first-best land use policies that remove all land use regulations are probably unrealistic and politically infeasible in planning practice. Many zoning codes are regulations and laws enacted by the local government. There are significant costs for property owners or city authorities to change zoning codes.

This section investigates two relatively realistic land use planning strategies that could mitigate planning failure and excessive congestion. The first policy is an imposition of residential densification policies against low-density zoning regulations and the second is against exclusionary zoning regulations by building employment centers in the suburbs.

Densification Policies

The densification policies discussed indicate that city authorities promote and allow for denser development in a particular residential area through regulation reforms and redevelopment, such as zoning changes from large-lot to small-lot zoning or from single-family to multi-family use and relaxing building height restrictions. The modeling analysis assumes that low-density zoning regulation is fully relaxed in a particular annular interval $[x_{d1}, x_{d2}]$. This interval is defined as a planning area in which land use planning policies are implemented. Thus, after imposing a densification policy in the planning area $[x_{d1}, x_{d2}]$, the equilibrium residential lot size $q_d(x)$ and density $\frac{1}{q_d(x)}$ is defined as follows:

$$\frac{1}{q_d(x)} = \begin{cases} \frac{1}{q^*(x)}, if \ \frac{1}{q^*(x)} < M \ or \ x \in \ [x_{d1}, x_{d2}] \\ M, \qquad\qquad\qquad\qquad\qquad others \end{cases} \tag{5.3}$$

where $q^*(x)$ is the equilibrium residential lot size without low-density zoning as defined in Eq.(3.2) and M is the density cap.

Excessive Congestion and Welfare Impacts.

Policies allowing denser development in particular residential areas in the suburbs reduce excessive congestion and improve social welfare. These effects vary with the planning area $[x_{d1}, x_{d2}]$ for imposing densification policies and the locations of such planning areas. Simulation results demonstrate that densification policies are more effective when they are imposed in a larger planning area or at locations closer to the urban core. For example, when a city has a zoning density cap of 1,000 households per square mile, the densification policy in the planning area [8, 9] improves the whole city's welfare by about $28 per household per year, accounting for 16% of the largest welfare improvement by removing all regulations (Table 5.3). As the planning area enlarges to [7, 9] or [8, 9], the average CV values increase to $47 and $43 per household per year, accounting for about 28% and 25% of total welfare improvement by removing all regulations. Such densification policies can make the average commute distance decrease by 1.6% to 3.6%, but at the same time increase the average traffic volume by 2.2% to 2.5%. These findings suggest that the densification policies defined here are not as effective as the "optimal" policies that remove all land use regulations, but they can still partially correct planning failures and reduce excessive congestion.

When the low-density zoning regulation becomes more restricted, as the density cap drops from 1,000 to 800 households per square mile, the same densification policies bring a larger amount of welfare improvement but become less efficient. For example, in the planning area [7, 9], densification policies produce a welfare gain of $119 for each household per year in the city under the relatively restrictive low-density zoning regulations (i.e., a cap of 800 hhs/sq mi.), about $72 higher than under less restrictive regulations (i.e., a cap of 1,000 hhs/sq mi.). Densification policies in more restrictive regulations are more effective for reducing commute distance and traffic volume on the roads. However, these policies reduce 11% of excessive congestion sourced from planning failures under the low-density zoning regulation with a cap of 800 hhs/sq mi., while they reduce 16% of excessive congestion under the low-density zoning with a cap of 1,000 hhs/sq mi. This finding indicates that a city with more restrictive low-density zoning regulations probably needs more planning areas allowing for denser development to reduce most excessive congestion and improve social efficiency.

Table 5.3 Impact of Densification Policies on Congestion and Welfare

	Densification policies allowing denser development at $[x_{d1}, x_{d2}]$ in cities with a density cap of 1000 hh./sq. mi.			Densification policies allowing denser development at $[x_{d1}, x_{d2}]$ in cities with a density cap of 800 hh./sq. mi.		
	[8,9]	[7,9]	[8,10]	[8,9]	[7,9]	[8,10]
CV comparative to FM cases under low-density zoning	27.81	46.98	42.72	60.29	119.37	116.67
% CV in the largest CV earned by removing all low-density zoning	16.37%	27.65%	25.14%	11.00%	21.78%	21.29%
% change of avg. commute distance before and after densification policies	-1.57%	-3.64%	-3.34%	-4.38%	-8.47%	-8.26%
% change of avg. traffic volume before and after densification policies	2.39%	2.19%	2.52%	-0.22%	-0.60%	0.01%

Land Use Impacts

After implementing densification policies in the planning areas, the planning areas will have higher residential densities and rents. Figures 5.3 and 5.4 show the simulated changes of densities and rents before and after implementing densification policies. The base cases are cities with low-density zoning regulation of two levels of density caps. Table 5.4 shows the percentage change in various land use characteristics after implementing densification policies. All of these results are estimated from simulations.

Inside the planning areas, densification policies can cause more than a 30% increase in residential densities when the density cap is set at 1,000 hhs/sq mi., and an increase of about 90% when the cap is 800 hhs/sq mi. (Figure 5.3). These result in an increase in average residential density but have no impact on job density and firm distribution (Table 5.4). Because densification policies can raise densities, the city becomes more compact after such policies are implemented. According to Table 5.4, densification policies in cities with more restrictive low-density zoning regulations generate higher average residential densities and more compact urban forms.

Table 5.4 Land Use Impact of Densification Policies

% change of variable values before and after densification policies	Densification policies allowing denser development at $[x_{d1}, x_{d2}]$ in cities with a density cap of 1000 hh./sq. mi.			Densification policies allowing denser development at $[x_{d1}, x_{d2}]$ in cities with a density cap of 800 hh./sq. mi.		
	[8,9]	[7,9]	[8,10]	[8,9]	[7,9]	[8,10]
avg. city Area	-2.52%	-4.33%	-4.45%	-4.03%	-7.67%	-8.17%
avg. job density	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
avg. residential density	2.29%	4.43%	4.57%	4.54%	9.03%	9.67%
avg. rent for firms	0.00%	0.00%	0.00%	0.00%	0.01%	0.01%
avg. rent for housing	0.26%	0.47%	0.46%	0.78%	1.54%	1.63%
avg. labor wage	0.04%	0.04%	0.04%	0.00%	0.01%	0.01%

Densification policies can raise land rents in the planning areas while other areas are restricted by low-density zoning regulations (Figure 5.4). In an aspatial perspective, the relaxation of density limits increases land and housing supply and lowers land value and housing prices given constant demand. However, when households' spatial decisions are endogenized as in this research, households in the outer suburbs outside the planning area would be likely to move to the planning area for closer commuting to their workplace in the urban core. These moving households also provide larger bid rents for the new housing built inside the planning area from the savings on their travel costs. In this case, when the land market reaches equilibrium, the land price in the planning area increases rather than decreasing. The rent-escalation effects of densification policies are greater in cities with more stringent low-density zoning regulations (Table 5.4)

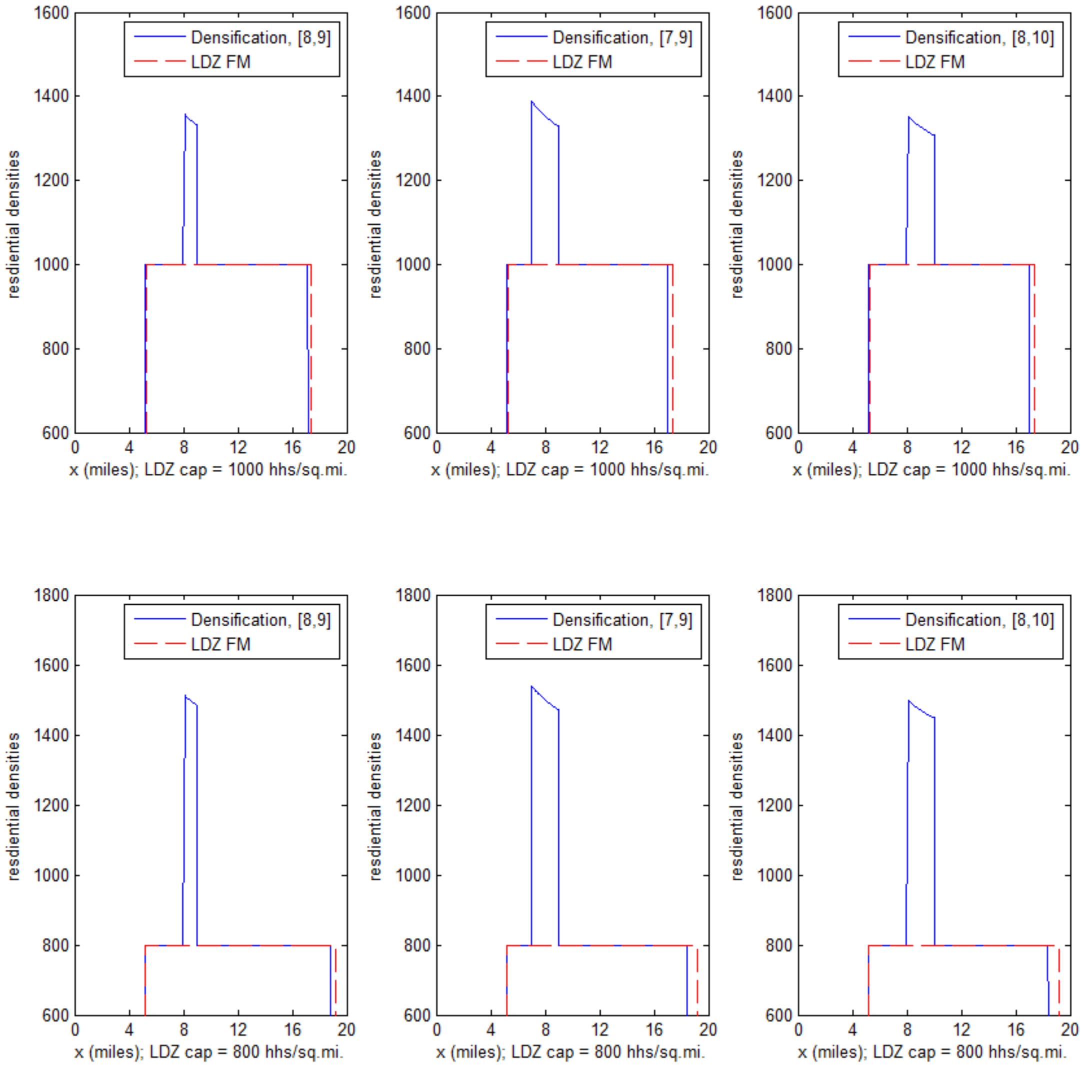


Figure 5.3 Land Use Effects After Relaxing Low-Density Regulations in Particular Areas

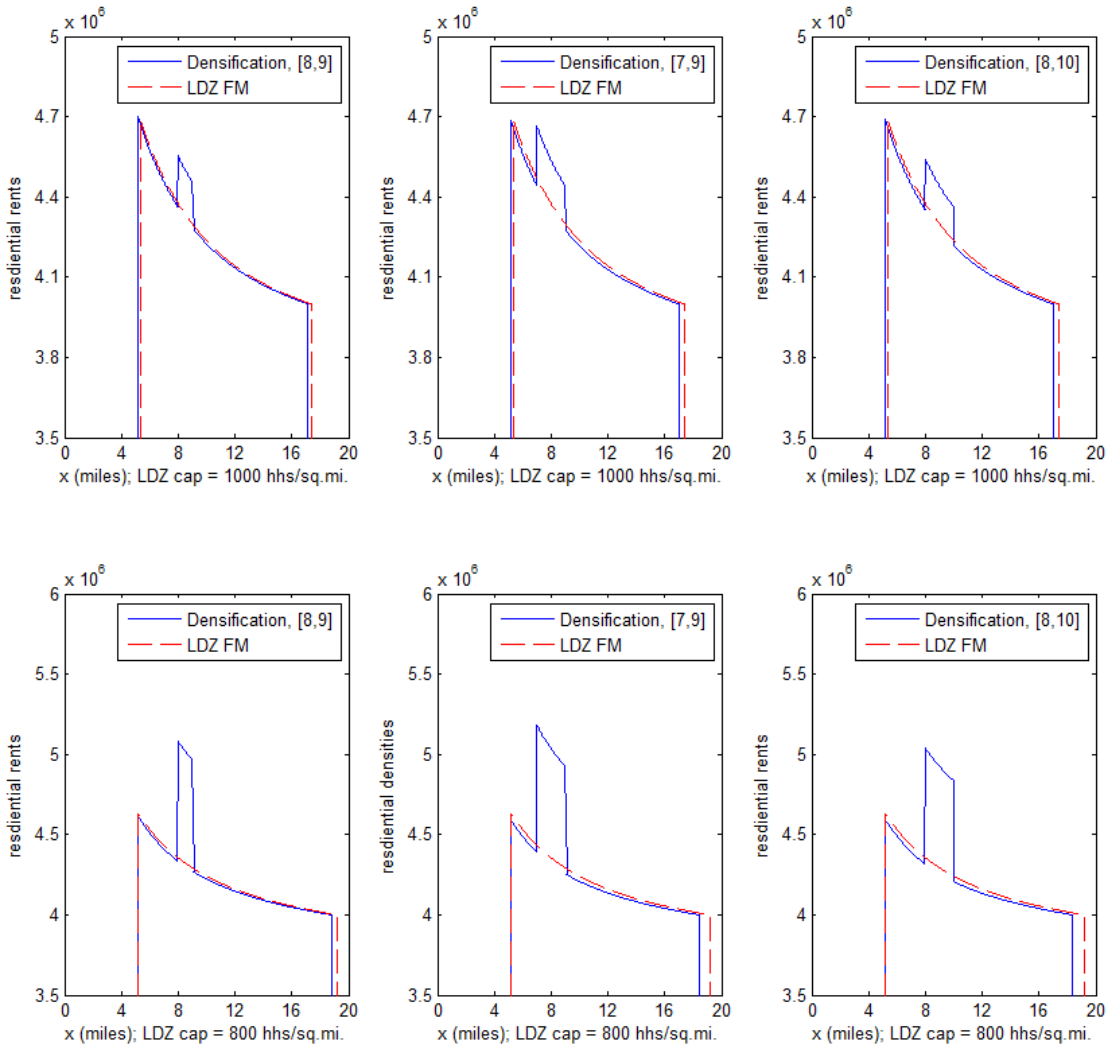


Figure 5.4 Land-Rent Effects After Relaxing Low-Density Regulations in Particular Areas

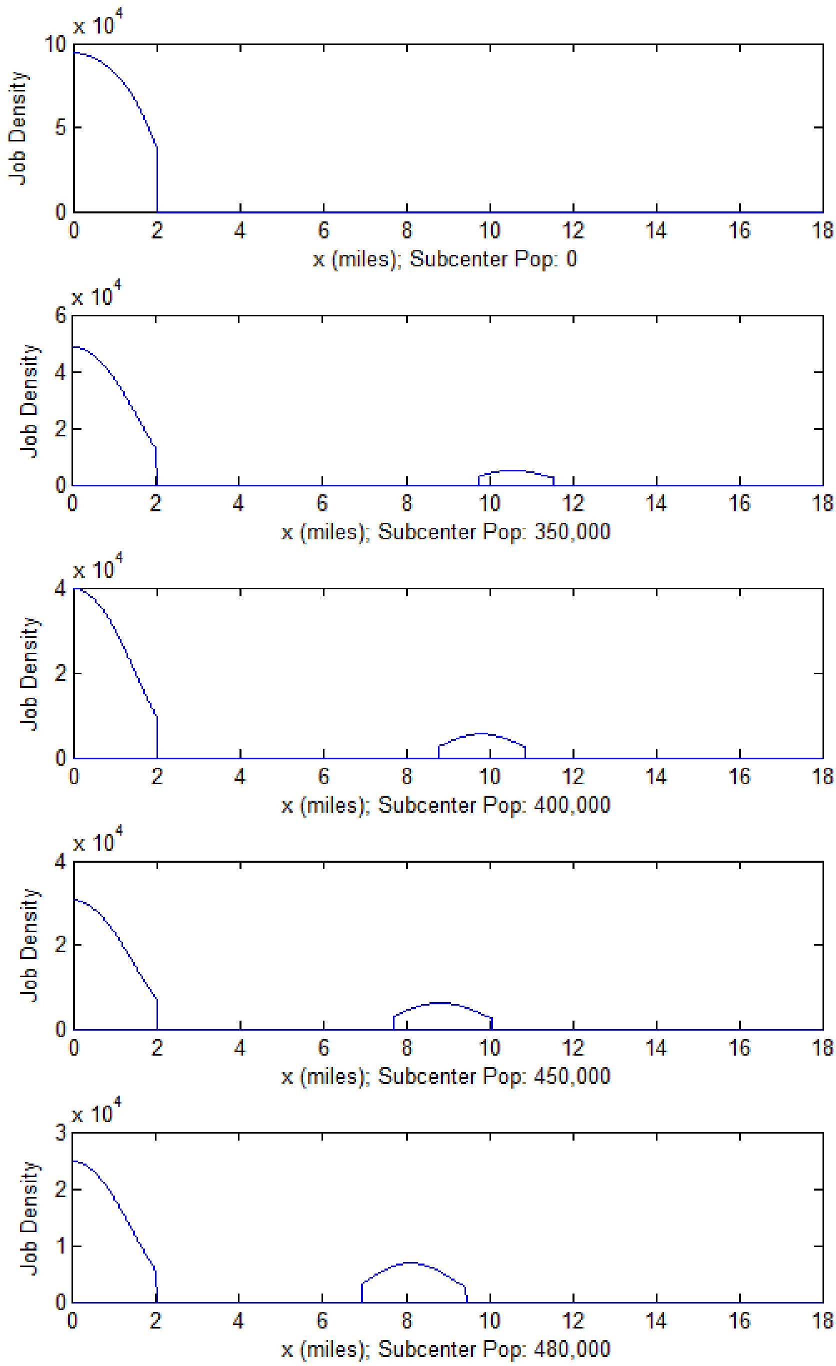
Job-Decentralization Policies: Building Employment Centers in the Suburbs

Chapters 3 and 4 suggest that job decentralization is often considered a market or socially desirable outcome for countering traffic congestion, as also found in many theoretical and empirical studies (e.g., Giuliano & Small, 1990; Gordon & Richardson, 1996; Crane & Chatman, 2004; Anas & Rhee, 2006). According to a report by Kneebone (2009, p. 1), in the US, "only 21% of employees in the top 98 metro areas work within 3 miles of downtown, while over twice that share (45%) work more than 10 miles away from the city center" and "employment steadily decentralized between 1998 and 2006: 95 out of 98 metro areas saw a decrease in the share of jobs located within 3 miles of downtown."

