Title

Essays on Urban and Public Economics

Author

Moon, Byunggeor

Publication Date

2018

UNIVERSITY OF CALIFORNIA,

IRVINE

Essays on Urban and Public Economics

DISSERTATION

submitted in partial satisfaction of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

in Economics

by

Byunggeor Moon

Dissertation Committee:

Professor Jan Brueckner, Co-Chair

Professor Linda Cohen, Co-Chair

Professor Michael McBride

2018

DEDICATION

To

My wife, Eunjung Park

My Family

and

My precious daughter, Emily Youngsuh Moon.

I love you and thank you, Eunjung and Emily

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENT

Firstly, I would like to express my sincere gratitude to my co-chair Prof. Jan Brueckner for the continuous support of my Ph.D study and related research, for his patience, motivation, and immense knowledge. His guidance helped me in all the time of research and writing of this thesis. I could not have imagined having a better advisor and mentor for my Ph.D study.

I would like to pay special thankfulness, warmth and appreciation to my co-chair Prof. Linda Cohen who made my research successful and assisted me at every point to cherish my goal. Her encouragement made it possible to achieve the goal.

I am especially indebted to Prof. Michael McBride, who have been supportive of my academic goals and who worked actively to provide me valuable inputs to pursue those goals.

CURRICULUM VITAE

Byunggeor Moon

2008 B.A. in Economics, Seoul National University, Seoul, Korea

2016 M.A. in Economics, University of California, Irvine

2018 Ph.D. in Economics, University of California, Irvine

FIELD OF STUDY

Urban Economics, Public Economics, Economics of Innovation

PUBLICATIONS

Housing investment, default risk, and expectations: Focusing on the chonsei market in Korea. Regional Science and Urban Economics, 71, 80-90.

ABSTRACT OF THE DISSERTATION

Essays on Urban and Public Economics

By

Byunggeor Moon

Doctor of Philosophy in Economics

University of California, Irvine, 2018

Professor Jan Brueckner, Co-Chair

Professor Linda Cohen, Co-Chair

This paper attempts to provide a theoretical and empirical analysis of important issues in urban and public economics. Chapter 1 notes that the Korean chonsei lease contract, in which the tenant provides a lump sum deposit equivalent to a large part of the housing value to the landlord for the contract period, is in effect a mortgage provided by the tenant to finance the landlord's housing investment. I presents a model for deriving the size of the equilibrium chonsei deposit by incorporating the default risk of chonsei for the first time. Chapter 2 explores the stringency of floor area ratio (FAR), one of the land use regulations in New York, and analyzes it by borough of New York. In particular, I overcome issues in previous studies by constructing a data set using publically available data through new methods. The results show that the FAR stringency in Manhattan is the highest in New York City and the stringency is lower as the distance from the center of the city increases.

Chpater 3 analyzes the impact of receiving R&D grant on firms' receiving subsequent investment from venture capital (VC) using internal data of Korea's R&D grant program. We address sample selection and endogeneity issues that arise when estimating the impact of

R&D policy through the matching method and instrument variable approach, respectively, using unique features of Korea's R&D grant programs. The results of our empirical analysis show that firms receiving R&D grants receive 10 to 15% less VC investment than firms that do not. Chapter 4 analyzes the impact of the financial costs of using innovation projects supported by government grants on firm's innovation project choice through game theory. Then, theoretical predictions are verified by using unique data. In particular, I utilize the quasi experimental environment brought about by the institutional characteristics of Korea's R&D grant program and estimate the effect of the cost difference faced by the firm on the type of innovation outcomes (product or process innovation) through the regression discontinuity design.

Chapter 1

Housing Investment, Default Risk, and Expectations: Focusing on the *Chonsei* Market in Korea

1.1 Introduction

There is a unique but widely used rental system called *chonsei* in Korea. Under the *chonsei* scheme, the tenant pays the landlord 40-60% of the value of the house in a lump sum (called *chonsei deposit*), rents the house, and receives the money back at the end of the contract period. In other words, the tenant lends the chonsei deposit to the landlord at the expense of general consumption, and the landlord can use the chonsei to invest in the house. Therefore, chonsei is like a mortgage supplied by the tenant, not a bank, and it has essentially the same structure as a mortgage.

As Navarro and Turnbull (2010) point out, in countries based on civil law, an antichresis contract requiring deposit of a lump sum can be used in a lease contract, and chonsei is a form of such a contract. In Korea, however, chonsei is a type of contract that accounts for over 50% of the housing rental market, and the total amount of chonsei deposits in the economy is estimated to be $420 billion in 2016. Therefore, considering the size and the proportion of chonsei in the housing market, the significance of the system in the Korean housing market is apparent.

This paper goes beyond simply studying the lease contract between the landlord and the tenant. Chonsei is a private mortgage provided by the tenant, and we contribute to existing mortgage-related research by analyzing the chonsei, taking into account the nature of the mortgage and the default risk. Therefore, it is necessary to review the existing studies on both chonsei and the mortgage market in order to approach chonsei as a financing method for housing investment.

Studies on the Korean chonsei market have focused on why the chonsei system has arisen and why chonsei has become the most representative type of lease contract. Renaud (1989) suggests the friendly policy of the Korean government toward the industrial sector as a reason for the chonsei system becoming popular. He argues that the Korean government focused on lending money to the industrial sector at low interest rates for economic growth, and that the real estate market has therefore experienced financial repression. In this situation, chonsei has become an important source of funding for those who want to buy real estate. That is, in a situation where it is difficult to access a mortgage loan, a lump-sum chonsei deposit has become an important catalyst facilitating the purchase of a home. In addition, a high rate of real estate price increases in line with economic growth has ensured a safe return of chonsei deposits, which makes chonsei widely used in the real estate rental market.

Son (1997) also explains why the chonsei system became popular in terms of the demand and supply of rental housing. In Korea, the demand for housing has increased rapidly due to rapid industrialization. On the supply side, however, the government focused on allowing individuals to own their own homes and was reluctant to support adequate rental housing. In this situation, the main supplier of rental housing became the private home buyer, and in a situation where the mortgage market is not relatively developed, home buyers became dependent on the tenant's chonsei.

The reason why the market for chonsei has developed in Korea has been discussed sufficiently, but there has not been much theoretical and equilibrium analysis of the market itself. The discussion of Ambrose and Kim (2003) is most relevant to our argument, since they first considered the default risk of chonsei. Kim (2013) explains through a general equilibrium model that choosing chonsei instead of monthly rent is the mutually optimal choice on the premise that either chonsei or monthly rent can be selected.[1]

This study focuses on the mortgage aspect of chonsei and considers the fact that landlords who receive chonsei may default. On the institutional side, when the landlord defaults, the tenant cannot foreclose as in the case of a mortgage. Therefore, the tenant has the risk of not receiving the deposit back. According to actual real estate auction statistics, among 31,363 apartments in the Seoul metropolitan area auctioned from 2010 to June 2013 showed that 76.2% of tenants did not receive back some or all of the chonsei deposit when the property was sold.

In the past when real estate prices rose rapidly, there was no major concern about landlords potentially defaulting on the chonsei deposit, because it was widely believed that the property value would be much higher than the deposit after the contract ended. However, after the real estate crisis, real estate prices experienced a sharp decline, and the continued decline of property values put tenants at risk of failing to receive their deposit after the contract period. Therefore, this study reflects the characteristics of chonsei and the possibility that chonsei can involve default like mortgage. To this end, it is necessary to examine the collection of studies on the default behavior of mortgages.

First, for default behavior, Campbell and Dietrich (1983) and Vandell and Thibodeau (1985) support the argument that default occurs when equity is negative, referred to as ruthless default. On the other hand, studies including Lekkas et al. (1993), Hendershott and Schultz (1993), Quigley and Van Order (1995), and Capozza et al. (1997) have shown that the default does not usually exhibit ruthless behavior. In order to explain this empirical behavior, Quigley and Van Order (1995) introduced the concept of default cost, which captures credit and financial-transaction impairment when default occurs. Brueckner (2000) derived an optimal contract that takes into account default cost under asymmetric information.

[1] For additional studies on chonsei, see Cho (2010), Gyourko and Han (1989), Kim (1990), and Lee and Chung (2010)

Brueckner, Calem, and Nakamura (2012) use the default model with default cost, assuming that house price appreciation leads to favorable expectations for future house prices, and showing that this expectation leads to relaxation of underwriting standards. They also provide an empirical test of this proposition. Brueckner, Calem, and Nakamura (2016) use the same approach to study the relationship between house price expectations and the use of alternative mortgage products (AMPs). These studies suggest that lenders and borrowers are more likely to anticipate that defaults will not occur when expectations for future home prices become more favorable, thereby increasing the use of risky mortgage products or relaxing underwriting standards.

For the relationship between default cost and mortgage loan-to-value, the most relevant research is Harrison, Noordewier, and Yavas (2004). They study mortgage data and FICO scores, finding that borrowers with low default cost select high LTV ratios.

This paper utilizes existing theoretical models, but derives an optimal contract between the landlord and the tenant through a new approach, and it also overcomes some limitations of the empirical methods found in existing discussions. Kim (2013) derives an equilibrium contract based on a process in which the landlord provides the tenant with housing consumption, but this paper derives an optimal contract between the tenant and the landlord with a focus on the mortgage nature of chonsei. This paper also extends the scope of the existing discussion by applying Brueckner's (2000) model of the mortgage market with default costs to individual chonsei lease contracts.

In addition, Brueckner, Calem, and Nakamura (2012) use appreciation of past house prices as an indicator of favorable expectations for future house prices, but this approach may not accurately reflect true expectations. This paper complements and develops the existing empirical analysis using the landlords' actual expectation of future house prices derived directly from a survey.

Through these new approaches, this paper derives the equilibrium size of the chonsei deposit in a model that considers the mortgage characteristics of chonsei along with default risk. We also examine the effect of expectations of future housing prices on the size of chonsei deposits, and then test these predictions empirically using real data. The theoretical model to be described in the next section shows the relationship between chonsei deposit, default cost, and expectations for future housing prices.

Theoretically, we find that the larger is the default cost, the smaller is the chonsei deposit and, conversely, the more favorable are the expectations of future house prices, the greater is the chonsei deposit. Empirically, the propositions of the theoretical model are verified by using bank data and survey data. Some limitations of existing studies are overcome using the characteristics of the survey data.

The discussion proceeds as follows. Section 2 provides an overview of the chonsei system in Korea. Section 3 describes the theoretical model showing an optimal chonsei deposit considering chonsei default, Section 4 presents the characteristics of the data for empirical work, and verifies the propositions of the theoretical model through the estimation results.

1.2 Overview of the Korean Chonsei

This section provides an overview of the chonsei system in Korea. The most common rental system in Korea, chonsei does not involve a monthly rent in most cases. Instead, the tenant provides the landlord with a lump sum of 40% to 60% of the property value for the contract period. The tenant then enjoys the housing consumption during the contract period, and the landlord invests the chonsei in the purchase of housing. After the contract period ends, the landlord returns the deposit to the tenant without interest.

It is unclear when the chonsei system started, but it has become the most popular housing lease system during Korea's period of rapid industrialization. Chonsei is by no means the only existing form of lease contract in the country,[2] but in 2016, chonsei was the most common type of contract, accounting for more than 50% of lease contracts.

[2]For example, if the landlord makes a contract with a monthly rent and chonsei, the contract is mixed chonsei. Pure chonsei and pure monthly rent are a form of mixed chonsei.