In this decentralization process, land use planning also plays an important role. While many land use regulations such as exclusionary zoning and low-density zoning regulations deter job densification, many land use planning strategies have instead been applied in practice to facilitate job decentralization. For example, building employment centers or subcenters in the suburbs has a long history in planning, from Ebenezer Howard's (1902) "garden city" to concepts of "satellite town" (Taylor, 1915) and "edge cities" (Garreau, 1991). This section discusses the potential benefits of such job-decentralization policies in cities under restricted exclusionary zoning regulations.[21]

The first question is how to facilitate a new subcenter. In our model the new subcenter is a new employment annulus in the suburbs. Simulations find that the strategy of simply designating a planning area exclusively for firm use is often meaningless, and it is difficult to attract firms to move from the urban core to the subcenter. The major reason is that firms moving to the subcenter will have very high risks of losing the benefits from agglomeration. A strategy discussed here is to subsidize companies that move to the subcenter. For example, as assumed in this model, the city authority will pay a subsidy for each laborer hired in the suburbs to each firm and the total subsidy will be financed by the land rent income from the urban area. Once some firms move to the subcenter the city land use will reach equilibrium when the urban residents' utility is equal to the suburban residents' utility.

[21] In cities without any land use regulations, building subcenters is a possible equilibrium but will bring more losses than the free market without subcenters.



Basic Status: No policy in a city with an exclusionary zoning limit at 2 miles

- Labor Subsidy: $535 per worker for firms moving to subcenter
- Moving Jobs: 350,000
- Earned CV: $161 per household for the whole city

- Labor Subsidy: $386 per worker for firms moving to subcenter
- Moving Jobs: 400,000
- Earned CV: $165 per household

- Moving Jobs: 450,000
- Marginal Cost: $230 per worker
- Earned CV: $212 per household

- Moving Jobs: 480,000
- Labor Subsidy: $148 per worker
- Earned CV: $233 per household

Figure 5.5 Evolution of Employment Subcenters

Simulation findings suggest that multiple equilibria can be found, and the welfare gains of these equilibria are determined by how much of the population moves to the subcenter (or stays at the urban center), which is determined by the magnitude of labor subsidies to firms in the subcenter. For example, based on our simulation scenarios, a subcenter equilibrium will not occur until the subcenter attracts more than around 350,000 jobs, above 58% of the total worker population. When the jobs in the subcenter exceed around 490,000, about 82% of the total worker population, the urban center will no longer exist (i.e., all firms will have moved to the subcenter). Therefore, this job-decentralization policy via building subcenters incorporates both land use planning and pricing policies.

Figure 5.5 presents the evolution of employment subcenters from the dynamic simulations. The base case is that all firms locate in the urban core area [0, 2] under the exclusionary zoning regulation. When local municipality provides a labor subsidy ($535 per year) to firms for each worker they hire in the subcenter, there are 350,000 jobs moved to the subcenter in the equilibrium situation. In this case, each household in the whole city, including those working at the urban and suburban centers, can earn about $161 per year. After that the local municipality provides a relatively low subsidy at $386 per year to attract 50,000 more jobs to the subcenter, and the location of the subcenter area will move toward the urban core. When the subcenter has grown larger, firms moving to the subcenter will lose lower agglomeration benefits and the marginal cost (i.e., the labor subsidy) for moving them away from the center will fall. Thus the labor subsidy drops to $23 and $148 when the jobs in the subcenter increase to 450,000 and 480,000. The corresponding welfare gains to each household in the whole city also increase to $212 and $232 yearly. These findings suggest that a policy incorporating land use planning and economic strategy has the potential to correct planning failure by facilitating job decentralization and re-agglomeration and could reduce the economic scale of the urban center but significantly improve social welfare of all households in both urban and suburban areas.

Congestion Impacts

Table 5.5 summarizes the impact of job-decentralization policies on travel and congestion outcomes. Job-decentralization policies are more effective for reducing excessive congestion when they attract more jobs to the subcenter. When the number of jobs in the subcenter account for 58% of total jobs in the city, the excessive congestion from planning failures can be reduced by about 33% of excessive congestion from planning failure. As the share of suburban jobs increases to 80%, the excessive congestion from planning failure can be alleviated by 47%.

These job-decentralization policies significantly reduce travel and congestion. For example, after 480,000 jobs are relocated to the suburbs the average commute distance for residents working in the urban and suburban centers will drop by over 55% and 42% respectively. Meanwhile the average traffic volume on roads in the urban and suburban areas will drop by 69% and 72% respectively. The more jobs move to the subcenter, the better the traffic conditions (shorter commute distance and less traffic) in the urban area will be. Although the traffic conditions become worse as the suburban jobs increase, they are still much better than in the monocentric case. These traffic conditions are even better than those in the social optimum after all market and planning failures are corrected. However, it is worth mentioning that while city residents can benefit hugely from the reduction in VMT and congestion, the city's agglomeration economy is weakened.

Table 5.5 Impact on Congestion from job-decentralization policies

Change of variable values before and after building subcenter	Subcenter with 350,000 jobs	Subcenter with 400,000 jobs	Subcenter with 450,000 jobs	Subcenter with 480,000jobs
Reduced diseconomy of excessive congestion comparative to FM cases under exclusionary zoning	161.70	165.40	211.80	233.48
% excessive congestion from planning failure	32.64%	33.38%	42.75%	47.13%
% Avg. Commute Distance in the Urban area	-34.96%	-42.04%	-50.04%	-55.49%
% Avg. Commute Distance in the Suburban area	-60.73%	-54.17%	-46.89%	-42.12%
% Avg. Traffic Volume in the Urban area	-45.94%	-54.19%	-63.17%	-68.98%
% Avg. Traffic Volume in the Suburban area	-82.98%	-79.63%	-75.52%	-72.44%

Land Use Impacts

Job-decentralization policies can generate efficient "sprawling." The city area will increase by over 20%, varying with the level of decentralization. This can greatly reduce job densities and land rents for firms. For example, when the subcenter attracts 80% of the total jobs the average job density in the urban and suburban centers will drop by 58% and 94% respectively. Despite the subcenter having four times the jobs in the urban center, the job density in the urban center is higher than in the subcenter. The decreases in job densities also lower the bid rents of firms, leading to a trend of rent decline similar to density decline. Compared to the impacts on firms, the effect on residential densities and rents is relatively small. The job-decentralization policies can lead to a decline of around 4% in the average residential density and the average land rent for housing in both urban and suburban areas.

Table 5.6 Impact on Land Use after Building Suburban Employment Center

% change of variable values after building subcenters		Subcenter with 350,000 jobs	Subcenter with 400,000 jobs	Subcenter with 450,000 jobs	Subcenter with 480,000jobs
city area		23.36%	23.96%	23.66%	21.99%
avg. job density	Urban	-80.00%	-75.00%	-66.67%	-58.33%
	Suburban	-92.11%	-92.80%	-93.44%	-93.81%
avg. residential density	Urban	-4.72%	-4.61%	-4.06%	-3.54%
	Suburban	-4.02%	-4.12%	-4.56%	-4.56%
avg. rent for firms	Urban	-80.26%	-75.28%	-66.95%	-58.61%
	Suburban	-92.28%	-92.96%	-93.60%	-94.00%
avg. rent for housing	Urban	-4.89%	-4.60%	-4.12%	-3.64%
	Suburban	-3.89%	-4.16%	-4.57%	-4.84%
avg. labor wage	Urban	-1.32%	-1.11%	-0.85%	-0.65%
	Suburban	-2.05%	-2.24%	-2.46%	-3.09%

SUMMARY

This chapter simulated the anti-congestion, welfare, and land use effects of four types of second-best but more practical land use planning strategies. Among them, UGBs and firm cluster zoning regulations were investigated as two policies for reducing excessive congestion sourced from market failures. Densification policies that allow denser development and job-decentralization policies that build new suburban employment centers were examined as two policies for reducing excessive congestion sourced from planning failures.

The UGB regulations may partially correct distortions in both transport and labor markets, but may worsen land market distortion via the residential rent-escalation effects, leading to trivial utility gains. Such UGB distortions in land markets appear in regions like Portland, Oregon and Knoxville, Tennessee, where housing rents/prices inside the UGBs rise faster than those of properties in areas without UGBs (Staley & Mildner, 1999; Cho et al., 2008). London, England and Auckland, New Zealand have also reported major rent escalations due to relatively low housing or land supply for new development (Cheshire & Sheppard, 2005; Cox, 2010). Home affordability remains a key topic for debate in growth management discussions (Downs, 2004; Nelson et al., 2004). Of course, real cities are much more complex than the models allowed here. Human health, ecological conservation, social interaction, and other variables are at play and may counteract some or much of the rent escalation losses that tend to come with tight UGBs.

The firm cluster zoning policies that regulate a zone's land use exclusively for firm/business use are probably more efficient than the UGB policies for reducing congestion and enhancing agglomeration. They generate welfare improvement closer to the first-best level and will not create much excessive congestion or excessive escalation of housing rents, and they avoid the housing affordability issue raised by the UGB policies. While planning practice should pay more attention to such an effective policy, at least in theory, urban economic models should allow for land use policy scenarios related to firms.

The objectives of policies for correcting planning failures differ from those for market failures. Remedies for market failures aim at anti-congestion and welfare effects by comparison with the free-market (bottom limit) and socially optimal (upper limit) cases, assuming no planning failure exists. In contrast, remedies for planning failures seek to compare anti-congestion and welfare effects with the free-market case with a bottom limit and without upper-limit regulations. For example, densification policies that relax low-density zoning regulations in a particular planning area partially correct planning failure from low-density zoning and reduce excessive congestion. A city with more stringent low-density zoning regulations probably allows more areas for denser development.

On the other hand, when firms are regulated against decentralization (e.g., due to exclusionary zoning), policies aimed at moving jobs to the subcenter need to incorporate both land use and pricing strategies. Simulation findings in this research suggest that a subsidy to firms (e.g., a labor subsidy for hiring workers) is the key trigger to firm decentralization. The optimal setting of job decentralization can reduce about half of excessive congestion and improve half of welfare. Job decentralization also brings other attractive land use and transportation consequences including significant drops in firm rents, VMT, and congestion, despite the disadvantage of declined agglomeration economies.

CHAPTER 6: INCORPORATING LAND USE AND ECONOMIC POLICIES: AN EMPIRICAL STUDY IN AUSTIN

Previous chapters have presented theoretical (Chapter 2) and analytical (Chapters 3–5) findings to demonstrate the importance of incorporating land use and economic policies for correcting market and planning failures that incur excessive congestion and social inefficiency. These findings suggest that a direct empirical analysis needs to estimate the amount of excessive congestion in metropolitan areas and explain the variations of the amount by the variations in land use regulations, pricing policies, and their interactions. This type of direct empirical models, however, appears to be difficult to build and estimate. One major reason is the lack of detailed data sets (e.g., local congestion data, individual travel consumption, land use regulation data, and pricing scheme data). The other major reason is the difficulty of measuring excessive congestion; it is difficult to identify the optimal travel demand of each individual traveler. Even we can accurately calibrate the analytical model developed in previous chapters using data of a realistic city, the corresponding empirical findings may not fit the reality well, since many modeling settings are oversimplified (e.g., continuous circular space, radial commute, no heterogeneity among residents and firms).

This chapter thus creates an alternative approach to empirical analysis, with a focus on travel-related benefits of a combination of land use and pricing policies. These benefits are not directly compared to the optimal level of travel, as the measurement of excessive congestion does, but estimated by comparing different policies, including land use alone policy, pricing alone policy, and an combination of both policies. Specifically, this chapter provides an empirical study to investigate the interactional effects of land use and economic factors on auto travel behaviors. Specifically, relying on data from Austin's 2005–2006 household activity-travel survey, this research develops a multilevel multinomial logit (MML) model to estimate the impacts of travel costs and land use variables on travel mode choice.

As reviewed in Chapter 2, while most empirical studies have focused on either congestion pricing or land use planning strategies for driving reduction, only limited studies have examined a combination of land use and pricing policies as a strategy to reduce excessive driving and congestion (e.g., Guo et al., 2011; Lee & Lee, 2014). For example, Guo et al. (2011) tracked a pilot mileage fee program in Portland, Oregon with 130 household participants who needed to pay either a congestion pricing fee or a flat-rate charge for travel over 10 months. They found that there was a "mutually supportive" relationship between congestion pricing and land use planning strategies. Congestion pricing can reduce vehicle-miles traveled (VMT) for those households in traditional neighborhoods with dense and mixed-use built environments. Also, compact land use policies may be more efficient for VMT reduction if congestion pricing is imposed. Lee and Lee (2013) investigated how gasoline prices and land use characteristics affect transit ridership in 67 urbanized areas from 2002 and 2010. They advocated a complementary land use and pricing policy and suggested that the effects of urban land use on transit ridership become greater when driving externalities are corrected.

This chapter proposes a new study with a focus on how travel costs and land use variables mutually affect people's travel mode choice among five alternatives: driving alone, shared driving, public transit, bicycling, and walking. Despite this analysis not directly measuring traffic congestion, it evaluates (1) whether pricing schemes that increase the cost of driving significantly reduce auto travel, (2) whether denser and more mixed-use developments lead to less auto dependence, and (3) what benefits from driving reduction are induced by incorporating land use and pricing policies.

RESEARCH DATA

Data for this study primarily comes from the 2005–2006 household activity-travel survey conducted in Austin, the metropolitan capital area of Texas. The survey recorded household information, individual characteristics, vehicle information, and a 24-hour activity-travel diary of each adult member in households. The number of responses was 1,499 households and 4,117 individuals, yielding 18,545 trip records. The study focuses on mode choice for nonwork trips, which play an increasingly important role in people's everyday lives. National household travel surveys (NHTSs) from 1969 to 2001 revealed that the share of nonwork trips increased from 75% to over 85%, coupled with a 72% increase in total trips (USDOT, 2003). Most of the trips, including those of short distances (less than 2 miles), were made by driving. Many nonwork trips and activities are discretionary and therefore more likely to be influenced by policies than commute trips. In particular, the cross-sectional investigation in this study found it difficult to estimate the long-term impact of land use on commuting behaviors, as done in Chapters 3–5.

Figure 6.1 Research Area and Locations of Household Sampled in 2005-2006 Austin Household Travel Surveyed

For evaluating the impact of land use planning on travel behaviors, the selected research samples are those residents living in the neighborhood planning areas (NPAs) and activity centers in Austin (Figure 6.1). This research selects neighborhoods in the NPAs rather than TAZs or census blocks as the basic spatial unit for calculating land use characteristics. First, the definition of neighborhood boundaries is always based on natural objects, like rivers, parks, and transport networks; on socio-demographic and census information; and/or is decided by public meetings and surveys. Residents living in a neighborhood are a group with close social and physical association. Second, the neighborhood or community is probably the basic spatial unity for marking city plans and many land use policies or regulations are imposed based on neighborhoods rather than TAZs or census blocks.

In addition, this study identifies 27 activity centers or mixed-use development (MXD) zones in the NPAs, because activity centers have more heterogeneous land use and social components than regular neighborhoods and they are often located at the intersection of several neighborhoods. In practice the selection of the research sample of MXDs took a "bottom up" approach, based on local knowledge of city officials, professional planners, staff from the Capital Area Metropolitan Planning Organization (CAMPO), and academic experts (Zhang, Kone, Tooley, & Ramphul, 2009). The sampling process involved three working steps. First, a list of 49 communities in the region was created and the contact information of representative planners or public officials collected. Planners or officials were then interviewed by phone to identify MXD's boundaries based on their professional and personal knowledge of their communities. Each interviewee was first given a definition of MXD: "A mixed-use development or district consists of two or more land uses between which trips can be made using local streets without having to use major streets. The uses may include residential, retail, office, and/or entertainment. There may be walk trips between uses." If the planner required further clarification, an additional set of characteristics of mixed-use districts, as defined by the ULI (Witherspoon, Abbett, & Gladstone, 1976), was provided along with known examples, such as the Triangle area in Austin.

Accordingly, there is a total of 27 MXD neighborhoods in Austin and 65 non-MXD neighborhoods in Austin (Figure 6.1). After canceling some missing data and zones without households being surveyed, this study includes a sample of 2,141 trips recorded by 975 individuals in 427 households located in 79 neighborhoods.

Land use data comes from ArcGIS-encoded zone data for the research areas, as obtained from the City of Austin and CAMPO. The survey also obtained the geographic coordinates of activity locations and trip ends (origins and destinations) of the surveyed travelers. For travel analysis, these trip ends were geocoded in ArcGIS. Network distance was estimated based on the assumption that the traveler took the shortest path in length between trip origin and destination.

DESCRIPTION ANALYSIS

Table 6.1 Mode Choice Shares of Nonwork Travel Mode and Related Mean Sample Values of Level-Of-Service Characters

Mode Choice	Frequency	Percent (%)	Cost (dollar/ mile)	Time (minutes/ mile)
Driving Alone (DA)	730	34.10	0.95	6.45
Shared Ride (SR)	1185	55.35	0.41	8.20
Public Transit(PT)	38	1.77	0.59	12.85
Walking (WA)	134	6.26	0	28.44
Bicycle (BI)	54	2.52	0	16.73
Total	2141	100.00		

Note: The cost and time values are those from CAMPO's skim file results.

This research categorizes travel modes into five types: driving alone (DA), shared ride (SR), public transit (PT), walking (WA), and bicycling (BI). Among them, shared ride is the dominant mode of the nonwork trips of Austin's residents, occupying as much as 55% of the total sample. This percentage is even greater than the mode of driving alone, occupying more than one-third of the total (Table 6.1). Nearly 90% of nonwork trips in Austin are conducted by automobile. Only 38 trips were recorded as using public transit, including bus and taxi, accounting for only 1.77%. More than 8% of the travelers employed nonmotorized modes, such as walking and bicycling. Table 6.1 provides averages of two level-of-service variables, travel cost and time. On average, the mode of driving alone for a nonwork trip cost the most money and the least time as compared with other modes. Travelers walking and bicycling paid no out-of-pocket cost but consumed the most time, 28 and 17 minutes per mile on average, respectively.

Table 6.2 Descriptive Statistics of the Multilevel Structure of Variables

Variable	Explanation	Mean/Share	Std. Dev.
Level-1: Individual/Household			
AGE20	Age of the Traveler (1: if age<=20; 0: otherwise)	0.276	
AGE60	Age of the Traveler (1: if age>60; 0: otherwise)	0.201	
FEMALE	Gender (1: if female; 0: otherwise)	0.533	
WHITE	Race (1: if white, 0: otherwise)	0.501	
EMP	Employment Status (1: if employed; 0: otherwise)	0.460	
HHSIZ	Household Size (persons)	2.599	1.500
VEPHM	Vehicles per Household Members	1.677	0.813
HHINC	Household Income (transferring form degree variables, thousand dollars)	46.083	40.080
Level2 : MXD/ Neighborhood			
PODEN	Population Density (persons/acre)	7.922	4.375
EPDEN	Employment Density (persons/acre)	4.475	5.869
SWDEN	Sidewalk Density (miles/acre)	0.045	0.018
LUMIX	Entropy Index of Land-Use Mix (0-1)	0.602	0.156
DTNAC	Distant to the Nearest Activity Center (miles)	0.922	0.992

The explanatory variables contain two-level factors, the level of individual/household and the level of neighborhood/MXD, as delineated in Table 6.2. The individual-level attributes include age, gender, race, employment status, household size, and vehicle ownership in the household, while the neighborhood-level variables consist of population density, employment density, sidewalk density, land use mixture entropy, and distance to the nearest activity center. The entropy index of land use mixture is calculated as $-\sum_j [P_j * ln(P_j)] / ln(J)$, where P_j is the proportion of developed land in the *j*th use type and *J* is the number of land use categories considered. In this study *J* = 6: residential, commercial, office, industrial, civic, and open space.

MULTILEVEL MULTINOMIAL LOGIT MODEL

The traditional analysis of travel mode choice relies on discrete choice models with an assumption of random utility maximization (RUM) (McFadden, 1974; Ben-Akiva & Lerman, 1985). Given that an individual *i (i = 1,2, …, I)* living in the neighborhood *j (j = 1,2,…, J)* chooses an alternative of travel mode *m (m = 1,2,…, M)*, he or she will have the following utility function:

$$U_{ijm} = \alpha_{jm} + \boldsymbol{\beta}_{\boldsymbol{j}}^{'} \boldsymbol{z}_{\boldsymbol{ijm}} + \boldsymbol{\gamma}_{\boldsymbol{m}}^{'} \boldsymbol{x}_{\boldsymbol{ij}} + \epsilon_{ijm} \tag{6.1}$$

where α_{jm} is a scalar utility term for alternative *m* associated with the neighborhood *j* of the individual. $\boldsymbol{z}_{\boldsymbol{ijm}}$ is an individual-specific covariate vector that varies with alternatives and may also vary over individuals and neighborhoods. The vector $\boldsymbol{x}_{\boldsymbol{ij}}$ varies with characteristics of individuals and neighborhoods.

The model developed here underscores the contextual effect of land use on travel mode choice. This contextual effect indicates that the impact of level-of-service factors on travel mode choice may vary with neighborhoods with diverse land use features. Specifically, the coefficients ($\boldsymbol{\beta}_{\boldsymbol{j}}$) of the alternative-associated variables, such as travel cost and time, are hypothesized to vary across neighborhoods. $\boldsymbol{\gamma}_{\boldsymbol{m}}$ is the coefficient vector of individual-associated variables such as age, income, and household size. They are assumed to vary across the alternatives only. ϵ_{ijm} is an unobserved standard extreme value random term, which represents all other factors affecting the utility of mode choice but not included in the regressors. One can assume ϵ_{ijm} to be independently and identically (IID) distributed.

Eq.(6.1) represents the individual-level variation of mode choice. The next step is to allow the intercept term α_{jm} and the coefficient vectors β_j for interacting with land use variables in the neighborhood level. Thus the Level-2 model is given as follows

$$\alpha_{jm} = \delta_m + \boldsymbol{\pi}_{\boldsymbol{m}}^{'} \boldsymbol{w}_{\boldsymbol{j}} + \theta_{jm} \tag{6.2}$$

$$\boldsymbol{\beta}_{\boldsymbol{j}} = \boldsymbol{\rho} + \boldsymbol{W}_{\boldsymbol{j}} \boldsymbol{\mu} + \boldsymbol{\varphi}_{\boldsymbol{j}} \tag{6.3}$$

In Eq.(6.2) δ_m is an alternative-specific constant of the average effect of unobserved variables on the utilities associated with the mode *m*. $\boldsymbol{w}_{\boldsymbol{j}} = (w_{j1}, w_{j2}, \ldots w_{jK})^{'}$ and w_{jk} are land use variables, such as density, walkable environment, accessibility, or land use mixture. $\boldsymbol{\pi}_{\boldsymbol{m}}$ is the corresponding coefficient vector related to mode *m*. $\theta_{jm} \sim N(0, \sigma_m^2)$ are random terms that represent unobserved idiosyncratic difference across neighborhoods. They are assumed to be normally and identically distributed across neighborhoods. The coefficients in Eq.(6.3) do not vary with the alternatives of travel mode. $\boldsymbol{\rho}$ is an intercept vector indicating the average effect of unobserved variables on the slope of level-of-service factors. $\boldsymbol{W}_{\boldsymbol{j}}$ is a diagonal matrix of land use variables:

$$\boldsymbol{W}_{\boldsymbol{j}} = \begin{pmatrix} w_{j1} & 0 & \ldots & 0 \\ 0 & w_{j2} & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & w_{jK} \end{pmatrix} \tag{6.4}$$

$\boldsymbol{\mu}$ is the corresponding coefficient vector of land use variables. $\boldsymbol{\varphi}_{\boldsymbol{j}}$ is a vector of random terms that capture unobserved variations across neighborhoods. All the elements of $\boldsymbol{\varphi}_{\boldsymbol{j}} = (\varphi_1, \varphi_2, \ldots, \varphi_K)^{'}$ are assumed to be normally and identically distributed across neighborhoods, i.e., $\varphi_k (k = 1, 2, \ldots, K) \sim N(0, \omega_k^2)$.

Combining the equations from (6.1) to (6.3), the integrated equation is as follows:

$$U_{ijm} = (\delta_m + \boldsymbol{\pi}_{\boldsymbol{m}}^{'} \boldsymbol{w}_{\boldsymbol{j}} + \boldsymbol{\rho}^{'} \boldsymbol{z}_{\boldsymbol{ijm}} + \boldsymbol{\gamma}_{\boldsymbol{m}}^{'} \boldsymbol{x}_{\boldsymbol{ij}} + \boldsymbol{\mu}^{'} \boldsymbol{W}_{\boldsymbol{j}} \boldsymbol{z}_{\boldsymbol{ijm}}) + (\theta_{jm} + \boldsymbol{\varphi}_{\boldsymbol{j}}^{'} \boldsymbol{z}_{\boldsymbol{ijm}} + \epsilon_{ijm}) \tag{6.5}$$

This equation includes two parts: fixed effects and random effects. The segment $[\ \delta_m + \boldsymbol{\pi}_{\boldsymbol{m}}^{'} \boldsymbol{w}_{\boldsymbol{j}} + \ \boldsymbol{\rho}^{'}\ \boldsymbol{z}_{\boldsymbol{ijm}} + \boldsymbol{\gamma}_{\boldsymbol{m}}^{'} \boldsymbol{x}_{\boldsymbol{ij}} + \boldsymbol{\mu}^{'}\ \boldsymbol{W}_{\boldsymbol{j}} \boldsymbol{z}_{\boldsymbol{ijm}}\ ]$ in Eq. (6.5) contains the fixed coefficients. The segment $[\theta_{jm} + \boldsymbol{\varphi}_{\boldsymbol{j}}^{'}\ \boldsymbol{z}_{\boldsymbol{ijm}} + \epsilon_{ijm}]$ contains the random error terms. The terms $\boldsymbol{W}_{\boldsymbol{j}} \boldsymbol{z}_{\boldsymbol{ijm}}$ are a list of interaction terms multiplying LOS variables $\boldsymbol{z}_{\boldsymbol{ijm}}$ by neighborhood-level land use variables $\boldsymbol{W}_{\boldsymbol{j}}$. From these interaction terms we can estimate how the effect of the travel cost and time on mode choice is adjusted or moderated by the land use context.

Because $\boldsymbol{z}_{\boldsymbol{ijm}}$ connects with the random error vector $\boldsymbol{\varphi}_{\boldsymbol{j}}$, the derived total error could vary with individual socio-economic attributes. The property of independence from irrelevant alternatives (IIA) in MNL models may fail in the multilevel structure. Individuals in the same neighborhood are probably interdependent due to unobserved heterogeneity within and between neighborhoods (Bhat, Carini, & Misra, 1999; Raudenbush & Bryk, 2002).

Letting the error terms θ_{jm} $(m = 2, \dots, M)$ and $\varphi_k (k = \ 2, \dots, K)$ be conditioned, the probability of choice of mode *m* for individual *i* nested within living neighborhood *j* can be written in the traditional MNL form:

$$P_{ijm} | \left(\theta_{j2}, \dots, \theta_{jM},\ \varphi_2, \dots\ , \varphi_K\right) = \frac{exp(\delta_m + \boldsymbol{\pi}_{\boldsymbol{m}}^{'} \boldsymbol{w}_{\boldsymbol{j}} + \boldsymbol{\rho}^{'}\ \boldsymbol{z}_{\boldsymbol{ijm}} + \boldsymbol{\gamma}_{\boldsymbol{m}}^{'} \boldsymbol{x}_{\boldsymbol{ij}} + \boldsymbol{\mu}^{'}\ \boldsymbol{W}_{\boldsymbol{j}} \boldsymbol{z}_{\boldsymbol{ijm}} + \theta_{jm} + \boldsymbol{\varphi}_{\boldsymbol{j}}^{'}\ \boldsymbol{z}_{\boldsymbol{ijm}})}{\sum_n^M exp(\delta_n + \boldsymbol{\pi}_{\boldsymbol{n}}^{'}\ \boldsymbol{w}_{\boldsymbol{j}} + \boldsymbol{\rho}^{'}\ \boldsymbol{z}_{\boldsymbol{ijn}} + \boldsymbol{\gamma}_{\boldsymbol{n}}^{'}\ \boldsymbol{x}_{\boldsymbol{ij}} + \boldsymbol{\mu}^{'}\ \boldsymbol{W}_{\boldsymbol{j}} \boldsymbol{z}_{\boldsymbol{ijn}} + \theta_{jn} + \boldsymbol{\varphi}_{\boldsymbol{j}}^{'}\ \boldsymbol{z}_{\boldsymbol{ijn}})}$$

(6.6)

The unconditional likelihood function for the multilevel choice model does not have closed-form solutions. In other words, the maximization of probability requires some integral approximation (Raudenbush & Bryk, 2002; Rabe-Hesketh, Skrondal, & Pickles, 2004; Grilli & Rampichini, 2007). This research adopts adaptive Gauss-Hermite quadrature to integrate the latent variables and obtain the marginal log-likelihood under the program gllamm (Rabe-Hesketh et al., 2004) running in the software STATA 12.

According to Raudenbush and Bryk (2002) and Rabe-Hesketh, Skrondal, & Pickles (2002), the adaptive Gauss-Hermite quadrature technique in maximum likelihood (ML) estimation can improve the accuracy of approximate methods compared to other techniques such as marginal quasi-likelihood (MQL), penalized quasi-likelihood (PQL), Markov Chain Monte Carlo (MCMC), and Gaussian quadrature (GQ). However, the adaptive quadrature estimation in the gllamm program could be very slow, particularly when the estimation includes many random effects (Rabe-Hesketh et al., 2002). Thus such an estimation often needs to first exclude random effects in those insignificant coefficients of the Level-1 variables (e.g., α_{jm} and $\boldsymbol{\rho}$ in Eq.(6.1)) and to examine which random effects can be combined. These statistical adjustments also need to follow theoretical assumptions, and the final model should be a balanced structure that achieves accuracy and efficiency.

VARIATIONS IN MODE CHOICE FROM INDIVIDUAL VERSUS NEIGHBORHOOD VARIABLES

The modeling analysis in this research includes two steps. The first step involves an estimation of a base model in which the Level-1 model includes only the alternative specific constant and level-of-service variables and the Level-2 model introduces only the constant and the error term. The base model, Model 1, is applied to the survey if there are neighborhood-level variations in the Level-1 coefficients. We use multiple modeling runs to examine coefficients of the LOS variables and the Level-1 intercepts. Results from the first-step analysis can show which coefficients of Level-1 variables significantly vary with neighborhoods. If no Level-1 variables have contextual effects, the multilevel model becomes unnecessary and collapses into the conventional single-level model. Thus Model 1 is mainly used for estimating variations and to help determine a statistically efficient modeling structure.

Specifically, if the variances of the random component $\sigma_m^2 (m = 1,2, \ldots, M)$ are zero, one can conclude that there are no between-neighborhood variations in the average log-odds of mode choice. The average log-odds can thus be predicted only by individual-level variables. We first set σ_{DA} and σ_{SR} as random effects. Based on estimations, the random intercept variance σ_{DA}^2 is 3.37 with a standard error of 0.96, and σ_{SR}^2 is 2.89 with a standard error of 0.80. These results demonstrate that the between-neighborhood variations in the log-odds ratios of selecting DA or SR are statistically significant compared to other modes. Also, the covariance between σ_{DA} and σ_{SR} is estimated as 2.88 with a standard error of 0.86. This suggests a high correlation between the log-odd of choosing DA and SR, indicating that these two random effects can be combined.

Also, the random effects of $\alpha_{j,PT}$, $\alpha_{j,WA}$, and $\alpha_{j,BI}$ are tested in Model 1. There are high correlations between $\alpha_{j,DA}$ and $\alpha_{j,SR}$ and between $\alpha_{j,WA}$ and $\alpha_{j,BI}$. These findings suggest a combination of these pairs of random effects. In addition, the random intercept variance of public transit σ_{PT}^2 is 7.84 with a standard error of 4.40. The insignificance implies that the random effect σ_{PT}^2 can be set as 0. Therefore these variation estimates suggest that the model only internalizes the neighborhood-level variation between the mode by auto (σ_{Auto}^2) and by nonautomobile.

Table 6.3 shows the results of the final intercept-only model. The utility variance of the auto mode, including DA and SR, across neighborhoods is 2.724 with a standard error of 0.577. Significantly, the null hypothesis H_0: $\sigma_{Auto}^2 = 0$ can be rejected by the *t*-test value (t = 4.75). According to Hox (2002) the Level-1 variation of the MNL structure is approximated as $\frac{\pi^2}{3}$ and the intralevel correlation for the null model is 0.453, i.e., 2.724/(2.724+$\frac{\pi^2}{3}$). These findings show that over 45% of the variations of mode choice between auto and nonauto are determined by neighborhood-level contexts. Nearly 55% of the variations occur at the individual level.

Similar tests are conducted to examine the LOS variable. Model 2 adds LOS variables into the final Model 1(Table 6.3). The slopes of these level-of-service variables in the Level-1 equation represent the effect of a unit change in the value of service variables on individual-level log-odds of mode choice, all else being equal. After the coefficients of the level-of-service variables are examined only ρ_1, the slope of the travel cost variable, varies significantly across neighborhoods. Therefore the following models will consider only the interaction effects between land use variables and the cost variable.

Table 6.3 Intercept-Only Model with Random-Slope Effects of Level-of-Service Variables

Fixed Effects	Note	Model 1: Base model Intercept-only		Model 2: + level-of-service var.	
		Coef.	*t*	*Coef.*	*t*
Alternative specific constants (Driving alone is base)					
Shared Ride (SR)	δ_{SR}	0.484	10.3	0.235	4.44
Public Transit (PT)	δ_{PT}	-4.008	-13.8	-4.515	-14.26
Walking (WA)	δ_{WA}	-2.748	-10.74	0.888	4.15
Bicycle (BI)	δ_{BI}	-3.657	-13.22	-3.532	-14.46
Level-of-service variables					
Cost ($) over household income (10,000$/yr) (Cost_Inc)	ρ_1			-9.466	-11.35
Travel time (min)	ρ_2			-0.104	-11.53
Random Effects					
var (θ_{Auto})	σ^2_{Auto}	2.724	4.721	7.823	7.03
var (φ_{Cost_Inc})	$\omega^2_{\rho_1}$			114.792	5.36
var (φ_{Time})	$\omega^2_{\rho_2}$			0.031	1.57
cov (φ_{Cost_Inc}, θ_{Auto})				-21.456	-7.26
cov (φ_{Time}, θ_{Auto})				-0.089	-0.99
cov (φ_{Time}, φ_{Cost_Inc})				1.405	5.15
Log-likelihood at convergence		-2103.025		-1845.037	

LAND USE VERSUS PRICING IMPACTS ON REDUCING DRIVING

The second step sequentially enters sociodemographic variables and land use variables into the final model with intercept-only random effects derived from the first step. For example, Model 3 estimates a model by introducing only sociodemographic and level-of-service variables (Table 6.4). Model 4 adds neighborhood-level land use variables but still restricts their random effects. Five land use variables are included: population and employment density, sidewalk density, land use mixed entropy, and distance to the nearest activity center (Table 6.4). Model 5 continues to add random effects between and within neighborhood levels. These three models are compared and applied to validate land use and pricing effects on reducing driving.