As Son (1997) and Kim (2013) explain, the rapid urbanization of Korea and the government's financial repression have made chonsei an attractive system for both the landlord and the tenant. Because the government's banking policy favored the industrial sector, landlords could not access enough funds to invest in housing. Chonsei came to fill this gap for the landlord. At the same time, tenants were able to save money by renting through chonsei, and they were able to use the money returned at the end of the contract period as funds to purchase new homes.[3] Rapid urbanization increased demand for housing, but supply was limited by the government's land use controls. As a result, investment in housing became a superior choice over other investments. The combination of the above-mentioned factors led to an increase in housing investment, which the landlord financed through the chonsei provided by the tenant.[4] In return for financing the landlord with the interest-free chonsei deposit, the tenant gained housing consumption and savings opportunities.[5]

1.3 The Model

This section derives the effect of default cost and expectations of future house prices on the equilibrium of lease contracts with a combination of chonsei deposit and monthly rent, an arrangement called "mixed" chonsei. We use the mortgage-market equilibrium model with default costs proposed by Brueckner (2000), and incorporate chonsei and monthly rent selection following Kim (2013).

This model adds novel aspects to existing discussion in several respects. First, we incor-

[3] In a mortgage scheme, the bank and the landlord enter into a contract, and the landlord pays interest and a part of the principal to the bank for a certain period and repays the principal upon termination of the contract. In comparison, chonsei differs in that the tenant and the landlord enter into a contract and provide housing consumption to the tenant for a period of time, the tenant earns "interest" in the form of reduced rent, and the landlord returns the principal to the tenant upon termination of the contract.

[4] Chonsei is also more advantageous to the landlord than a loan because it is free of interest. It also enables large-scale funding compared to monthly rent. Chonsei is also beneficial for the landlord in that, unlike a mortgage, chonsei does not affect an individual's credit rating because it takes the form of a lease, not a loan contract.

[5] Kim (2013) explains that a chonsei tenant can save more compared to a landlord or an owner-occupier. The data used by Lee and Chung (2010) also shows that the deposit-to-monthly-rent conversion rate is from 11.76% to 14.4% a year between 2002 and 2008, compared to a bank interest rate of 4.05% to 5.46%. These discussions show that tenants can save using chonsei rather than using monthly rental contracts.

porate default risk for the chonsei deposit. As explained by Ambrose and Kim (2003), the tenant is not authorized foreclose due to the nature of the chonsei system, so there is a risk of losses from chonsei default. In the model, chonsei default risk is introduced for the first time to make the existing model more sophisticated and fill the existing theoretical gap.

Next, the influence of changes in parameters such as default costs and future house-price expectations on the chonsei deposit, and the conditions under which monthly rent may coexist with chonsei, are derived through the model. While showing the change in the chonsei deposit as parameters change, the model also articulates the conditions under which pure chonsei contracts (rental contract using only chonsei) and mixed chonsei contracts (rental contract using both chonsei deposit and monthly rent) exist.

In addition, we theoretically analyze the interaction of the landlord and tenant in the chonsei system, providing a further example of how housing finance operates under default risk. Chonsei is a system of housing financing between individuals rather than between financial institutions (bank) and borrowers, and our discussion extends the existing literature by analyzing the behavior of housing finance under such a relationship.

1.4 Model Setup

This section derives an optimal lease contract with the tenant incorporating default cost. The landlord and tenant agree to a lease contract using monthly rent (R) and a chonsei deposit (D). Reflecting the characteristics of the chonsei contract, the lease agreement is made over two periods, and the chonsei deposit paid at period-0 is assumed to be returned to the tenant at period-1. However, if chonsei default occurs, the tenant is not authorized to foreclose on the house. Tenant may, however, file a claim with the landlord through a lawsuit. In such cases, the tenant will be able to get back only a portion of the house value that is less than the deposit.

At period-0, the landlord pays P_0 , buys the house, and leases it to the tenant. The reason for chonsei default is that the price of the house is stochastic. We assume that the value of the house in period-1 has a distribution with density of $f(\cdot)$. In the following discussion, it is assumed that housing prices have a uniform distribution to facilitate intuitive explanations. However, assuming a general distribution, many of the conclusions of the discussion remain unchanged.

The default cost (C) of the landlord is the cost incurred when chonsei default occurs at the end of the contract term, including credit restrictions and other limits on financial activities.[6] If the landlord defaults, the tenant cannot foreclose on the house, but we assume that the tenant receives a fraction α of the house value, where $\alpha P < D$.[7] Incorporating default cost, default is optimal for the landlord when

$$P - D \leq (1 - \alpha)P - C \tag{1.1}$$

The left side of (1) is housing equity in period-1, which is the residual value of period-1 housing net of the deposit if no default occurrs. Thus, if the value of $P - D$ is sufficiently small, so that its absolute value is less than the residual value of period-1 housing net of the repayment for the tenant (αP) and the default cost, it is optimal for the landlord to default. Rewriting (1) as $P \leq \frac{D-C}{\alpha}$, $\frac{D-C}{\alpha}$ can be called the default price, and the larger is $\frac{D-C}{\alpha}$, the more likely that default will occur.[8]

Under the assumption that D is given, default's occurrence is associated with C, and

[6]Therefore, it can be assumed that the landlord's credit rating is related to default cost. If a default occurs, this is undesirable for the landlord, as it will result in a decline in the landlord's credit rating, which will negatively affect the landlord's reputation. The higher is the credit rating, the greater is the cost of default, which ruins the credit rating.

[7]Although D is endogenous, all possible equilibrium values are assumed to satisfy this relationship (otherwise default is ruled out). Although it is somewhat unusual to assume a condition involving values of an endogenous variable, doing so seems appropriate in the current model.