By comparing Models 3–5, one can conclude that the estimates of LOS coefficients are relatively robust. As expected, both coefficients of the cost and time variables are negative and both are significant at the 1% confidence level. When the median household income is set at $52,780 in Austin, from the 2010 census,[22] the estimated value of time is $4.69 per hour for nonwork trips in Model 3 and $4.65 per hour after land use attributes are added in Model 4.

The coefficients of sociodemographic variables demonstrate that individuals living in a larger household are more likely to use the vehicle. Individuals in households with more cars have a higher probability of selecting driving alone and shared ride modes. The nonwhite group appears more likely to use the nonauto mode but this effect is not statistically significant. The coefficient of FEMALE is significant at the 0.10 confidence level, suggesting that women in the Austin area depend more on vehicle travel. Also, employed travelers are more likely to drive for nonwork activities than the unemployed. The young group under 20 years old prefers to travel by public transit,

[22] http://www.clrsearch.com/Austin_Demographics/TX/Household-Income

walking, and bicycling, whereas the elder group over 60 years old appears to rely more on driving.

After adding land use variables into Model 3, Model 4 has a higher log-likelihood value at convergence, from –1,809.93 to –1,800.02 (Table 6.4). The variance components for the intercept also change from 3.45 to 3.10, implying that the selected land use variables explain 10% (i.e., (3.45 – 3.10)/ 3.45 = 0.101) of the variations in the average log-odds of mode choice across neighborhoods. In both Model 3 and Model 4, all the effects of LOS variables and sociodemographic variables are in the same direction; the magnitude of related coefficients and t-statistics are similar. These findings suggest that these variables preserve their significance to predict mode choice after land use variables are controlled. Also, land use variables have an independent influence on mode choice even after the effects of social demographics, travel cost, and time are monitored. These findings are consistent with previous studies in other cities (e.g., Cervero, 2002; Cervero & Kockelman, 1997; Zhang, 2004) despite the fact that these studies did not consider the contextual effects.

The coefficients of the land use variables have the expected signs, although only two factors are significantly above the 0.05 confidence level (Table 6.4). For example, residents living in higher population and employment density areas correlate with a lower probability of driving for their nonwork activities but the impact of population density is not statistically significant. People living in a neighborhood that has highly mixed land use or that is near the mixed-use activity center are more likely to employ nonauto modes, such as walking and public transit, to accomplish their day-to-day operations. The *t*-statistics of relative coefficients are both larger than 1.65, approaching the edge of the 90% significant level.

Table 6.4 Models with Socio-Economic Attributes and Land Use Variables

Fixed Effects	Note	Model 3: + Fixed effects of Level-1 variables		Model 4: + Land-use contextual variables	
		Coef.	*t*	*Coef.*	*t*
Alternative specific constants (Driving alone is base)					
Shared Ride (SR)	δ_{SR}	0.375	7.60	0.374	7.58
Public Transit (PT)	δ_{PT}	0.499	0.85	-1.723	-0.67
Walking (WA)	δ_{WA}	2.935	5.03	0.701	0.27
Bicycle (BI)	δ_{BI}	0.155	0.26	-2.081	-0.8
Level-of-Service Variables					
Cost ($) over household income (10,000$/year)	ρ_1	-3.173	-10.86	-3.198	-10.85
Travel time (min)	ρ_2	-0.047	-9.49	-0.047	-9.49
Socio-demographic Characters (specific to Auto mode)					
Household Size	γ_1	0.331	4.32	0.302	3.63
Vehicles per Household Members	γ_2	5.819	9.59	5.789	9.4
Non-White	γ_3	-0.135	-0.43	-0.078	-0.24
Female	γ_4	0.396	1.9	0.381	1.81
Employed	γ_5	1.211	3.43	1.267	3.56
Age less than 20	γ_6	-1.219	-4.29	-1.2	-4.19
Age over 60	γ_7	0.798	1.56	0.87	1.69
Land-use Contextual variables (For the intercept of Auto)					
Population Density (persons/acre)	π_1			-0.016	-0.24
Employment Density (persons/acre)	π_2			-0.061	2.04
Sidewalk Density (miles/acre)	π_3			-3.268	-0.17
Entropy Index of Land-Use Mixture	π_4			-2.776	1.65
Distant to the Nearest Activity Center (mile)	π_5			0.123	1.66
Random Effects	*Note*	*Coef.*	*t*	*Coef.*	*t*
var (θ_{Auto})	σ^2_{Auto}	3.445	3.239	3.185	3.10
var (φ_{Cost_Inc})	$\omega^2_{\rho_1}$				
cov (φ_{Cost_Inc}, θ_{Auto})					
Log-likelihood at convergence		*-1809.934*		*-1800.022*	

Table 6.5 Models with Random Effects and Interaction Effect of Land Use and Cost Variables

Fixed Effects	Note	Model 5 + Random effects of the slope of Cost_Inc.		Model 6: + land-use variables to explain the slope variation	
		Coef.	*t*	*Coef.*	*t*
Alternative specific constants (Driving alone is base)					
Shared Ride (SR)	δ_{SR}	0.108	2.03	0.081	1.52
Public Transit (PT)	δ_{PT}	0.078	0.04	-5.091	-2.62
Walking (WA)	δ_{WA}	3.551	1.64	-1.562	-0.81
Bicycle (BI)	δ_{BI}	-0.075	-0.03	-5.324	-2.75
Level-of-Service Variables					
Cost ($) over household income (10,000$/year)	ρ_1	-7.473	-13.97	-7.768	-2.86
Travel time (min)	ρ_2	-0.065	-10.46	-0.068	-10.98
Socio-demographic Characters (specific to Auto mode)					
Household Size	γ_1	0.209	2.57	0.275	3.57
Vehicles per Household Members	γ_2	8.147	10.09	7.932	10.83
Non-White	γ_3	0.086	0.23	0.343	1.06
Female	γ_4	0.324	1.4	0.355	1.53
Employed	γ_5	1.321	3.31	1.572	3.78
Age less than 20	γ_6	-1.839	-5.83	-1.79	-5.68
Age over 60	γ_7	2.658	3.74	3.206	4.68
Land-use Contextual variables (For the intercept of Auto)					
Population Density (persons/acre)	π_1	-0.007	-0.12	-0.102	-2.35
Employment Density (persons/acre)	π_2	-0.056	-1.92	-0.098	-2.57
Sidewalk Density (miles/acre)	π_3	-11.72	-0.72	-6.936	-0.48
Entropy Index of Land-Use Mixture	π_4	-0.929	-0.37	-1.976	-1.99
Distant to the Nearest Activity Center (mile)	π_5	0.138	1.5	0.487	4.38
Land-use Contextual variables (For the slope of variable of Cost over Household Income)					
Population Density (persons/acre)	μ_1			-0.068	-0.63
Employment Density (persons/acre)	μ_2			-0.075	-0.80
Sidewalk Density (miles/acre)	μ_3			-48.078	-3.18
Entropy Index of Land-Use Mixture	μ_4			2.563	0.68
Distant to the Nearest Activity Center (mile)	μ_5			-1.49	-8.20

Table 6.5 (*Continued*)

Random Effects	*Note*	*Coef.*	*t*	*Coef.*	*t*
var (θ_{Auto})	σ^2_{Auto}	9.084	4.08	18.743	5.63
var (φ_{Cost_Inc})	$\omega^2_{\rho_1}$	106.177	6.28	69.893	6.39
cov (φ_{Cost_Inc}, θ_{Auto})		-18.262	-5.98	-20.909	-5.99
Log-likelihood at convergence		-1687.114		-1674.608	

A Need to Incorporate Land Use and Pricing Policies

This section focuses on the interaction effect between land use and travel cost variables and investigates whether neighborhood land use contexts modify the impact of travel pricing on mode choice and whether land use policies for reducing driving are more efficient when travel costs are raised. Two models are discussed: Model 5 allows the slope of the *Cost_Inc* variable ρ_1 to vary with neighborhoods and examines the random effects; Model 6 introduces land use variables to explain the random effects detected in Model 5.

Table 6.5 presents the estimated results of Model 5 and Model 6. After the random effect of the coefficient of the cost variable is relaxed (i.e., from Model 4 to Model 5), the log-likelihood value proliferates from –1,800.02 to –1,687.11. This indicates that the model randomizing *Cost_Inc*'s impact on mode choice is a significant improvement. The variance components of the slope of *Cost_Inc* in Model 5 and Model 6 are 106.18 and 69.89. This finding suggests that the five land use variables explain above 34% (i.e., (106.17 – 69.89)/ 106.18 = 0.342) of the between-neighborhood variance of the slope. Without investigating such contextual effects, the model may underestimate the impact of land use and travel cost on mode choice.

The underlying assumption in Model 6 is that a land use variable has a *direct* effect on mode choice as well as an "interacted" effect through modifying the influence of travel cost on mode choice. For example, a higher level of population or employment density is *directly* associated with a lower probability of driving, including driving alone and sharing a ride (Table 6.5). However, the interacted effect of density variables is insignificant. These findings suggest that denser development or densification policies may lead to less driving and more nonauto trips, but the effectiveness of pricing policies for reducing driving demand and congestion may make no difference in denser and auto-oriented neighborhoods.

In contrast, the interacted effect of the variable of sidewalk density is significant while the direct effect is not. This implies that pricing policies may be more efficient for reducing driving in pedestrian-friendly neighborhoods than in auto-oriented neighborhoods. This finding is consistent with some empirical studies that argue the congestion pricing policies are more welcome in areas supporting other travel alternatives like public transit and walking (Mahendra et al., 2012).

Similar to density variables, the local mixed-use index also has only significantly direct effects but insignificantly interacted effects on mode choice (Table 6.5). However, the variable of the distance to the nearest activity center, representing regional accessibility, has both significantly direct and interacted effects on mode choice. This implies that people living in proximity to an activity center will participate in less auto travel than those living far away from activity centers, and pricing policies may be more effective for those living in areas with higher levels of regional accessibility, e.g., near activity centers.

Table 6.6 Elasticities of Nonwork Mode Choice with Respect to Land-Use Contextual Variables

Land-use contextual variable		Driving Alone (DA)	Shared Ride (SR)	Public Transit (PT)	Walking (WA)	Bicycle (BI)
Population Density (persons/acre)	Direct	-0.016	-0.010	0.252	0.212	0.151
	Interacted					
Employment Density (persons/acre)	Direct	-0.008	-0.005	0.260	0.126	0.084
	Interacted					
Sidewalk Density (miles/acre)	Direct					
	Interacted	-0.030	-0.007	0.022	0.297	0.194
Entropy Index of Land-Use Mix	Direct	-0.024	-0.015	0.233	0.304	0.187
	Interacted					
Distance to the Nearest Activity Center	Direct	0.012	0.007	-0.192	-0.145	-0.087
	Interacted	0.031	0.007	-0.015	-0.273	-0.163

Notes: Only coefficients significant at 0.05 level are shown in the table. All Elasticities reported above are probability-weighted average individual elasticities for each mode (see Ben-Akiva & Lerman, 1985, p. 113). The elasticities of the 'interacted' effects of land use variables are calculated at the weighted average of travel cost over income (Cost_Inc).

Table 6.6 reports the *direct* and *interacted* probability-weighted average elasticities with respect to land use variables for *all* travel modes calculated by the estimates in Model 6. The table lists only significant elasticities. The magnitudes of elasticities of driving with respect to land use variables are small (i.e., below 0.03) while those elasticities of nondriving are relatively larger, ranging from 0.02 to 0.3. This finding suggests that land use policies may be more efficient to promote nonauto travel, especially walking and bicycling, than to restrict auto trips. For example, doubling the population density in a neighborhood can lead to a 25%, 21%, and 15% increase in the likelihood of riding transit, walking, and bicycling, respectively. However, the corresponding decreased percentages for choosing driving-alone or shared-ride mode are just 1.6% and 0.8%. As a result, nondriving modes are much more sensitive to densification policies than auto modes.

The interacted effects of land use policies may be as large as, but no more than, the direct impact. For example, the direct elasticity of driving alone with respect to the distance to the nearest activity center is 0.012 while the interacted elasticity is 0.031. The interacted elasticities of walking and bicycling are also larger than the corresponding direct elasticities. These interaction effects between land use and travel cost variables not only indicate that pricing policies could be more efficient by incorporating with land use policies, but also suggest that land use policies are more effective for reducing driving demand when the cost of driving is high.

These results support the theoretical and analytical findings in previous chapters, all of which call for incorporating land use and economic policies. Suppose there is congestion pricing levied on major roads of Austin, and the cost of driving increases. Congestion pricing may have less impact on those living in auto-oriented neighborhoods because they have no alternatives to driving. However, for those living in denser neighborhoods in proximity to activity centers, congestion pricing will probably make them leave their cars and use other travel modes. These evaluations are not limited to the congestion pricing. For example, when parking charges are raised, the cost differential decreases and driving becomes less likely. In areas with higher levels of sidewalk density the effect of raising parking charges on reducing driving is greater. For another example, when transit fare is raised, the cost differential between nondriving and driving will then increase. Consequently the probability of choosing transit will decrease as indicated by the negative coefficient of ρ_1. The positive coefficient of μ_3 in the Level 2 cost model indicates that the effect of raising transit fare on choosing transit, however, is smaller in areas near activity centers with high regional access, such as downtown, than in areas with low regional access, like a suburban neighborhood. These results indicate that the effectiveness of an economic policy like pricing may vary with different land use contexts—more compact and mixed-use land use can facilitate the function of pricing while low-density sprawling land use may damage the effects of economic policy associated with sustainable travel behavior.

SUMMARY

This chapter presented an empirical study using land use and travel data from the Austin metropolitan area to investigate land use versus pricing effects on travel mode choice as well as the interaction impact between land use and travel cost variables. This empirical research sought to echo the theoretical and analytical studies in previous chapters, which demonstrated a need for incorporating land use and economic policies to reduce excessive auto travel and congestion. Land use policies are necessary to correct planning failures, while pricing policies are remedies for market failures. This study also enriches recent empirical studies in planning for complementary land use and pricing policies (Guo et al., 2011; Lee & Lee, 2013).

This chapter first developed a multilevel logit model with LOS and sociodemographic variables estimated at the individual level and land use variables estimated at the neighborhood level. Random effects between and within levels were also investigated. The key assumption was that land use characteristics not only affect travel mode choice directly, but also play a role in shaping neighborhood contexts in which the impact of travel costs on mode choice would be modified (i.e., land use contextual effects). This assumption was examined using several comparable models.

The results suggest that neglect of land use contextual effects may lead to inaccurate estimation and misleading evaluation of transportation policy. Multilevel relationships between individual travel behavior and the neighborhood environment cannot collapse into the traditional single-level regression framework. Over 45% of the variations in driving mode choice are determined by neighborhood land use variables, and about 55% are between individuals.

Both land use and cost variables exert significant impacts on mode choice. The higher the travel cost to income ratio, the lower the probability of driving. Denser and more mixed-use developments significantly decrease the likelihood of auto travel and reduce driving frequency. These findings suggest that either land use or pricing policies alone reduce auto travel demand. The interaction effects between some land use variables and the cost variable are statistically significant. These land use variables include sidewalk density and access to the activity center. These findings support those found in previous chapters, which illuminate that land use policies could narrow down the marginal external costs that should be corrected by pricing, improving the effectiveness of pricing policies. In contrast, pricing policies reduce excess travel and congestion demand produced by planning failure; the same land use policies are more efficient once market failures are corrected.

CHAPTER 7: CONCLUSIONS AND FUTURE WORK

PRIMARY FINDINGS AND CONTRIBUTIONS

This dissertation research places traffic congestion in a broader context of land use and economic linkages and contends that congestion relief requires incorporating land use and pricing policies. Anti-congestion policies should target at the socially optimal level of traffic, rather than the free-flow status without congestion. This research concentrates on three research questions: when congestion is excessive, what causes excessive congestion, and which policies are most efficient for excessive congestion reduction.

Excessive congestion occurs when the individually desirable amount of auto travel exceeds the socially optimal level, in which the marginal social cost (MSC) of travel equals the marginal social benefit (MSB). This research articulates two underlying causes of excessive congestion: market failures from congestion and agglomeration externalities and planning failures from land use regulations such as exclusionary zoning and low-density zoning. While much literature recognizes the congestion externality as a source of excessive congestion (e.g. OECD, 2007; Anas and Lindsey, 2011), less explores how agglomeration externalities and planning failures shape the excessive congestion. This dissertation research filled the gap by developing a spatial general equilibrium framework internalizing congestion, agglomeration, and planning failures. The modeling framework was calibrated using data from the US cities and solved relying on computational simulations. This research relied on computational simulations of several policy scenarios to examine the moderate effect of agglomeration externalities on excessive congestion, estimate the diseconomy of congestion, theorize how market and planning failures together determine excessive congestion, and evaluate the optimal and practical policies for reducing excessive congestion. Simulation and empirical studies focused on two categories of congestion-relief policies: land use planning and congestion pricing. The following sections summarize primary findings in this dissertation.

Agglomeration Economies Moderate Excessive Congestion

Simulation findings suggest that excessive congestion could generate a large cost to road users and the society while the agglomeration economy could partially compensate such a cost. The existence of agglomeration externalities moderates the level of excessive congestion because a certain degree of congestion is desired by enhancing agglomeration. These findings demonstrate the importance of integrating congestion and agglomeration in an analytical framework. Anti-congestion policies can certainly reduce the congestion diseconomy but meanwhile would erode the agglomeration economy (Chapter 2 and 3).

An efficient policy for reducing congestion needs to balance the benefits from congestion reduction and the loss of reduced agglomeration economies. For example, the Pigouvian congestion toll alone policy is no longer socially optimal in cities with agglomeration externalities since this policy could reduce too many agglomeration economies over the benefits earned from congestion relief. In some simulation cases, an imposition of congestion pricing can even lead to a welfare loss than without such a pricing policy in the free market.

The optimal congestion pricing policies need to internalize agglomeration externalities and are often hard to design in practice (see discussions in Chapter 3). The optimal toll levels across locations may lie below the Pigouvian congestion toll level and even become negative in high-productivity areas. In the latter case, the toll is essentially a travel subsidy. These findings are consistent with several economic studies (Arnott, 2007; Wrede, 2009; and Borck and Wrede, 2009); however, rather little empirical research has examined the effects of such a travel subsidy.

The Congestion Diseconomy Could Be Much Smaller Than the Total Cost of Congestion

This dissertation research enriches recent debates on the cost of congestion. Although the estimation of the total cost of congestion has received substantial critiques (Goodwin, 2004; OECD, 2007), many anti-congestion projects in practice remains largely relying on the estimation of the total cost of congestion (e.g., Grant-Muller and Laird, 2006; OECD, 2007; Bilbao-Ubillos, 2008; Litman, 2009). This research suggests that policies aiming to reduce all cost of congestion can largely erode agglomeration economies and cause huge welfare loss. The free-flow speeds and traffic are probably never socially desirable. Studies should focus more on the estimation of the diseconomy of congestion, with considering both costs and benefits of travel captured in the transportation, land use, and economic systems. Simulations here show that the congestion diseconomy is about 5% to 23% of the total cost of congestion, varying wth the levels of congestion and agglomeration. Thus, the congestion diseconomy, or the social loss of excessive congestion, could be much smaller than the total cost of congestion; a large share of congestion cost is offset by travel benefits, including commute profits from wage income, shopping-travel benefits from goods consumption, and crowding benefits from agglomeration.

Planning Failures Can Cause Excessive Congestion

This dissertation research is among the first to present a theoretical analysis on how planning failures cause excessive congestion. The objectives of policies for correcting planning failure differ from those remedies for market failure. While policies for reducing excessive congestion from market failures often aim to achieve the optimal level of congestion, those policies for correcting planning failures aim at removing all regulations that deter market outcomes. The target level of congestion is thus the equilibrium level in the free market without any regulations. However, realistic cities are full of regulations and externalities. In most cases, both market and planning failures contribute to excessive congestion and social inefficiency. Evaluating congestion-relief policies needs an innovative analytical framework enabling to internalize both failures.

The identification of planning failure needs to justify whether market-desired land use patterns are constrained by land use regulations. The simulation in this research finds a type of optimal land use patterns with the firm cluster decentralized away from the urban core. However, this type of urban form is seldom found in reality. One potential reason is that planning failure from exclusionary zoning regulations restricts the decentralization of firms and jobs. Also, simulations also detect that low-density zoning regulations restrict denser development and lead to urban sprawl.

Based on policy scenarios, planning failures from exclusionary zoning and low-density zoning regulations could increase travel distance and auto dependence, produce excessive driving demand, and accumulate excessive traffic on the streets and highways. These findings correspond to the empirical studies from Cervero (1996) and Levine (2006). Planning failure can play a dominant role leading to excessive congestion and social inefficiency when regulations largely restrict the market-desirable development and when the external cost of congestion is small. Since planning failures are insensitive to pricing signals, remedies for planning failures require regulatory reform and innovative land use planning.

Even First-Best Pricing Could Be Low Effective When Planning Failures Dominate

Proponents of pricing policy often believe that economic policies are superior to land use policies and regard land use planning as a second-best and replaceable strategy for congestion mitigation (e.g., Brueckner, 2007; Kono, 2012). This perspective could be true if no planning failure exists in cities. Unfortunately, most municipalities have local land use regulations, and our actual markets are never the free market result or the social optimum. Most regulations are insensitive to pricing signals at least at the short term. Facing both markets and planning failure, neither congestion pricing nor land use planning alone could fully reduce excessive congestion. Even the first-best pricing policy (e.g., a combination of Pigouvian congestion toll and Pigouvian labor subsidy) may be of very low effectiveness for congestion relief. For example, in a simulated city with 600,000 workers and jobs, if all firms are restricted to the urban core of 2-mile radius, even the first-best pricing policy can only reduce 4% of excessive congestion. Planning failure dominates in such a city, land use planning strategies via regulatory reform or promoting alternative development could reduce up to about 96% of excessive congestion and social inefficiency (Chapter 3).

This research does not suggest that land use planning strategies are always superior to pricing policies. Market failures are still the dominant cause leading to excessive congestion in many cases. Designing more efficient policies should bridge perspectives from economics and planning and incorporate land use and pricing policies in practice.

Efficient Pricing Policies in Practice Should Concern Agglomeration Benefits and Land Use Impacts

First-best pricing policies can maximize social welfare and eliminate all excessive congestion but are difficult to implement in practice, especially when recognizing spatial variations (Chapter 3). The first-best toll lies below its related marginal externality cost (or benefit), as also found in Arnott (2007) and Thissen et al.'s (2011) empirical analysis for the Netherlands. However, the specific optimal tolls levied on drivers can be both positive and negative, varying over space. While both first-best tolling and subsidy policies are equivalent in theory (Chapter 3), some may suggest that it is easier to subsidize firms than charge drivers; because the public prefers to earn the subsidy rather than pay the tolls and subsidizing a few firms may be much easier than tolling the masses (e.g., Arnott, 2007). However, my findings challenge this belief since the aggregate optimal subsidy will equal the aggregate optimal toll. If the optimal toll is a true negative tax, firms need to pay labor tax, rather than receive a positive subsidy when hiring/paying a worker. These findings demonstrate that it is important to evaluate the potential impact of pricing on agglomeration economies.

This dissertation research also investigates several more practical pricing policies, such as VMT tax and a cordon toll and compares their land use, travel, and rent impacts under both monocentric and polycentric settings (Chapter 4). The practical pricing policies like VMT Tax and Cordon Toll can partially reduce excessive congestion around 30% of total excessive congestion from simulations in this research. They can also produce significant decreases in average commute distance and travel costs and deliver more compact city form. Both the VMT tax and the cordon toll can generate somewhat higher household utility although their welfare improvements are less than that of the Pigouvian congestion toll policy. They can also partially reduce excessive congestion.

The VMT tax is predicted to generate a more compact urban form than the Pigouvian congestion toll policy, by incentivizing firms and households to locate more closely, to reduce commuting distance, while the Pigouvian congestion toll toll may allow firms and/or households to trade a longer travel distance for less congestion. The compactness effects are also reflected in the findings that all three congestion pricing policies can reduce daily travel distance by more than 10% (with results ranging from 10% to 20%, varying across settings and policies).

All congestion pricing policies facilitate job decentralization. Simulation results show how Pigouvian tolling of travel in the polycentric setting can cause many jobs (17%

in this example) to leave the central business district (CBD) and relocate to a relatively dense but suburban ring. To achieve city-wide welfare gains, efficient land use regulations should permit such job decentralization. Simulations also illuminate how simple, distance-based tolls generate lower welfare improvements, but stimulate similar land use effects. A cordon toll may agglomerate firms in a smaller CBD if monocentricity is required by the model or re-agglomerate companies in a polycentric sub-center ring of development.

Efficient Land Use Policies in Practice Should not Deter Job Decentralization and Residential Densification

This dissertation research surveys four land use planning strategies, including urban growth boundaries (UGBs) and firm cluster zoning for correcting market failures and residential densification policies and a job-decentralization policy by building a suburban employment center for correcting planning failures (Chapter 5). Simulation results discover that the UGB regulations may partially correct distortions in both transport and labor markets, but may worsen land market distortion via the residential rent-escalation effects. The congestion-relief and welfare improvement impact of UGBs are trivial. Thus, this finding suggests that UGBs are not an efficient policy for reducing congestion. The firm cluster zoning policies by regulating a zone’s land use exclusively for firm/business use are probably more efficient than the UGB policies for reducing congestion and enhance agglomeration. They can generate welfare improvement closer to the first-best levels and will not bring much excessive congestion and excessive escalation of housing rents, avoiding the housing affordability issue raised by the UGB policies.

Densification policies by relaxing low-density zoning regulations in a particular planning area can partially correct planning failure from low-density zoning and reduce excessive congestion. A city with more stringent low-density zoning regulations probably allows more areas for denser development. On the other hand, when firms are regulated against decentralization (e.g. due to exclusionary zoning), policies aiming at moving jobs to the subcenter need to incorporate both land use and pricing strategies. Simulation findings in this research suggest that a subsidy to firms (e.g., labor subsidy for hiring workers) is a key trigger to firm decentralization. The optimal setting of job decentralization can reduce about half of excessive congestion and improve half of welfare. Job decentralization also brings other attractive land-use and transportation consequences, including significant drops in firm rents, VMT, and congestion, despite the disadvantage from declined agglomeration economies.

Congestion Relief Should Incorporate Land Use and Economic Policies

The existence of externalities and regulations cause market and planning failures, leading to excessive congestion. While pricing policies are difficult to remedy planning failures, an integration of land use planning and pricing strategies are needed to correct both failures. This dissertation research first relies on the analytical model and computational simulations to quantify the effectiveness of land use planning-alone, congestion pricing-alone, and a complementary land use and pricing policies for reducing excessive congestion (Chapter 3). Simulation findings suggest that incorporating land use and pricing policies is more efficient than the other policies, although the efficiency of the land use planning-alone policy is proximity to the combination policy when planning failure dominates. Similarly, when cities have less restrictive land use regulations, the congestion pricing-alone policy can generate efficiency close to the combination policy.

Chapter 6 also provides an empirical study relying on land use and travel data from the city of Austin. Despite this study does not directly measure the excessive congestion, it investigate the land use versus pricing effects on travel mode choice, as well as the interaction impact between land use and travel cost variables. Modeling results suggest that the blind to land-use contextual effects may result in inaccurate estimation and misleading evaluation of transportation policy. Multilevel relationships between individual travel behavior and neighborhood environment cannot collapse into the traditional single-level regression framework. Neighborhood-level land use variables, determine over 45% of the variations in driving mode choice, and about 55% are within neighborhoods.

Both land use and cost variables exert significant impacts on model choice. The higher the travel expenses over income, the less probability of driving occurs. Denser and more mixed-use development can significantly decrease the likelihood of auto travel and reduce driving frequency. These findings suggest that either land use or pricing policies alone can reduce auto travel demand. The interaction effects between some land use variables and the cost variable are statistically significant. These land use variables include sidewalk density and access to the activity center. These findings support those found in previous chapters, which illuminate that land use policies could narrow down the marginal external costs that should be corrected by pricing, improving the effectiveness of pricing policies. In contrast, pricing policies can reduce excessive travel and congestion demand produced by planning failure; the same land use policies can be more effective once market failures are corrected.

Methodological Innovation

This research serves as a fresh contribution to three important methodological debates surrounding multiple urban externalities and planning failures. The first debate focuses on the modeling framework applied in analyzing interactions between externalities. Both analytical and simulation results in this chapter support previous studies' results, supporting the notion that it is important to use general equilibrium frameworks, rather than non-spatial or partial equilibrium models, and internalize spatial interactions when analyzing urban externalities. The model here further suggests that it is critical to endogenize firms' land use decisions (e.g., decentralization and agglomeration), which are always neglected in the traditional monocentric model. The Pigouvian congestion toll-alone or Pigouvian labor subsidy-alone policies could be the optimal policies in the partial equilibrium model that are internalizing congestion or agglomeration externalities only. However, in more realistic cities with both externalities, the Pigouvian congestion toll alone or Pigouvian labor subsidy alone policies could lead to significant land market distortions and welfare loss. Only by considering the land use decisions of both firms and households can one quantify such policy impacts. This work does not imply that aspatial, partial equilibrium, or monocentric models should be not used for policy analysis, but that decision makers should recognize the potential distortions when using such models in cities full of distinctive externalities.

Second, while less research has formulated the economics of planning failure, this research develops models internalizing both planning and market failures. This modeling development allows the economic model, e.g., the spatial general equilibrium model, for investigating more realistic land use and transportation planning issues.

Third, the empirical study in this dissertation developed a multilevel multinomial logit (MML) model to examine the interaction effects between land use and travel cost variables. Neighborhood-level land use characteristics are assumed to formulate a neighborhood context that modifies the impact of travel cost on travel mode choice. This model is innovative in the application of land use and travel studies.

LIMITATIONS AND FUTURE WORK

Explicit understanding of how market and planning failures cause excessive congestion is important for designing efficient policies of congestion mitigation. The theoretical, simulation, and empirical analysis in this dissertation have investigated the efficiency of land use and pricing policies on congestion reduction, economic development, land use change, and social welfare improvement. These investigations are combined to offer an innovative interpretation of congestion. However, traffic congestion that people face every day is much more complicated than that formulated in theoretical and empirical models. The following is a discussion of key opportunities for extension of this research and modeling analysis.

First, more anti-congestion policies can be discussed using the simulation framework developed here. For example, the theoretical and simulation models can be extended to investigate the supply-side policies, such as expanding highways and building new roads, and evaluate their effects on congestion reduction. The model can internalize land use (e.g., areas and locations) for transportation infrastructure (e.g., streets and highways), as done by Wheaton (1998) and De Lara et al. (2013). One can solve for this model relying on simulations and determine the optimal road space. Also, the model could be extended to consider more than one travel mode (like transit), to reflect differences in congestibility and mode-based pricing impacts. This extension is useful for evaluating whether public transit development helps to alleviate highway traffic congestion. Several monocentric studies have explored the effects of public transit on congestion and land development (e.g., Baum-Snow, 2007; Kilani, Leurent, & De Palma, 2010; Buyukeren and Hiramatsu, 2015).

Second, allowing for travel mode and trip scheduling flexibility is important in appreciating congestion toll effects. A model that enables a gradual, dynamic city evolution is important to explore. The one-shot, static equilibrium typical of papers in urban economics is never achieved in practice. In reality, most cities already exist, and populations regularly expand, in the midst of great uncertainty and imperfect information, along with speculation and another complex -- but very realistic – human behaviors. Several recent studies have explored this topic (e.g., Boucekkine, Camacho, & Zou, 2009; Desmet and Rossi-Hansberg, 2010). Zhang and Kockelman (2015) developed a dynamic spatial general equilibrium model to enable more heterogeneous agents, land use detail, population growth, and transitional dynamics, and investigates the zoned-out effects on land use and housing affordability. These applied models are powerful for simulating the reality, although they are difficult to derive the optimal congestion level. They can evaluate more flexible land use and pricing policies applied in our living cities and easily connect congestion with urban dynamics of demographics and land use.

Third, more empirical studies are needed. It is challenging to measure excessive congestion directly since it is a relative concept. Empirical studies should break congestion issues into several aspects, including direct impacts of congestion on mobility and accessibility and indirect impacts on housing selection, land development, and agglomeration economies. Moreover, only limited studies have examined the effectiveneses of an incorporation of both land use and pricing policies as a strategy for reducing driving, VMT, and travel delay (Langer and Winston, 2008; Guo et al., 2011; Lee and Lee, 2014). Also, while many empirical studies have focus on the wider economic effects of transportation infrastructures (e.g., Graham, 2005; Lakshmanan, 2011), fewer of them have tackled the wider economic effects of congestion-relief policies. Empirical studies should examine the impact of congestion on surrounding land and economic development. Since many cities have imposed congestion tolls, it is important to evaluate realistic tolling impacts on traffic performance, land development, and social welfare.

Appendices

A1: Proof of Proposition 3.1

(a) Since utility maximization and expenditure minimization are fully equivalent, the minimum expenditure at the equilibrium utility $\bar{u}$ equals the net income $y(x, x_w)$, i.e., $y(x, x_w) = e(r_h(x), \bar{u})$. Since $r_h(x)$ is only relevant to location x, one has $y(x, x_w) \equiv y(x)$. Under utility maximization, $c^*(x, x_w) = c^*\big(y(x, x_w)\big) \equiv c^*\big(y(x)\big) = c^*(x)$, and $q^*(x, x_w) = q^*\big(r_h(x), y(x, x_w)\big) \equiv c^*\big(r_h(x), y(x)\big) = c^*(x)$.

(b) From the first-order conditions of this utility maximization problem, one can derive the following: $c(x) + q(x)u_q/u_c = y(x)$. In combination with $u\big(c(x), q(x)\big) = \bar{u}$, one calculates that $q^*(x) = q^*(y(x), \bar{u})$ and $c^*(x) = c^*(y(x), \bar{u})$.

(c) Since $t(x, x) = 0$, $y(x) = y(x, x) = w(x) - t(x, x) = w(x)$.

(d) Since $w(x) \equiv w(x_w) - t(x, x_w), \forall\, x_w > 0$, $w(x_w) - w(x) = \int_x^{x_w}[t(s) + \tau(x)]ds$. Thus,$w'(x) = t(x) + \tau(x)$. From (c), $y'(x) = t(x) + \tau(x)$.

A2: Proof of Proposition 3.2

The solutions to the social optimum is achieved by determining each of six factors, $\{n(x), q(x), c(x), \theta_f(x), F(x), t(x)\}$, at each location x so as to maximize the households' utility level under constraints (A1)-(A5), as defined in Problem A.

Problem A. Choose functions $n(x), q(x), c(x), \theta_f(x), F(x)$ so as to maximize

$$u\big(c(x), q(x)\big)$$

subject to

(A1) $\int_0^{\bar{x}} \left\{2\pi x \left[\theta_f(x)\delta n(x)^\kappa F(x)^\gamma - \frac{\theta_h(x)}{q(x)}c(x) - (1-\theta_t)R_A\right] - t(x)D(x)\right\} dx \geq 0$

(A2) $\theta_h(x) + \theta_f(x) + \theta_t = 1$

(A3) $F(x) = \zeta \int_0^{\bar{x}} \int_0^{2\pi} r\theta_f(r)n(r)e^{-\zeta l(x,r,\psi)} d\psi dr$

(A4) $|t(x)| = \varphi + \rho\left(\frac{|D(x)|}{2\pi x\theta_t}\right)^{\sigma}$

(A5) $D'(x) = 2\pi x\left(\frac{\theta_h(x)}{q(x)} - \theta_f(x)n(x)\right)$

for all $x \in [0, \bar{x}]$, with boundary conditions:

(A6) $D(0) = 0 \text{ and } D(\bar{x}) = 0$

(A7) $r(\bar{x}) = R_A$

(A8) $\int_0^{\bar{x}} 2\pi x\frac{\theta_h(x)}{q(x)}dx = N$

Equations (A1)-(A8) are present in the body text of this paper, with the exception of constraint (A1), which guarantees a non-negative net social surplus. Given that aggregate land rents (net of the opportunity costs) are equally returned to each household (in this closed system), the net surplus equals the total value of production, minus general consumption, minus and opportunity costs of land, and minus workers' commute costs.