[8]In this paper, we assume that α has a fixed arbitrary value which is less than one. In the chonsei scheme, the tenant is unable to foreclose when chonsei default occurs, and the house is auctioned through a court case. Therefore, the final amount the tenant gets paid back will be a certain percentage of the market price of the house. Government policy on housing may affect α. For example, in 2013, the government introduced chonsei deposit insurance. If the tenant uses this insurance and pays the insurance premium, the government pays all or part of the chonsei deposit. Therefore, the introduction of this policy also provides an environment in which the tenant can receive the chonsei deposit even if default occurs.

the smaller is C, the more likely chonsei default is to occur, and therefore the more risky is the landlord. If one landlord is more risky than other landlords, that landlord will have a relatively small C value. The observability of C can vary depending on whether it is private information, but because, when the landlord and tenant make a contract, they share official information (such as the debt of the landlord) and informal information (such as the occupation, residential area of the landlord) through the real estate agency, C is assumed to be information that the tenant can observe.

The cumulative distribution function of future house prices considering house price expectation can be defined as follows:

$$F(P,\delta) = \int_0^P f(p,\delta)dp \tag{1.2}$$

where $f(p,\delta)$ is the density function of the house price and δ is a parameter that shifts the density to the right. The support of the density is $(\overline{P}, \underline{P})$. As a parameter indicating favorable expectations for future house prices, the larger is the value of δ , the higher P is expected to be, with $F_P > 0$.

The landlord purchases the house at period-0 and rents it to tenant using the chonsei deposit (D) and monthly rent (R).[9] The landlord's wealth changes according to the house price of the next period, and the objective function is

$$\pi = -P_0 + D + R + \theta(R + \int_{\underline{P}}^{\frac{D-C}{\alpha}} ((1-\alpha)P - C)f(P,\delta)dP + \int_{\frac{D-C}{\alpha}}^{\overline{P}} (P-D)f(P,\delta)dP) \tag{1.3}$$

Equation (3) describes the present value of landlord's profit. The landlord buys the house by paying P_0 and then rents to tenant receiving D in period-0 and R in both periods.[10] If

[9] Since the landlord uses all of the received chonsei to buy a house, we do not consider interest generated with the chonsei deposit. Also, tenants do not save since their deposit is returned in the future, which means that interest rates also do not enter their object function, considered below.

[10] The tenant in reality would pay a monthly rent periodically between periods zero and one. However, since our model consists of two periods, the R can be viewed as a discounted value of rent over multiple short periods.

the value of the house in the next period is less than $\frac{D-C}{\alpha}$, chonsei default occurs. However, since the tenant will not be able to foreclose on the house, the tenant will get back αP, not P, and the landlord will receive $(1-\alpha)P$ net of default cost (C) when the house is sold at the end of period-1. If the house value is greater than $\frac{D-C}{\alpha}$, the landlord pays back the chonsei deposit and has wealth equal to $P-D$. $\theta < 1$ is the discount factor of the landlord.

The landlord considers monthly rent (R) and chonsei (D) together when renting out a house, and if there is no monthly rent, the contract is pure chonsei. On the other hand, if the landlord makes a contract with a mixture of monthly rent and chonsei, the contract is mixed chonsei.

The landlord's objective function involves the following assumptions. First, it is assumed that the landlord is risk neutral and the purchase price of the house (P_0) at 0-period is fixed. Thus, P_0 is not a decision variable in the landlord's objective function. It is also assumed that the landlord has only one default opportunity. Thus, the lease with the tenant is a one-period contract and does not depend on changes in house prices in multiple periods in the future.

The tenant's utility function can also be derived in conjunction with the landlord's profit function. Based on the assumptions of the landlord's objective function, the tenant assumes that a part of period-1 housing value (αP) is returned if chonsei default occurs. The objective function of the tenant is then

$$U = -D - R + \eta(-R + \int_{\underline{P}}^{\frac{D-C}{\alpha}} \alpha P f(P,\delta) dP + \int_{\frac{D-C}{\alpha}}^{\overline{P}} D f(P,\delta) dP) \tag{1.4}$$

where $\eta < 1$ is the tenant's discount factor. In other words, tenant assumes that if the house value is less than $\frac{D-C}{\alpha}$, only αP is returned, with D returned only if the house value is greater than $\frac{D-C}{\alpha}$. Thus, the utility of tenant is determined by the expected return of the chonsei deposit considering the landlord's default.

1.5 Zero-profit and indifference curves of the landlord and tenant

The analysis discussed in this section develops the characteristics of the landlord's zero-profit curve and tenant's indifference curves from the objective functions described in (3) and (4). Each curve shows the relationship between D and R under the contract for given profit and utility, and equilibrium is a point of tangency between these two curves.

If we set the profit of the landlord to zero, we can obtain the zero profit curve. To do so, the derivatives of π with respect to D and R can be obtained by using Leibniz's rule:

$$\pi_D = 1 + \theta(-(1 - F(\tfrac{D-C}{\alpha}, \delta))) \tag{1.5}$$

$$\pi_R = 1 + \theta \tag{1.6}$$

Using (5) and (6), the slope of the zero profit curve of the landlord in (D, R) space[11] is obtained as

$$\frac{\partial R}{\partial D}\,|_L = -\frac{\pi_D}{\pi_R} = MRS_L = -\frac{1 + \theta(-(1 - F((\frac{D-C}{\alpha}, \delta)))}{1 + \theta} \tag{1.7}$$

To demonstrate the equilibrium through a concrete case, we assume that P follows a uniform distribution with support of $[\underline{P} + \delta, \overline{P} + \delta]$. In this case, the density of P is $1/(\overline{P} - \underline{P})$ and equation (6) can be expressed as

$$\frac{\partial R}{\partial D}\,|_L = -\frac{1 + \theta((\frac{1}{\alpha}(D - C) - (\overline{P} + \delta))/(\overline{P} - \underline{P}))}{1 + \theta} \tag{1.8}$$

In order to derive the characteristics of the tenant's indifference curve, we differentiate

[11] In this model, the purchase price of P_0 is fixed and thus is not a decision variable of the landlord. Thus, the landlord's zero profit curve is fixed with respect to P_0, with zero profit maintained by variation in D and R.