The Hamiltonian function of the Problem A is given by:

$$H(n, F, q, c, \theta_f, t, D, \beta_1, \beta_2, \beta_3)$$
$$= \lambda(x)u(c(x), q(x))$$
$$+ 2\pi x\left[\theta_f(x)\delta n(x)^{\kappa}F(x)^{\gamma} - \frac{1-\theta_t-\theta_f(x)}{q(x)}c(x) - (1-\theta_t)R_A\right]$$
$$- t(x)D(x) + \left(\beta_1(x)F(x) - \zeta\int_0^{\bar{x}}\int_0^{2\pi}\beta_1(r)r\theta_f(r)n(r)e^{-\zeta l(x,r,\psi)}d\psi dr\right)$$
$$+ \beta_2(x)\left(t(x) - \varphi - \rho\left(\frac{|D(x)|}{2\pi x\theta_t}\right)^{\sigma}\right)$$
$$+ \beta_3(x)2\pi x\left(\frac{1-\theta_t-\theta_f(x)}{q(x)} - \theta_f(x)n(x)\right)$$

From the conditions of the maximum principle, some of the first-order conditions are derived as:

(A9) $\frac{\partial H}{\partial n} = \frac{\partial H}{\partial n(x)} + \frac{\partial H}{\partial n(r)} =$

$2\pi x\theta_f(x)[\delta\kappa n(x)^{\kappa-1}F(x)^{\gamma} - \beta_3(x)] - \zeta\int_0^{\bar{x}}\int_0^{2\pi}\beta_1(r)r\theta_f(r)e^{-\zeta l(x,r,\psi)}\,d\psi dr = 0$

(A10) $\frac{\partial H}{\partial F} = 2\pi x\theta_f(x)\gamma\delta n(x)^{\kappa}F(x)^{\gamma-1} - \beta_1(x) = 0$

(A11) $\frac{\partial H}{\partial D} = -\beta_3{}'(x) \rightarrow \beta_3{}'(x) = t(x) + \rho\sigma\left(\frac{|D(x)|}{2\pi x\theta_t}\right)^\sigma$

From (A9) and (A10), one can obtain the following relationship in the firm cluster:

(A12) $\delta\kappa n(x)^{\kappa-1}F(x)^\gamma = \beta_3(x) - \zeta\gamma\delta\int_0^{\bar{x}}\int_0^{2\pi} r\theta_f(r)n(r)^\kappa F(r)^{\gamma-1}e^{-\zeta l(x,r,\psi)}\, d\psi dr$

When firms' profits are maximized, from Eq. (3.16), one can derive the following:

(A13) $\delta\kappa n(x)^{\kappa-1}F(x)^\gamma = w(x) - s(x)$

In a socially optimal city, both conditions (A12) and (A13) should be satisfied. Thus,

(A14) $\beta_3(x) = w(x) - s(x) + \zeta\gamma\delta\int_0^{\bar{x}}\int_0^{2\pi} r\theta_f(r)n(r)^\kappa F(r)^{\gamma-1}e^{-\zeta l(x,r,\psi)}\, d\psi dr$

Comparing the first-order condition (A11) and Eq.(A14), one can derive the following equations:

(A15) $w'(x) = \left(s(x) - \zeta\gamma\delta\int_0^{\bar{r}}\int_0^{2\pi} r\theta_f(r)n(r)^\kappa F(r)^{\gamma-1}e^{-\zeta l(x,r,\psi)}\, d\psi dr\right)' + t(x) + \rho\sigma\left(\frac{|D(x)|}{2\pi x\theta_t}\right)^\sigma$

When household's utility is maximized, from Proposition 1d and (A15), one can obtain the following relationship:

(A16) $\tau(x) - \rho\sigma\left(\frac{|D(x)|}{2\pi x\theta_t}\right)^\sigma = \left(s(x) - \zeta\gamma\delta\int_0^{\bar{r}}\int_0^{2\pi} r\theta_f(r)n(r)^\kappa F(r)^{\gamma-1}e^{-\zeta l(x,r,\psi)}\, d\psi dr\right)'$

In order to fulfill Eq. (A16) for each location *x*, we have three strategies:

(a) A combination of two instruments:

(A17) $$\begin{cases} \tau(x) = \tau_{pct}(x) = \rho\sigma\left(\frac{|D(x)|}{2\pi x\theta_t}\right)^\sigma \\ s(x) = s_{pls}(x) = \zeta\gamma\delta\int_0^{\bar{r}}\int_0^{2\pi} r\theta_f(r)n(r)^\kappa F(r)^{\gamma-1}e^{-\zeta l(x,r,\psi)}\, d\psi dr \end{cases}$$

(b) When $s(x) = 0$, $\tau(x) = \tau_{pct}(x) - s_{pls}{}'(x)$, which represents the first-best toll at location *x*. Given Eq. (A5), the total toll revenues thus equal:

(A18) $\int_0^{\bar{x}} \tau(x)D(x)dx = \int_0^{\bar{x}} \left(\tau_{pct}(x) - s_{pls}'(x)\right) D(x)dx =$

$\int_0^{\bar{x}} \tau_{pct}(x)D(x)dx - \int_0^{\bar{x}} 2\pi x\theta_f(x)n(x)s_{pls}(x)dx$

Therefore, revenues provided by optimal tolling across the region equal the total congestion externality costs of the work commute traffic (or total revenues from the Pigouvian congestion toll policy) minus total agglomeration externality benefits (or total payments under the Pigouvian labor subsidy policy).

(c) When $\tau(x) = 0$, $s'(x) = s_{pls}'(x) - \tau_{pct}(x)$. Thus, $s(x_i, x) = s_{pls}(x_i, x) - \int_{x_i}^{x} \tau_{pct}(x)dx$, which represents the first-best subsidy to workers living at x_i but working at x. Given Eq. (A5) and the fact that $\theta_h(x) = 0$, the total first-best subsidies equals the following:

(A19) $\int_0^{\bar{x}} 2\pi x\theta_f(x)n(x)s(x)dx = -\int_0^{\bar{x}} s(x)D'(x)dx = \int_0^{\bar{x}} s'(x)D(x)dx - s(x)D(x)\Big|_0^{\bar{x}} = \int_0^{\bar{x}} \left(s_{pls}'(x) - \tau_{pct}(x)\right) D(x)dx = \int_0^{\bar{x}} 2\pi x\theta_f(x)n(x)s_{pls}(x)dx - \int_0^{\bar{x}} \tau_{pct}(x)D(x)dx$

Thus, total optimal subsidy to workers equals the overall benefits of agglomeration to the region's firms minus total external congestion costs.

A3: A NESTED FIXED-POINT ALGORITHM

The following computational procedures describe detailed algorithms for solving for the optimum and equilibria defined in Chapters 3-5.

Step 0: Set the model's parameters (following Table 3.2) and tolerances. In our simulation, the tolerances $\epsilon_1 \sim \epsilon_6$ are all set at 1.

Step 1: Given an initial function θ_f^0, there exist a set of equilibrium functions $\{F^*, w^*, q^*, n^*, D^*, t^*, \tau^*\}$ and equilibrium values $\{y_{rent}^*, y_{toll}^*, y_{suby}^*\}$ that solve Problem A.

Step 1.0: Define the initial values of the function $\theta_f^0(x)$. In order to check the existence of multiple equilibria, simulations in this paper often use several different initial functions of $\theta_f^0(x_i)$, such as:

$$\theta_f^0(x_i) = \begin{cases} 1-\theta_t, x_i \epsilon [1{,}100] \\ 0, \quad others \end{cases} \text{ or } \begin{cases} 1-\theta_t, x_i \epsilon [1{,}600] \\ 0, \quad others \end{cases}$$
$$\text{or } \begin{cases} 1-\theta_t, x_i \epsilon [100{,}700] \\ 0, \quad others \end{cases} \text{ or } \begin{cases} 1-\theta_t, x_i \epsilon [300{,}800] \\ 0, \quad others \end{cases}$$

Step 1.1: Given a set of initial values, F^0,$y_{rent}^0, y_{toll}^0, y_{suby}^0$, one can find an unique wage at the city center $w^*(x_1)$ and an unique utility level u^* that satisfies the first-order conditions and the Maximum Principle conditions of Problem A.

Step 1.1.0: Define the initial values of F^0,y_{rent}^0, y_{toll}^0 and y_{suby}^0. Our simulations set y_{rent}^0, y_{toll}^0 and y_{suby}^0 as 2000, 0, and 0. The initial values of $F^0(x_i)$ vary with the setting of $\theta_f^0(x_i)$. For example, $F^0(x_i) = \theta_f^0(x_i) \times 10^6$.

Step 1.1.1: Given an initial utility u_0, select an initial wage at x_1, $w_0(x_1)$, calculate $q_0(x_1)$ and $n_0(x_1)$ by Eqs. (3.7) and (3.16), then $D_0'(x_1)$ using Eq. (3.23). Given $D_0(x_1)$ is known, calculate $D_0(x_2) = D_0(x_1) + D_0'(x_1)\Delta x$. Given $D_0(x_2)$, calculate $t_0(x_2)$ by Eq. (3.5) and $\tau_0(x_2)$ under different policy scenarios as defined in Table 3.1. Given $t_0(x_2)$, $\tau_0(x_2)$, and $w_0(x_1)$, calculate $w_0(x_2)$=$w_0(x_1)$+$(t_0(x_2) + \tau_0(x_2))\Delta x$. Repeat the previous calculation, one can derive a set of paths $\{w_0(x), q_0(x), n_0(x), D_0(x), t_0(x), \tau_0(x)\}$, $\forall x_1 \leq x \leq x_I$. These iterative calculations stop at x_I, that satisfies:

$$D_0(x_{I-1}) \leq 0 \; and \; D_0(x_I) \geq 0$$

Step 1.1.2: Calculate the edge household bid-rent $r_h(x_I)$. If the boundary condition satisfies

$$\begin{cases} |r_h(x_I) - R_a| < \epsilon_1, if \; the \; instrument \; is \; not \; UGB \; policies \\ x_I = x_{ugb}, \quad if \; the \; instrument \; is \; UGB \; policies' \end{cases}$$

return $w^*(x_1) = w_0(x_1)$ and go to Step 1.1.3. Instead, repeat Step 1.1 to find a continuous series of central wage $w_0(x_1)$, $w_1(x_1)$, …, $w_{n_w}(x_1)$ until finding the $w^*(x_1)$.

Step 1.1.3: Based on $w^*(x_1)$, calculate a set of equilibrium function $\{w^*, q^*, n^*, D^*, t^*, \tau^*\}$. If the city population reaches the given number, i.e., satisfying:

$$\left|\sum_{i=1}^{I} 2\pi x \theta_f^0(x_i) n^*(x_i) \Delta x - N\right| < \epsilon_2$$

return $u^* = u_0$ and go to Step 1.2. Else, adjust the value of u^0 and repeat the Step 1.1.1and 1.1.2 to find a continuous series of u_0^0, u_1^0, …, $u_{n_u}^0$until the population condition is satisfied

Step 1.2: Based on u^* and $\{w^*, q^*, n^*, D^*, t^*, \tau^*\}$, compute land rent as follows:

$$r(x) = \begin{cases} r_f(x), & if\ \theta_f^0(x) > 0\ and\ r_f(x) > R_a \\ r_h(x), & if\ \theta_f^0(x) = 0\ and\ r_h(x) > R_a \\ R_a, & if\ \ r_h(x) \le R_a and\ r_f(x) \le R_a \end{cases}$$

Calculate y_{rent} and $F(x)$ using Eq.(3.22) and Eq. (3.13)[23]. And calculate y_{toll} and y_{suby} according to the definition in different policy scenarios (Table 3.1). If the following conditions are satisfied:

$$|y_{rent} - y_{rent}^0| < \epsilon_3$$
$$|y_{toll} - y_{toll}^0| < \epsilon_4$$
$$|y_{suby} - y_{suby}^0| < \epsilon_5$$

[23] The calculation here of the integral in *F(x)* follows LRH's (2002) to use an approximation over a radial coordinate system, while Dong and Ross (2015) suggested that the approximation of the production externality function *F*(*x*) over a rectangular grid system is more precise than a radial coordinate system. Dong and Ross argued that the radial coordinate approximation could lead to inaccurate simulated outcomes, such as a decrease in job density near the city center, which should never occur in theory. Our simulation experience suggests that the two coordinate systems could generate the same approximation of *F(x)* if the interval of angle (or grid) is small enough. While both approximation approaches could bring inaccuracy, we believe the imprecision generated by radial coordinate approximation is tolerable in our modeling simulations.

$$\max_{\forall x_i} |F(x_i) - F^0(x_i)| < \epsilon_6$$

return $y^*_{rent} = y_{rent}$, $y^*_{toll} = y_{toll}$, $y^*_{suby} = y_{toll}$, $F^* = F$ and go to Step 2. Else, replace y^0_{rent}, y^0_{toll}, y^0_{suby}, and F^0 with y_{rent}, y_{toll}, y_{toll}, and F, and go back to Step 1.1.

Step 2: Based on the equilibrium functions $\{F^*, w^*, q^*, n^*, D^*, t^*, \tau^*\}$ and equilibrium values $\{y^*_{rent}, y^*_{toll}, y^*_{suby}\}$, calculate a new land use share function $\theta_f(x)$ using Eqs. (3.20) and (3.21). If $\theta_f(x) = \theta^0_f(x)$, the simulation ends. Else, set $\theta^0_f(x) = \theta_f(x)$ and go back to Step 1. If a converged θ_f were not found, one could try different initial values of θ_f and/or F.

A4: Discussion on Mixed Urban Configurations under Multiple Externalities

The existence of mixed-use equilibrium has been discussed in several studies (e.g., Ogawa and Fujita, 1982; Lucas and Rossi-Hansberg [LRH], 2002; Rossi-Hansberg, 2004; Duranton and Puga, 2014). These require urban models that endogenize both firms' and households' location decisions and their interactions, which are difficult to examine through traditional monocentric models. The model developed in this paper extends to internalize both agglomeration and congestion externalities, thus enabling to discuss the existence of mixed urban configurations with multiple externalities. Our theoretical and simulation analyses suggest that the partially or completely mixed land use pattern could be an equilibrium solution when the congestion level increases or the agglomeration scale decreases, as found in those existing literature (e.g., Ogawa and Fujita, 1982; Lucas and Rossi-Hansberg, 2002; Duranton and Puga, 2014). However, our findings also show that mixed-use equilibrium allocation is never *Pareto*-optimal in either the free-market or first-best cases. There exists a non-mixed use equilibrium that produces an increase in *Pareto* efficiency, compared to the mixed-use equilibrium allocation.

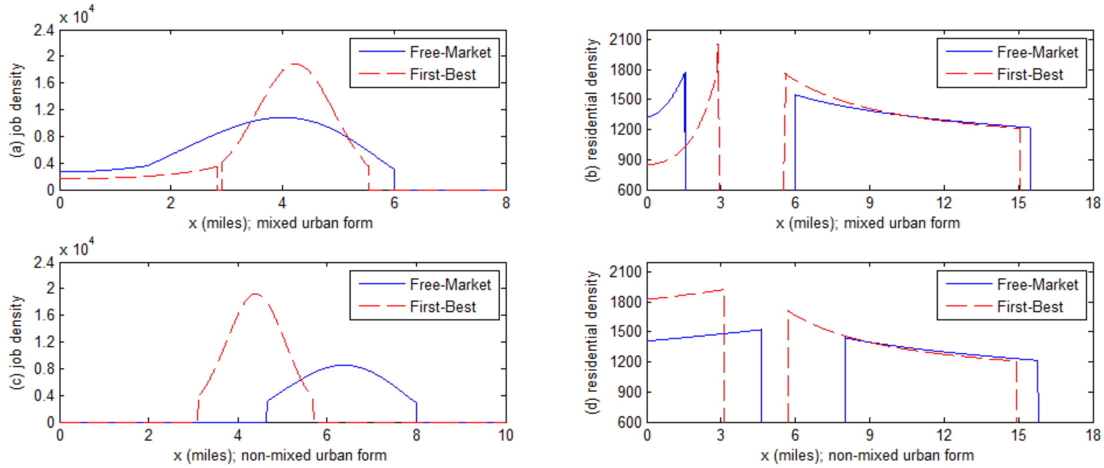


Figure A1 Spatial Distributions of Job and Residential Densities in Mixed and Non-Mixed Equilibria (φ=60, ρ=0.00001, σ=1.5, γ=0.06, N=600,000)

Figure A1 provides a simulation example. Under the same parameter sets, the solutions to the free-market equilibrium and first-best optimum are not unique. The solution can be a mixed urban form (Figure A1a and A1b) or a non-mixed, annular urban form (A1c-A1d). In the free-market cases, the utility level in the annular urban equilibrium (u=5195) is larger than that in the mixed urban equilibrium (u=5185), leading to a CV gain of $72 per household per year. The corresponding first-best policies can obtain a CV of $125 per household per year in the mixed-use optimum and $171 per household per year in the annular urban optimum. These findings suggest that the non-mixed equilibrium allocations are more efficient than the equilibrium mixed-use allocations in both the laissez-faire cities and the cities with optimal policies correcting externalities.

In theory, the question of whether mixed land use patterns is Pareto-optimal is discussed in three situations. The first is a free market where both congestion and agglomeration externalities are not internalized. The second one is that the society recognizes both externalities, but do correct them by introducing policy instruments. The third one is the social optimum, where the externalities are internalized and fully corrected.