the utility function with respect to D and R using Leibniz's rule:

$$U_D = -1 - \eta(\frac{1}{\alpha}Cf(\tfrac{D-C}{\alpha}, \delta) - (1 - F(\tfrac{D-C}{\alpha}, \delta)) \tag{1.9}$$

$$U_R = -(1 + \eta) \tag{1.10}$$

The slope of the indifference curve is obtained as follows:

$$\frac{\partial R}{\partial D} |_T = -\frac{U_D}{U_R} = MRS_T = -\frac{1 + \eta(\frac{1}{\alpha}Cf(\frac{D-C}{\alpha}, \delta) - (1 - F(\frac{D-C}{\alpha}, \delta)))}{1 + \eta} \tag{1.11}$$

With a uniform house-value distribution, (11) can be written as

$$\frac{\partial R}{\partial D} |_T = -\frac{1 + \eta((\frac{1}{\alpha}D - (\overline{P} + \delta))/(\overline{P} - \underline{P}))}{1 + \eta} \tag{1.12}$$

In the situation where D is given, a bigger R increases the profit of the landlord, but it negatively affects tenant because the tenant has to use more resources for the contract. Thus, the utility of the tenant in the D-R plane is higher as the curve is located lower, while the profit of the landlord is higher as the curve is located higher.

In order to derive the equilibrium, the curvature of the zero profit curve and the indifference curve should be confirmed. The curvature of each curve can be derived by differentiating the MRS of the landlord and the tenant with respect to D, using (8) and (12):

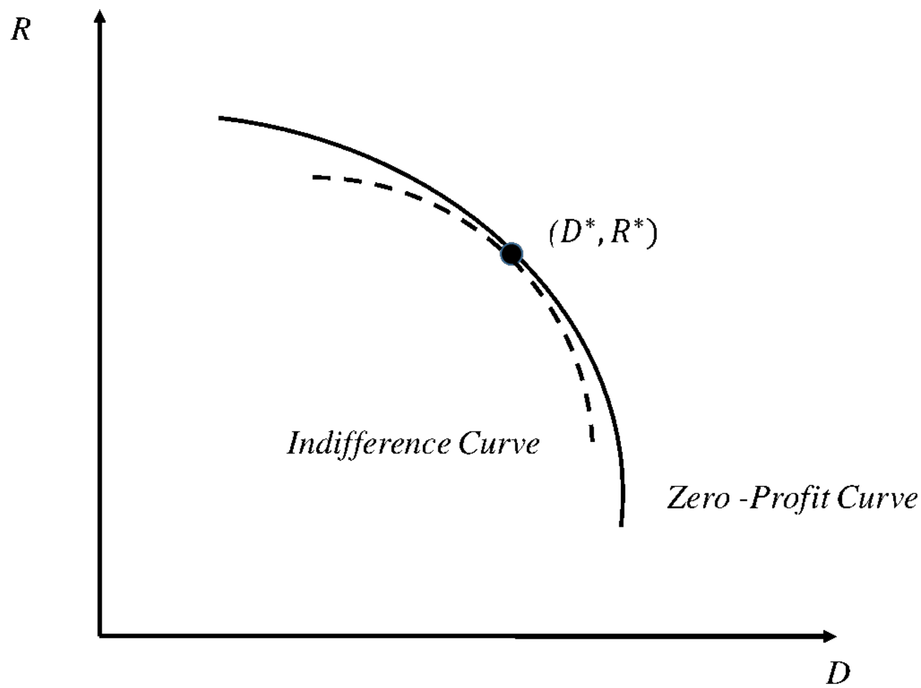
$$\frac{\partial MRS_L}{\partial D} = -\frac{1}{\alpha}\frac{\theta}{1 + \theta} < 0 \tag{1.13}$$

$$\frac{\partial MRS_T}{\partial D} = -\frac{1}{\alpha}\frac{\eta}{1 + \eta} < 0 \tag{1.14}$$

Since (13) and (14) have negative values, the zero-profit and indifference curves are concave.

With this background, the optimal rental contract between the landlord and tenant can be characterized. The equilibrium will be at a point where the lowest indifference curve touches the zero profit curve. For this point to be a tangency, the indifference curve must be more concave than the zero profit curve, as seen in Figure 1. Therefore, given (13) and (14), $\eta > \theta$ must hold in order to satisfy the relative concavity condition for a mixed equilibrium.[12] This conclusion follows because (13) and (14) are increasing in the discount factors so that $\frac{\partial MRS_L}{\partial D} > \frac{\partial MRS_T}{\partial D}$ requires $\eta > \theta$.[13]

Figure 1.1: Equilibrium Contract



1.6 Equilibrium analysis

In this section, the equilibrium contract is derived.To find the equilibrium, we set the tenant's indifference curve slope equal to the landlord's zero profit curve slope, using (8) and (12):

$$\frac{\partial R}{\partial D} \mid_L = MRS_L = -\frac{1 + \theta((\frac{1}{\alpha}(D - C) - (\overline{P} + \delta))/(\overline{P} - \underline{P}))}{1 + \theta} \tag{1.15}$$

[12]$\eta > \theta$ must hold for the mixed equilibrium to be admissible. Otherwise the equilibrium will always be a corner solution.

[13]The discussion focuses on the case where the slopes of the landlord's zero-profit curve and tenant's indifference curve are negative. However, whether each slope has a negative value or a positive value does not make any difference in the analysis. As can be seen from equations (13) and (14), the curvature of each curve is determined independently of the slope, and even if the slope is positive, the indifference curve must be relatively more concave than the zero-profit curve for a tangency equilibrium.

$$= \frac{\partial R}{\partial D} \mid_T = MRS_T = -\frac{1 + \eta((\frac{1}{\alpha}D - (\overline{P} + \delta))/(\overline{P} - \underline{P}))}{1 + \eta}$$

Then, (15) is solved for D to derive

$$D^* = 2\alpha(\overline{P} + \delta) - \alpha(\underline{P} + \delta) - \frac{\theta(1 + \eta)}{(\eta - \theta)}C \tag{1.16}$$

The optimal D^* depends on the default cost C and double the future maximum house value $(\overline{P} + \delta)$ minus future minimum house value $(\underline{P} + \delta)$.[14] Differentiating D^* with respect to C and δ yields

$$\frac{\partial D^*}{\partial C} = -\frac{\theta(1 + \eta)}{(\eta - \theta)} < 0 \tag{1.17}$$

$$\frac{\partial D^*}{\partial \delta} = \alpha > 0 \tag{1.18}$$

As can be seen, a favorable shift in the future house price distribution (an increase in δ) raises D^*. This conclusion makes sense because the higher δ makes it less likely that the deposit will be lost via default. The relative concavity condition shows that η is greater than θ, so the denominator of the factor multiplying default cost in (17) is positive. As a result, the optimal D^* has a negative relation to default cost.[15]

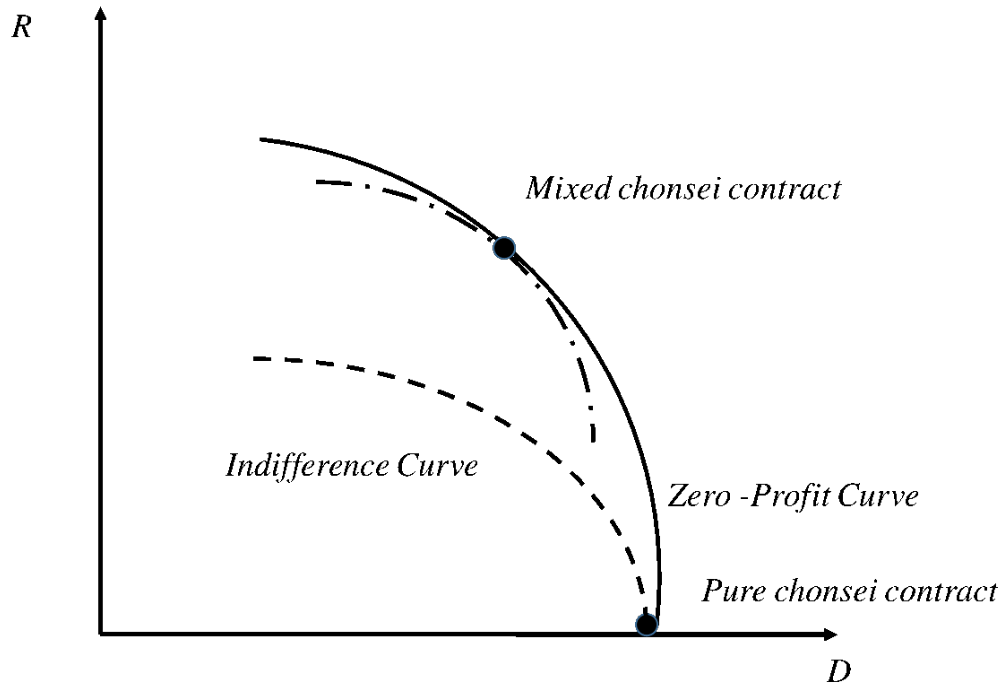
This relationship can be easily seen through the changing position of the zero profit curve as default cost changes. To determine the change of the curve, the MRS expression of the landlord is differentiated with respect to C:

$$\frac{\partial}{\partial C}\frac{\partial R}{\partial D} \mid_L = \frac{1}{\alpha}\frac{\theta}{1 + \theta} \tag{1.19}$$

[14] The optimal D^* is also dependent on α, and D must be greater than αP. If $D > \alpha(\overline{P} + \delta)$, then $D > \alpha P$ is always satisfied. When we apply D^* to this condition, then $D^* > \alpha(\overline{P} + \delta)$ if $\alpha > \frac{\frac{\theta(1+\eta)}{(\eta-\theta)}C}{\overline{P} - \underline{P}}$. This tells us that α cannot have too small a value.

[15] We also do comparative statics for the likelihood of default. We confirm that the default probability increases with an increase in δ, but the change of the default probability with an increase in C is ambiguous, which is determined by the relative size of η and θ. See Appendix 2 for a detailed discussion.

Figure 1.2: Change of equilibrium with change of C



Equation (19) shows that zero profit curve becomes flatter (with a less negative slope) as default cost increases. Intuitively, as C rise, the landlord requires less compensation in the form of higher R when D falls, a consequence of a lower likelihood of default. Since C has no effect on the slope of the indifference curves and the landlord's zero profit curve becomes flatter, the equilibrium D will decrease, as seen in Figure 2.[16] Summarizing the preceding analyses yields

Proposition 1 : If the relative concavity condition and maintained assumptions are satisfied, an increase in the landlord's default cost (a higher credit rating) has the effect of reducing the chonsei deposit D, while a favorable change in the expected density of future house prices (a higher δ) increases the chonsei deposit.

Based on the solution for D^*, the equilibrium of monthly rent (R^*) can be derived by sub-

[16]Differentiating MRS_L and MRS_T against δ reveals that the landlord's zero profit curve and the indifference curve of the tenant are both flattened, but the indifference curve of the tenant becomes even more flat ($\frac{\eta}{1+\eta} > \frac{\theta}{1+\theta}$).

stituting D^* in the landlord's zero profit condition (π=0 in (3)).[17] We can say that contract moves away from pure chonsei as C increases and toward pure chonsei as δ increases.[18]

1.7 Empirical approach

To empirically verify the model's predictions, we need variables that reflect the landlord's credit information (C), lease information (P, D), and expectations for future house prices (δ). However, since the information on individual leases is not collected and accurate data on credit information cannot be constructed in relation to the lease contract, the literature so far has not studied the relationship between credit information and the chonsei deposit used as a means of housing investment.