In the free-market case, the constraints (A3) and (A4) in Problem A are relaxed. Suppose firms exist at location x, i.e., $\theta_f(x) > 0$, the solutions to Problem A satisfy a condition on $n^*(x)$:

(A20) $\delta\kappa n^*(x)^{\kappa-1}F(x)^{\gamma} - \beta_3(x) = 0,$

and the solutions to the firms' profits maximization problem require the optimal $n^*(x)$ satisfies:

(A21) $\delta\kappa n^*(x)^{\kappa-1}F(x)^{\gamma} = w(x)$

Thus, the optimal $\beta_3^*(x)$ in the free-market equilibrium should equal $w(x)$, i.e.,

(A22) $\beta_3^*(x) = w(x), if \theta_f(x) > 0$

If households exist at location x, i.e., $\theta_h(x) > 0$, from the first-order conditions on $c(x)$ and $q(x)$ of Problem A, one can derive the optimal $c^*(x)$ and $q^*(x)$ satisfy the following condition:

(A23) $\frac{c^*(x)-\beta_3^*(x)}{q^*(x)} = \frac{\partial u}{\partial q} / \frac{\partial u}{\partial c}$

By comparing the condition (A23) and the conditions of utility maximization, i.e., Eqs. (3.7) and (3.8), one can derive:

(A24) $\beta_3^*(x) = y(x) = w(x) + \bar{y}, if \theta_h(x) > 0$

Combining Eqs. (A22) and (A24):

(A25) $\beta_3^*(x) = \begin{cases} w(x) + \bar{y}, if \ \ \theta_h(x) > 0 \\ w(x), \quad\quad if \ \ \theta_f(x) > 0 \end{cases}$

Thus, if $\bar{y} \neq 0$, there exist no mixed land use at any location x. If the governmental incomes including rent and toll revenues net of subsidy expenditures were redistributed back to residents, a mixed urban form would be never Pareto-optimal.

However, if the governmental incomes are considered to be owned by an absent landlord and/or city authority and are not redistributed, a mixed land use pattern could be an optimal solution. This is why a completely or partially mixed urban configuration could be a Pareto-optimal solution to the models of Ogawa and Fujita (1982) and Lucas and Rossi-Hansberg (2002).

Under the second situation, Problem A includes the constraints (A3) and (A4) and sets $\tau(x) = 0$ and $s(x) = 0$. Similar to the free-market case, one can compute the optimal $\beta_3^*(x)$ as follows:

(A26) $\beta_3^*(x) =$

$$\begin{cases} w(x) + \bar{y}, & if \;\; \theta_h(x) > 0 \\ w(x) + \zeta\gamma\delta \int_0^{\bar{x}} \int_0^{2\pi} r\theta_f(r)n(r)^{\kappa}F(r)^{\gamma-1}e^{-\zeta l(x,r,\psi)}\, d\psi dr, & if \;\; \theta_f(x) > 0 \end{cases}$$

Obviously, there are no mixed land use at any location x even the governmental incomes equal zero. Thus, if externalities are realized in the city market but no policy instruments are adopted, the optimal urban configuration has no mixed land use areas. This finding in fact is consistent with the Theorem 1 in Rossi-Hansberg (2004), though his research only internalizes agglomeration externalities.

Under the third situation, Problem A includes the constraints (A3) and (A4) and both $\tau(x)$ and $s(x)$ are set at their optimal levels (equaling their corresponding marginal externalities). The optimal $\beta_3^*(x)$ equals that in the free-market case, as follows:

(A27) $$\beta_3^*(x) = \begin{cases} w(x) + \bar{y}, & if \;\; \theta_h(x) > 0 \\ w(x), & if \;\; \theta_f(x) > 0 \end{cases}$$

Thus, similar to the free-market case, the socially optimal land use patterns would have no mixed areas, if the rent and toll revenues net of subsidy expenditures were partially or totally returned back to residents.

A5: PROOF OF PROPOSITION 4.1

Problem 1's Hamiltonian function is as follows:

$$H_1(n, q, c, \theta_f, \beta_1, \beta_2, \beta_3)$$
$$= u\left(c(x), q(x)\right)/\lambda(x)$$
$$+ 2\pi x\left[\theta_f(x) f\big(n(x)\big) A\big(F(x)\big) - \frac{1-\theta_t-\theta_f(x)}{q(x)} c(x)\right.$$
$$\left. - (1-\theta_t)R_A\right] - t(x)D(x)$$
$$+ \beta_1(x)[c(x) + r_h(x)q(x) - w(x) - \bar{y}]$$
$$+ \beta_2(x)\left[f\big(n(x)\big)A\big(F(x)\big) - w(x)n(x) - r_f(x)\right]$$
$$+ \beta_3(x)2\pi x\left(\frac{1-\theta_t-\theta_f(x)}{q(x)} - \theta_f(x)n(x)\right)$$

From the Maximum Principle (Pucci and Serrin, 2007), the first-order conditions are as follows:

(A28) $\frac{\partial H_1}{\partial n} = 2\pi x\theta_f(x)\left[f_n\big(n(x)\big)A\big(F(x)\big) - \beta_3(x)\right] + \beta_2(x)\left[f_n\big(n'(x)\big)A\big(F(x)\big) - w(x)\right] = 0$

(A29) $\frac{\partial H_1}{\partial c} = \frac{U_c}{\lambda(x)} - 2\pi x\frac{1-\theta_t-\theta_f(x)}{q(x)} + \beta_1(x) = 0$

(A30) $\frac{\partial H_1}{\partial q} = \frac{U_q}{\lambda(x)} + 2\pi x\frac{1-\theta_t-\theta_f(x)}{q^2(x)}c(x) + \beta_1(x)r_h(x) - \beta_3(x)2\pi x\frac{1-\theta_t-\theta_f(x)}{q^2(x)} = 0$

(A31) $\frac{\partial H_1}{\partial \theta_f} = f\big(n(x)\big)g\big(F(x)\big) + \frac{c(x)}{q(x)} - \frac{\beta_3(x)}{q(x)} - \beta_3(x)n(x) = 0$

(A32) $\frac{\partial H_1}{\partial D} = -\beta_3{}'(x)$, *and thus* $\beta_3{}'(x) = t(x)$.

(a) (A28) → $f\big(n'(x)\big)F(x)^\gamma - w(x) = 0$ and $f\big(n'(x)\big)g\big(F(x)\big) - \beta_3(x) = 0$. Then, $\beta_3(x) = w(x)$, and $f\big(n'(x)\big) = w(x)/g\big(F(x)\big)$, so $n^*(x) = n^*(w(x))$.

(b) Given $r_h(x) = \frac{y(x)-c(x)}{q(x)}$, (A29)/(A30) = $c(x) + q(x)u_q/u_c = y(x)$. Thus, given $u\big(c(x), q(x)\big) = \bar{u}$, one can solve for $q^*(x) = q^*(w(x), \bar{u})$ and $c^*(x) = c^*(w(x), \bar{u})$.

(c) (A31) → $\frac{\partial H_1}{\partial \theta_f} = r_f(x) - r_h(x)$. Thus, if $r_f^*(x) > r_h^*(x) \to \frac{\partial H_1}{\partial \theta_f} > 0$, the larger the $\theta_f(x)$, the larger the *H*. Since $0 \le \theta_f(x) \le 1 - \theta_t$, $\theta_f^*(x) = 1 - \theta_t$. Similarly, if

$r_f^*(x) < r_h^*(x)$, then $\theta_f^*(x) = 0$. If $r_f^*(x) = r_h^*(x)$, then $0 < \theta_f^*(x) < 1 - \theta_t$, and both firms and households will locate at location x, which is a mixed use area.

(d) From (a) and (A32) we have $w_3{}'(x) = t(x)$.

A6: PROOF OF PROPOSITION 4.2

The search for a second-best optimal congestion toll (e.g., a VMT tax and a cordon toll) is equivalent to solving Problem 4.1's optimization by adding constraints on $t(x)$ and $w(x)$ and imposing the following condition:

(A33) $$\int_0^{\bar{x}} \tau(x)D(x)\,dx = \epsilon \int_0^{\bar{x}} \tau_{mce}(x)D(x)dx\,, \epsilon \neq 1\,:$$

Eq. (4.10)'s constraint represents the internalized travel cost, while proposition 4.1d guarantees that the wage gradient equals the marginal private travel cost plus a congestion toll, which is s a critical condition for Pareto efficiency. The condition (A33) implies that second-best tolls cannot correct all the aggregate congestion externalities; such tolls cover just an ϵ ($\epsilon < 1$) share of those external costs.

With the condition (A33), the first-order condition of the corresponding Hamiltonian function with respect to $D(x)$ is as follows:

(A34) $$\beta_3'(x) = t(x) + \rho\sigma\left(\frac{|D(x)|}{2\pi x\theta_t}\right)^{\sigma} + \tau(x) - \epsilon(1+\sigma)\rho\sigma\left(\frac{|D(x)|}{2\pi x\theta_t}\right)^{\sigma}$$

Since $\beta_3(x) = w(x)$ still holds here (as noted in Appendix A5), (A34) become the following:

(A35) $$w'(x) = t(x) + \rho\sigma\left(\frac{|D(x)|}{2\pi x\theta_t}\right)^{\sigma} + \tau(x) - \epsilon(1+\sigma)\rho\sigma\left(\frac{|D(x)|}{2\pi x\theta_t}\right)^{\sigma}$$

Comparing (A35) and the Pareto condition on the wage gradient (Eq. 9), one can derive that the optimal toll $\tau^*(x)$ needs to reflect/correct for $\frac{1}{1+\sigma}$ of overall congestion externalities.

Glossary

CP: Congestion Pricing

CV: Compensating Variation

EC: Excessive Congestion

EZ: Exclusionary Zoning

FB: First Best

FCZ: Firm Cluster Zoning

FM: Free Market

HOT: High-Occupancy Toll

LDZ: Low-Density Zoning

LOS: Level of Service

LUP: Land Use Planning

MCP: Marginal Cost Pricing

MML: Multilevel Multinomial Logit

MPB: Marginal Private Benefit

MPC: Marginal Private Cost

MSB: Marginal Social Benefit

MSC: Marginal Social Cost

PCT: Pigouvian Congestion Toll

PLS: Pigovian Labor Subsidy

TFF: Traffic Free Flow

TFP: Total Factor Productivity

UGB: Urban Growth Boundary

VMT: Vehicle-Miles Traveled

Bibliography

Anas, A. 2012. The optimal pricing, finance and supply of urban transportation in general equilibrium: A theoretical exposition. *Economics of Transportation*, 1, 64–76.

Anas, A. 2013. The location effects of alternative road-pricing policies. In Ingram, G. K. & Karin L. B.'s (eds.) *Infrastructure and Land Policies*. Cambridge, MA: Lincoln Institute of Land Policy.

Anas, A., & Hiramatsu, T. 2013. The economics of cordon tolling: General equilibrium and welfare analysis. *Economics of Transportation*, 2(1), 18-37.

Anas, A., & Kim, I. 1996. General equilibrium models of polycentric urban land use with endogenous congestion and job agglomeration. *Journal of Urban Economics*, 40, 232–256.

Anas, A., & Lindsey, R. 2011. Reducing urban road transportation externalities: Congestion pricing in theory and in practice. *Review of Environmental Economics and Policy*, 5(1), 66-88.

Anas, A., & Liu, Y.2007. A regional economy, land use, and transportation model (RELU-RAN): Formulation, algorithm design, and testing, *Journal of Regional Science*, 47 (3), 415–455.

Anas, A. & Rhee, H. 2006. Curbing excess sprawl with congestion tolls and urban boundaries. *Regional Science and Urban Economics*, 36, 510–541.

Anas, A., & Xu, R. 1999. Congestion, land use and job dispersion. *Journal of Urban Economics*, 45, 451–473.

Arnott, R. J. 1979. Unpriced transport congestion. *Journal of Economic Theory*, 21(2), 294-316.

Arnott, R., 2007. Congestion tolling with agglomeration externalities. *Journal of Urban Economics*, 62 (2), 187–203.

Arzaghi, M., & Henderson, J. V. 2008. Networking off madison avenue. *The Review of Economic Studies*, 75(4), 1011-1038.

Badoe, D., & Miller, E. 2000. Transportation –land use interaction: Empirical findings in North America, and their implications for modeling. *Transportation Research D*, 5, 235-263.

Baum-Snow, N. 2007. Did highways cause suburbanization?. *The Quarterly Journal of Economics*, 775-805.

Ben-Akiva, M., & Lerman, S. 1985. *Discrete Choice Analysis*. Cambridge, MA: MIT Press.

Berliant, M., Peng, S., & Wang, P. 2002. Production externalities and urban configuration, *Journal of Economic Theory*, 104, 275–303.

Bhat, C., Carini, J. & Misra, R. 1999. On modeling the generation and organization of household activity stops. *Transportation Research Record*, 1676, 153-162.

Bhatt, K. 1993. Implementing congestion pricing: winners and losers. *ITE Journal,* 63(12), 33-37.

Bhatt, K. 2011.Congestion pricing: An overview of experience and impacts. In, Gregory K. Ingram and Yu-Hung Hong (Eds.), *Climate Change and Land Policies*. Cambridge, MA: Lincoln Institute of Land Policy.

Bilbao-Ubillos, J. 2008. The costs of urban congestion: Estimation of welfare losses arising from congestion on cross-town link roads. *Transportation Research A: Policy and Practice*, 42(8), 1098-1108.

Boarnet, M. G. 1997. Infrastructure services and the productivity of public capital: the case of streets and highways. *National Tax Journal*, 50(1), 39–57.

Boarnet, M. G., & Chalermpong, S. 2001. New highways, house prices, and urban development: A case study of toll roads in Orange County, CA. *Housing policy debate*, 12(3), 575-605.

Boarnet, M. G. & Crane, R. 2001. *Travel by Design: The Influence of Urban Form on Travel*. New York. Oxford University Press.

Bogart, W. T. 1998. *The Economics of Cities and Suburbs*. Upper Saddle River, NJ: Prentice Hall.

Borck, R., & Wrede, M. 2009. Subsidies for intracity and intercity commuting. *Journal of Urban Economics*, 66(1), 25-32.

Boucekkine, R., C. Camacho, & Zou, B. 2009. Bridging the gap between growth theory and the new economic geography: The spatial ramsey model. *Macroeconomic Dynamics*, 13, 20–45.

Brinckerhoff, P. 2009. *Design Guidance for Use in Evaluation of HOT Lane Corridors.* Metropolitan Transportation Commission, 1, 5.

Brueckner, J. K. 2000. Urban sprawl: Diagnosis and remedies. *International Regional Science Review*, 23(2), 160-171.

Brueckner, J. 2007. Urban growth boundaries: An effective second-best remedy for unpriced traffic congestion? *Journal of Housing Economics*,16, 263-273.

Buyukeren A. C., Hiramatsu T., 2015. Anti-congestion policies in cities with public transportation. *Journal of Economic Geography*, 1–27.

Cao, X. J. 2009. Disentangling the influence of neighborhood type and self-selection on driving behavior: An application of sample selection model. *Transportation*, 36(2), 207-222.

Cao, X., Mokhtarian, P. L., & Handy, S. L. 2009. Examining the impacts of residential self-selection on travel behaviour: A focus on empirical findings. *Transport Reviews*, 29(3), 359-395.

Cervero, R. 1996. Jobs-housing balance revisited: trends and impacts in the San Francisco Bay Area. *Journal of the American Planning Association*, 62(4), 492-511.

Cervero, R., & Kockelman, K. 1997. Travel demand and the 3Ds: Density, diversity, and design. *Transportation Research Part D: Transport and Environment*, 2(3), 199-219.

Cervero, R. 1998. The Planned City: Coping With Decentralization: An American Perspective. Available at: http://escholarship.org/uc/item/5b29d0fd

Cervero, R. 2002. Land use and mode choice: Toward a normative framework. *Transportation Research D*, 7, 265-284.

Cheshire, P. & Sheppard, S. 2005.The Introduction of Price Signals into Land Use Planning Decision-Making: A Proposal. London: LSE Research Online. Available at: http://eprints.lse.ac.uk/archive/00000568.

Chi, G. 2012. The impacts of transport accessibility on population change across rural, suburban and urban areas: A case study of Wisconsin at sub-county levels. *Urban Studies*, 49(12), 2711-2731.

Cho, S. H., Poudyal, N., & Lambert, D. M. 2008. Estimating spatially varying effects of urban growth boundaries on land development and land value. *Land Use Policy*, 25 (3), 320-329.

Clawson, M. 1971. *Suburban Land Conversion in the United States*. Baltimore: The Johns Hopkins Press.

Combes, P. P., Duranton, G., Gobillon, L., & Roux, S. 2010. Estimating agglomeration economies with history, geology, and worker effects. In *Agglomeration Economics*. University of Chicago Press.

Cox, W. 2010. Property values 11 times higher across Portland's urban growth boundary. Available at: http://www.newgeography.com/content/001808-property-values-11-times-higher-across-portlands-urban-growth-boundary, 2013, July.

Graham D. J. 2005: *Wider Economic Benefits of Transport Improvements: Link Between Agglomeration and Productivity*. Imperial College, London.

Crane, R. 1999. *The Impacts of Urban Form on Travel: A Critical Review* (Working paper WP99RC1). Boston, MA: Lincoln Institute of Land Policy.

Crane, R., & Chatman, D. G. 2003. As jobs sprawl, whither the commute?. *Access Magazine,* 1(23):1-12.

Datta-Chaudhuri, M. 1990. Market failure and government failure. *The Journal of Economic Perspectives*, 25-39.

De Lara, M., De Palma, A., Kilani, M., & Piperno, S. 2013. Congestion pricing and long term urban form: Application to Paris region. *Regional Science and Urban Economics*, 43(2), 282-295.

Desmet, K. & E. Rossi-Hansberg. 2010, On spatial dynamics. *Journal of Regional Science*, 50, 43-63.

Downs, A. 2004. *Growth Management and Affordable Housing: Do They Conflict*. New York: Brookings Institution Press.

Downs, A. 2005. *Still Stuck in Traffic: Coping with Peak-Hour Traffic Congestion*. New York: Brookings Institution Press.

Duranton, G., & Puga, D. 2014. *Urban Land Use*. CEPR Discussion Papers.

Ellickson, R. C. 1973. Alternatives to zoning: Covenants, nuisance rules, and fines as land use controls. *University of Chicago Law Review*, 40(4), 681-781.