In this section, we analyze this relationship using an extensive and unique bank data set to overcome these limitations. As shown in the model, default cost affects the chonsei deposit, and we can reasonably assume that the higher the credit rating of the landlord, the higher the default cost. The data includes credit information such as the credit score, letter grade of credit using the score, income, occupation, etc. of the landlord who applied for a mortgage. The data also include house price, the location of the house, and how much chonsei deposit was received from the tenant.

In addition, by using the method of Brueckner, Calem, and Nakamura (2012), which sets the prior house-price appreciation in the area to be the expectation index for future house price increases, we match the location information of the data with the local house price index and analyze the effect of house price appreciation on the chonsei deposit.

The methodology used by Brueckner, Calem, and Nakamura (2012), however, has limitations because rise in house prices in the region may not accurately reflect the expectation of future price increases, as was pointed out in their study. In order to overcome the limitations of this indirect method, we also examine the relationship between expectations and the

[17] See Appendix 3 for a detailed discussion.

[18] Appendix 1 shows mixed chonsei and pure chonsei graphically.

chonsei deposit using survey data on housing finance conducted every year by the Korean government. Because these data are based on questionnaires, there is no information about the individual's credit rating. However, since the expectations of future house prices are directly measured, the predictions of the model can be verified more directly than in existing studies.

In summary, we verify the propositions of the theoretical model for the relationship between the chonsei deposit, default cost, and price expectations using two sets of data. In particular, we use survey data to supplement the existing studies by directly exploring the relationship between expectations and the chonsei deposit.

1.8 Data sources

The first data set consists of data on mortgage applicants from one of the largest banks in terms of asset size, from 2011 to 2016, and focuses on observations where the landlord has previously purchased and leased the house to a tenant using chonsei without a mortgage. This prior information, which pertains to an earlier situation where the landlord did not have a mortgage, is used for purposes of testing the model's predictions.[19] In these data, the borrower's credit score is assessed when a new mortgage is created in a particular year, using a similar method to the way the FICO score is set (the sample score range is from 0 to 2,000). Based on this credit score, letter grading from D to AAA is assigned. Income and age of the landlord are also shown in the data set.

Using information about housing location in these data, we link the house price index of the district[20] unit that the Korean government creates based on actual selling prices. Brueck-

[19]The information about the landlord when earlier buying a house using chonsei may not be accurate at the time of his credit rating evaluation, but due to the limitations of the data, accuracy cannot be confirmed. However, the advantage of the dataset lies in the homogeneity of the landlords in the sample, all of whom originally purchased their house using only chonsei without a mortgage. This provides a good environment for empirically confirming the changes in D as parameters change.

[20]District is a higher administrative unit than town, the smallest administrative unit. Real estate regulations are applied differently across districts, and the house price index is also generated at the district level.

Table 1.1: Summary statistics of the bank data

VARIABLES	Observations	Mean	Std. Dev	Description
CHONSEI	13,588	14,874	12,372	The amount of chonsei deposit
VALUE	13,588	50,866	41,974	Market price of leased house
CREDITSCORE	13,588	860.8	135.3	Credit score (0-2000 points) assessed in a manner similar to FICO
CREDITGRADE	13,588	11.59	1.723	Credit rating based on credit score (D to AAA), a number from 0 to 14 assigned to the 14 letter grade
SIZE	13,588	136.4	463.9	Size of house (in square meters)
SEX	13,588	0.603		1-Male, 0-Female
AGE	13,588	54.27	12.24	The age of householder
INCOME	13,588	354.32	760.95	Monthly income of householder
JOBDUMMY	13,588	0.569		1- have a job, 0-No occupation at the time of application
JOBTYPE	13,588	0.147		1-self employed, 0-otherwise

The sample is comprised of 6 years, which span from 2011 to 2016
Variables with missing Std. Dev. are dummy variables.
Monetary units are in 10,000 Korean won, which is approximately equivalent to 10 US dollars

ner, Calem, and Nakamura (2012) argued that high housing price appreciation generates a price-expectations shift like that portrayed in the theoretical model. Using this methodology, we examine the effect of favorable price expectations on the chonsei deposit. This data set includes 13,588 observations of applicants who previously invested in housing using only chonsei without a mortgage. Table 1 reports summary statistics for the key variables.

The second data set is based on the Korean government's annual housing finance survey. The database consists of 5,000 households nationwide sampled in the same way as the census method each year, and it has data from 2013 to 2016. Some survey respondents are landlords and some are tenants. If the respondent is a tenant, some information, such as the housing prices or the expectation of changes in housing prices, are lacking. Therefore we only use the subsample of respondents who are landlords, for whom such information is available.

The data are based on 1,863 observations of houses rented to tenants and consist of information on the amount of the chonsei deposit, monthly rent, and the expectations of the future house price (the variable equals 1 if future house prices are expected to rise and equals zero if they are expected to fall or stay the same). The data also contain information about the contract type (the variable equals 1 if the contract type is pure chonsei and equals zero if it is mixed chonsei). In addition, variables such as the location of the house and income are also surveyed. This data set helps to overcome the limitations of using past appreciation as a proxy for house price expectations. The main features of these data are shown in Table 2.

Table 1.2: Summary statistics of the survey data

VARIABLES	Observations	Mean	Std. Dev	Description
CHONSEI	1,863	12,313	11,124	The amount of chonsei deposit
CONTRACTTYPE	1,863	0.684		Lease contract type (1 = pure chonsei, 0 = mixed chonsei)
EXPECTATION	1,863	0.366		Answers to expectations that future housing prices will rise (1 = will rise, 0 = will fall or similar)
VALUE	1,863	26,119	30,893	Market price of leased house
SEX	1,863	0.846		1-Male, 0-Female
AGE	1,863	46.34	8.193	The age of householder
INCOME	1,863	569.50	277.69	Monthly income of household

The sample is comprised of 4 years, which span from 2013 to 2016
Variables with missing Std. Dev. are dummy variables.
Monetary units are in 10,000 Korean won, which is approximately equivalent to 10 US dollars

The two data sets differ in some respects. First, the price of the house and the age of the landlord in the bank data are larger than those in the survey data, but the income of the landlord in the bank data is lower. This discrepancy emerges because the two data sets are based on different populations. The bank data is based on people applying for a mortgage, most likely because they bought an expensive home relative to their income. On the other hand, in the survey data, the respondents are randomly selected from the population. Our analysis, however, focuses on the impact of parameters such as credit score[21] or expectation of future house prices on D, and a consistent result using two different data sets can show robustness of what the theory predicts.