Ewing, R., & R. Cervero. 2001. Travel and the built environment: A synthesis. *Transportation Research Record*, 1780, 87-114.

Ewing, R. 2008. Highway-induced development: Research results for metropolitan areas. *Transportation Research Record: Journal of the Transportation Research Board*, 2067(1), 101-109.

Ewing, R. & Cervero, R. 2010. Travel and the built environment. *Journal of the American Planning Association*, 76 (3), 265-294.

Ewing, R., Greenwald, M., Zhang, M., Walters, J., Feldman, M., Cervero, R., & Thomas, J. 2011. Traffic generated by mixed-use developments: Six-region study using consistent built environmental measures. *Journal of Urban Planning and Development*, 137(3), 248-261.

Falcocchio, J. C., & Levinson, H. S. 2015. *Road Traffic Congestion: A Concise Guide.* Cham: Springer International Publishing.

Fischel, W.A. 1985. *The Economics of Zoning Laws: A Property Rights Approach to American Land Use Controls*. Baltimore: The Johns Hopkins University Press.

Fischel, W. A. 1999. Does the American way of zoning cause the suburbs of metropolitan areas to be too spread out?. *Governance and Opportunity in Metropolitan America*, 15, 1-19.

Frank, L. D., Saelens, B. E., Powell, K. E., & Chapman, J. E., 2007, Stepping towards causation: Do built environments or neighborhood and travel preferences explain physical activity, driving, and obesity? *Social Science & Medicine*, 65, 1898-1914

Federal Highway Administration. 1997. *Freeway Management Handbook*. Washington, DC, Report No. FHWA-SA-97-064. Avaliable at http://www.library.unt.edu/gpo/OTA/pubs/fmh/mastoccd.pdf

Fujishima, S. 2011. The welfare effects of cordon pricing and area pricing: simulation with a multi-regional general equilibrium model. *Journal of Transport Economics and Policy*, 45 (3), 481–504.

Fujita, M. & Ogawa, H. 1982. Multiple equilibria and structural transition of non-monocentric urban configurations. *Regional Science and Urban Economics*, 12: 161–196.

Fujita, M. & Thisse, J. 2002. *Economics of Agglomeration*. Cambridge University Press.

Garreau, J. 1991. *Edge City: Life on the New Frontier*. New York: Doubleday.

Glaeser, E. L., & J. Gyourko. 2003. The impact of building restrictions on housing affordability. *FRBNY Economic Policy Review*, 2, 21–29.

Glaeser, E. L., Gyourko, J., & Saks, R. 2005. Why have housing prices gone up?. National Bureau of Economic Research. Working paper No. w11129.

Goodwin, P. 2004. The economic costs of road traffic congestion. Available at http://discovery.ucl.ac.uk/1259/.

Gordon, P., & Richardson, H. W. 2001. The sprawl debate: Let markets plan. *The Journal of Federalism*, 31(3), 131-149.

Graham, D. J. 2007. Variable returns to agglomeration and the effect of road traffic congestion. *Journal of Urban Economics*, 62(1),103–120.

Grant-Muller, S. M., & Laird, J. J. 2007. *Costs of Congestion: Literature based Review of Methodologies and Analytical Approaches*. Edinburgh: Scottish Executive.

Grilli, L., & Rampichini, C. 2007. A multilevel multinomial logit model for the analysis of graduates' skills. *Statistical Methods and Applications*, 16(3), 381-393.

Guiliano, G. 1994. Equity and fairness considerations of congestion pricing. In: *Curbing Gridlock*, Vol.2. National Research Council, Special report 242, Washington D.C., 250-279.

Giuliano, G., & Small, K. A. 1991. Subcenters in the Los Angeles region. *Regional science and urban economics*, 21(2), 163-182.

Gulipalli, P. K., & Kockelman, K. M. 2008. Credit-based congestion pricing: A Dallas-Fort Worth application. *Transport Policy*, 15(1), 23-32.
Guo, Z., Agrawal, A. W., & Dill, J. 2011. Are land use planning and congestion pricing mutually supportive? Evidence from a pilot mileage fee program in Portland, OR. *Journal of the American Planning Association*, 77(3), 232-250.
Gupta, S., Kalmanje, S., & Kockelman, K. M. 2006. Congestion pricing simulations: Traffic, Simulations and Welfare Impacts for Austin, Texas. *Transportation Planning and Technology,* 29(1), 1–23.
Gyourko, J., Mayer, C., & Sinai, T. 2006. Superstar cities. National Bureau of Economic Research, working paper, No. w12355.
Hamilton, B. W. 1975. Zoning and property taxation in a system of local governments. *Urban Studies*, 12(2), 205-211.
Handy, S. L., Cao, X., & Mokhtarian, P. L. 2006. Self-selection in the relationship between built environment and walking: Evidence from Northern California. *Journal of the American Planning Association*, 72 (1), 55-74.
Harvey, G. W. 1994. Transportation pricing and travel behavior. *Transportation Research Board Special Report*, 242.
Henderson, J.V. 2003. Urbanization, economic geography and growth. In P. Aghion and S. Durlauf's (eds.) *Handbook of Economic Growth* (Vol. 1), North Holland.
Howard, E. 1902. *Garden Cities of Tomorrow*. London: S. Sonnenschein & Co., Ltd.
Hsieh, C. T., & Moretti, E. 2015. Why Do Cities Matter? Local Growth and Aggregate Growth (No. w21154). National Bureau of Economic Research.
Hymel, K. 2009. Does traffic congestion reduce employment growth?. *Journal of Urban Economics*, 65(2), 127–135.
Ieromonachou, P. & S. Potter, et al. 2006. Norway's urban toll rings: Evolving towards congestion charging? *Transport Policy*, 13 (5): 367-378.
Kalmanje, S., & Kockelman, K. M. 2004. Credit-based congestion pricing: Travel, land value, and welfare impacts. *Transportation Research Record: Journal of the Transportation Research Board*, 1864(1), 45-53.
Kilani, M., Leurent, F., & De Palma, A. 2010. Monocentric city with discrete transit stations. *Transportation Research Record: Journal of the Transportation Research Board*, 2144, 36-43.
Knaap, G., Talen, E., Olshansky, R., & Forrest, C. 2000. Government policy and urban sprawl. Illinois Department of Natural Resources, Office of Realty and Environmental Planning.
Kockelman, K. 1997. Travel behavior as function of accessibility, land use mixing, and land use balance: Evidence from San Francisco Bay Area. *Transportation Research Record: Journal of the Transportation Research Board,* 1607, 116-125.
Kono, T., & Joshi, K. K., 2012. A new interpretation on the optimal density regulations: closed and open city. *Journal of Housing Economics*, 21, 223-234.
Kono,T., Joshi, K.K., Kato, T., & Yokoi, T. 2012. Optimal regulation on building size and city boundary: An effective second-best remedy for traffic congestion externality. *Regional Science and Urban Economics*, 42, 619–630.

Lakshmanan, T. R. 2011. The broader economic consequences of transport infrastructure investments. *Journal of Transport Geography*, 19(1), 1-12.

Langer, A., & Winston, C. 2008. Toward a comprehensive assessment of congestion pricing accounting for land use. *Brookings-Wharton Papers on Urban Affairs*, 2008(1), 127-175.

Le Grand, J. 1991. The theory of government failure. B*ritish Journal of Political Science*, 21(4), 423-442.

Lee, B., & Lee, Y. 2013. Complementary pricing and land ue policies: Does it lead to hgher transit use?. *Journal of the American Planning Association*, 79(4), 314-328.

Levine, J. 2006. Zoned out. *Regulation, Markets and Choices in Transportation and Metropolitan Land Use*. Washington: Resources for the Future.

Levine, J., & Inam, A. 2004. The market for transportation-land use integration: Do developers want smarter growth than regulations allow?. *Transportation*, 31(4), 409-427.

Levine, J., Inam, A., & Torng, G. W. 2005. A choice-based rationale for land use and transportation alternatives evidence from Boston and Atlanta. *Journal of Planning Education and Research*, 24(3), 317-330.

Litman, T. 2009. Transportation cost and benefit analysis. Victoria Transport Policy Institute, 31.

Litman, T. 2011. London congestion pricing: Implications for other cities. Victoria Transport Policy Institute. Available at http://www.vtpi.org/london.pdf

Litman, Todd. 2014. Land Use Impacts on Transportation: How Land Use Factors Affect Travel Behavior, Victoria Transport Policy Institute. Available at http://www.vtpi.org/landtravel.pdf.

Lucas, R.E. Jr., & Rossi-Hansberg, E. 2002. On the internal structure of cities. *Econometrica*, 70 (4), 1445–1476.

Mahendra, A., Grant, M., & Swisher, M. 2012. Effective Approaches for Advancing Congestion Pricing in a Metropolitan Region–A Primer on Lessons Learned and Best Practices (No. FHWA-HOP-12-030).

Matas, A., & Raymond, J. L. 2003. Demand elasticity on tolled motorways. *Journal of Transportation and Statistics*, 6(2), 91.

Mayer, C. J., & Somerville, C. T. 2000. Land use regulation and new construction. *Regional Science and Urban Economics*, 30(6), 639-662.

McFadden, D. 1974. Conditional Logit Analysis of Qualitative Choice Behavior. In P. Zarembka (ed.), *Fronteirs in Eocoriometrics*. New York: Academic Press.

Mun, S., Konoshi, K., & Yoshikawa, K. 2003. Optimal cordon pricing, *Journal of Urban Economics*, 54, 21–38.

Mun, S., Konoshi, K., & Yoshikawa, K. 2005. Optimal cordon pricing in a non-monocentric city. *Transportation Research Part A,* 39, 723–736.

Nelson, Arthur C., Raymond J. Burby, Edward Feser, Casey J. Dawkins, et al. 2004, Urban containment and central-city revitalization. *Journal of the American Planning Association*, 70, 411-425.

Odeck, J., & Bråthen, S. 2008. Travel demand elasticities and users attitudes: A case study of Norwegian toll projects. *Transportation Research Part A: Policy and Practice*, 42(1), 77-94.

Oxford English Dictionary. 2015. The definition of “excessive”. Available at http://www.oxforddictionaries.com/us/definition/american_english/excessive.

Organization for Economic Co-Operation and Development (OECD). 2007. Managing Urban Traffic Congestion. Available at http://www.internationaltransportforum.org/jtrc/infrastructure/congestion/CongestionSummary.pdf

Parry, I. W., & Bento, A. 2001. Revenue recycling and the welfare effects of congestion pricing. *The Scandinavian Journal of Economics*, 103 (4), 645-671.

Pendall, R. 1999. Do land-use controls cause sprawl?. *Environment and Planning B: Planning and Design*, 26(4), 555-571.

Pines, D., & Sadka, E. 1985. Zoning, first-best, second-best and third-best criteria for allocating land for roads. *Journal of Urban Economics*, 17, 167-183.

Puga, D. 2010. The magnitude and causes of agglomeration economies. *Journal of Regional Science*, 50(1), 203-219.

Pugh, G., & Fairburn, J. 2008. Evaluating the effects of the M6 Toll Road on industrial land development and employment. *Regional Studies*, 42(7), 977-990.

Quigley, J., & S. Rafael. 2005. Regulation and the high cost of housing in California. *American Economic Review Papers and Proceedings*, 95 (2), 323–328.

Rappaport, J. 2014. Monocentric city redux. Federal Reserve Bank of Kansas City Working Paper, No.14-09.

Raudenbush, S. W., & Bryk, A. S. 2002. *Hierarchical Linear Models: Applications and Data Analysis Methods*. New York, Sage.

Rabe-Hesketh, S., Skrondal, A., & Pickles, A. 2002. Reliable estimation of generalized linear mixed models using adaptive quadrature. *The Stata Journal*, 2(1), 1-21.

Rabe-Hesketh, S., Skrondal, A., & Pickles, A. 2004. Generalized multilevel structural equation modeling. *Psychometrika*, 69(2), 167-190.

Rosenthal, S., & Strange, W. 2004. Evidence on the nature and sources of agglomeration economies. In V. Henderson and J.F. Thisse (eds.) *Handbook of Regional and Urban Economics*, volume 4. Amsterdam: North Holland, 2119–2171.

Rosenthal, S. S., & Strange, W. C. 2008. The attenuation of human capital spillovers. *Journal of Urban Economics*, 64(2), 373-389.

Rossi-Hansberg, E. 2004. Optimal urban land use and zoning. *Review of Economic Dynamics*, 7, 69–106.

Santos, G. 2005. Urban congestion charging: a comparison between London and Singapore. *Transport Reviews,* 25 (5), 511-534.

Santos, G., & Shaffer, B. 2004. Preliminary results of the London congestion charging scheme. *Public Works, Management and Policy,* 9 (2), 164–181.

Schwanen, T., & Mokhtarian, P. L., 2004, The extent and determinants of dissonance between actual and preferred residential neighborhood type. *Environment and Planning B: Planning and Design*, 31(5), 759-784.

Solow, R. M. 1972. Congestion, density and the use of land in transportation. The *Swedish Journal of Economics*, 74, 161-173.

Spivak, J. 2013. Toll Roads: The Route to Redevelopment? Available at http://urbanland.uli.org/infrastructure-transit/toll-roads-the-route-to-redevelopment/

Staley, S., Edgens, J. G., & Mildner, G. C. 1999. *A Line in the Land: Urban-Growth Boundaries, Smart Growth, and Housing Affordability*. Los Angeles: Reason Public Policy Institute.

Sweet, M. 2011. Does traffic congestion slow the economy?. *Journal of Planning Literature*, 26(4), 391–404.

Sweet, M. 2014. Traffic congestion's economic impacts: evidence from US metropolitan regions. *Urban Studies*, 51(10), 2088-2110.

Taylor, B. D. 2002. Rethinking traffic congestion. *Access Magazine*, 1(21), 8-16..

Taylor, G. R. 1915. *Satellite Cities: A Study of Industrial Suburbs*. New York: D. Appleton and company.

Talen, E., & Knaap, G. 2003. Legalizing smart growth: An empirical study of land use regulation in Illinois. *Journal of Planning Education and Research*, 22(4), 345-359.

Tiebout, C. M. 1956. A pure theory of local expenditures. *The Journal of Political Economy*, 416-424.

Tillema, T., van Wee, B., Rouwendal, J., & van Ommeren, J. 2008. Firms: Changes in trip patterns, product prices, locations and in the human resource policy due to congestion pricing. *Pricing in Road Transport: A Multi-Disciplinary Perspective*, Cheltenham, UK, 106-127.

Turner, M., Haughwout, A., & van der Klaauw, W. 2014. Land use regulation and welfare. *Econometrica*, 82(4), 1341-1403.

Texas A&M Transportation Institute (TTI). 2012. Congestion Data for Your City. Available at http://mobility.tamu.edu/ums/congestion-data/.

Ubbels, B., & Verhoef, E. T. 2008. Governmental competition in road charging and capacity choice. *Regional Science and Urban Economics*, 38(2), 174-190.

Urban Land Institute (ULI). 2013. *When the Road Price is Right: Land Use, Tolls, and Congestion Pricing*. Washington, D.C.: Urban Land Institute.

U.S. Department of Transportation (USDOT), 2003. Summary statistics on demographic characteristics and total travel: 1969, 1977, 1983, 1990, and 1995 NPTS, and 2001 NHTS. Available at http://nhts.ornl.gov/2001/html_files/trends_ver6.shtml

U.S. Government Accountability Office (USGAO). 2012. Traffic Congestion: Congestion pricing Can Help Reduce Congestion, but Equity Concerns May Grow. GAO-12-119. Available at http://www.gao.gov/products/GAO-12-119

Vadali, S. 2008. Toll roads and economic development: exploring effects on property values. *The Annals of Regional Science*, 42(3), 591-620.

Van Wee, B. 2009. Self-selection: A key to a better understanding of location choices, travel behaviour and transport externalities?. *Transport Reviews*, 29(3), 279-292.

Verhoef, E.T., 2005. Second-best congestion pricing schemes in the monocentric city. *Journal of Urban Economics*, 58, 367–388.

Verhoef, E.T., & Nijkamp, P. 2004. Spatial externalities and the urban economy. In Capello, R. and P. Nijkamp's (eds.) *Urban Dynamics and Growth: Advances in Urban Economics*. Elsevier, Amsterdam, 87-120.

Victoria Competition and Efficiency Commission (VCEC). 2006. Making the Right Choices: Options for Managing Transport Congestion, Final Report, http://www.vcec.vic.gov.au/Inquiries/Completed-inquiries/Transport-Congestion/Final-Report

Weisbrod, G., Vary, D. & Treyz, G. 2001. *Economic Implications of Congestion*. Washington, DC: National Academy Press.

Wheaton, W.C. 1998. Land use and density in cities with congestion. *Journal of Urban Economics*, 43, 258-272.

Wheaton, W.C. 2004. Commuting, congestion, and employment dispersal in cities with mixed land use. *Journal of Urban Economics*, 55: 417–438.

Winston, C. 2000. Government failure in urban transportation. *Fiscal Studies*, 21(4), 403-425.

Witherspoon, R., Abbett, J. P., & Gladstone, R. M. 1976. Mixed-use developments: new ways of land use (Vol. 71). Urban Land Institute.

Zhang, M. 2004. The role of land use in travel mode choice: evidence from Boston and Hong Kong. *Journal of the American planning association*, 70(3), 344-360.

Zhang, M., Kone, I. A., Tooley, S., & Ramphul, R. CAMPO Transit-Oriented Development Study. Available at http://www.soa.utexas.edu/files/csd/CAMPO2009TODReport20091201AllSubmit.pdf

Zhang, W., & Kockelman, K. M. 2014. Urban sprawl, job decentralization, and congestion: the welfare effects of congestion tolls and urban growth boundaries. Presented in the 93rd Annual Meeting of the Transportation Research Board, Washington D.C.

Zhang, W. & Kockelman, K. M. 2015.Land use and rent dynamics in a city with land use externalities and forward-looking agents. Presented at the Conference of Advancing Metropolitan Modelling for the Analysis of Urban Sustainability Policies, Riverside, CA.