1.9 Empirical Strategy

The main empirical model using the bank data relates the chonsei deposit for the household i in district j in year t, denoted as $CHONSEI_{ijt}$, to the one-year lag of annual house-price appreciation, denoted as $HPICHG_{ijt-1}$ (where HPI is the house price index), and the credit score of the landlord, denoted by $CREDITSCORE_{ijt}$, indicating default cost. The $CREDITGRADE_{ijt}$ variable was created by assigning a number from 0 to 14 to the 14 letter grades (D to AAA) based on $CREDITSCORE$, which is used as an independent variable

[21] Credit score is considered exogenous because it is calculated based on the individual's credit and payment history. To check for the possibility that the credit score is endogenous, we used a two-stage instrumental variable regression analysis that uses the income and age of the landlord as the instrument variables. However, the *F-statistics* of the first stage regressions using the 1) income only, 2) age only, and 3) both income and age as instrument variables are as low as 8.531, 3.387 and 6.579, respectively, which means that instrumental variables are weak instruments.

for further estimation. $HPICHG_{ijt-1}$ is calculated as $(HPI_{ijt-1} - HPI_{ijt-2})/HPI_{ijt-2}$. Thus, the key estimation model is

$$log(CHONSEI_{ijt}) = \alpha_i + \lambda_j + \tau_t + \beta_1 CREDITSCORE_{ijt} + \beta_2 HPICHG_{ijt-1} + \beta_4 X_{ijt} + \epsilon_{ijt} \tag{1.20}$$

where ϵ_{ijt} is the error term and X_{ijt} is the vector of additional covariates that represent the characteristics of the house and the household. α_i, λ_j, and τ_t represent a household fixed effect, a region fixed effect, and a year fixed effect, respectively. X_{ijt} includes a housing price variable, $log(VALUE_{ijt})$, which controls for the impact of the house price on $CHONSEI_{ijt}$. We expect $\beta_2 > 0$ and $\beta_1 < 0$.

The second empirical model using the survey data identifies the effect of house price expectations by relating the chonsei deposit for the household i in district j in year t to the expectations of future house prices as captured by $EXPECTATION_{ijt}$:

$$log(CHONSEI^1_{ijt}) = \alpha^1_i + \lambda^1_j + \tau^1_t + \beta^1_1 EXPECTATION_{ijt} + \beta^1_2 {X^1}_{ijt} + \epsilon^1_{ijt} \tag{1.21}$$

Like the first empirical method, X^1_{ijt} includes includes a housing price variable, $log(VALUE^1_{ijt})$ and other housing and household-related covariates. We expect $\beta^1_1 > 0$ in this estimation.

The third empirical model using the survey data identifies the effect of house price expectations on the contract type (pure vs mixed chonsei) for household i in district j in year t :

$$CONTRACTTYPE_{ijt} = \alpha^2_i + \lambda^2_j + \tau^2_t + \beta^2_1 EXPECTATION_{ijt} + \beta^2_2 {X^2}_{ijt} + \epsilon^2_{ijt} \tag{1.22}$$

$CONTRACTTYPE_{ijt}$ equals 1 for pure chonsei and 0 for mixed chonsei. Equation (25) is a linear probability model, but we also estimate a probit model. We expect $\beta^2_1 > 0$ in this estimation, given that a higher δ moves the contract toward pure chonsei ($\frac{\partial D^*}{\partial \delta} > 0$ and $\frac{\partial R^*}{\partial \delta} < 0$)

1.10 Empirical results

1.10.1 Estimation results using the bank data

The results of the estimation using the bank data are shown in Table 3. Our interest is the coefficients of $CREDITSCORE$ and $HPICHG$. First, $CREDITSCORE$ and $CREDITGRADE$ have negative coefficients at statistically significant levels, confirming what Proposition 1 predicts. Columns (1), (3), and (5) show the separate effects of credit score and $HPICHG$, and columns (2) and (4) show the effects of both variables together. The estimation results show that, as the credit score increases, the chonsei deposit decreases. That is, when the credit score increases by 100 points, the chonsei deposit decreases by 0.9%. The results of the estimation using the credit rating expressed in the letter form are shown in columns (3) and (4). It can be seen that when the credit rating goes up by one unit, the chonsei deposit decreases by about 1.8%.

Next, looking at the coefficients of $HPICHG$ in columns (2), (4), and (5), we can see that $HPICHG$ and the chonsei deposits are positively related at a statistically significant level, again confirming the model's predictions from Proposition 1. From the estimation results, we can see that the chonsei deposit increases by about 0.3% when house price appreciation rises by 1 percentage point, and that this result is robust across specifications.

The relationships between the chonsei deposit and other covariates suggest that a male landlord has a greater chonsei deposit than a female landlord. In addition, the chonsei deposit increases with the age and decreases with the income of landlord. Also, a 1% increase in house value leads to about a 0.8% increase in the chonsei deposit.[22]

[22] We also check the regression where the ratio of chonsei and house price (CTV = chonsei / Value) is the dependent variable (see the appendix). The results of the estimation are consistent with the regression with chonsei as dependent variable. One difference is that when the CTV is used as a dependent variable, the income of the landlord shows a significant relationship with CTV. This shows that the landlord considers the proportion of chonsei in the house price rather than the size of the chonsei according to his income.

Table 1.3: The relationship between the chonsei deposit, landlord's default cost, and house price appreciation

	(1)	(2)	(3)	(4)	(5)
	Estimation using Credit Score		Estimation using Credit Grade		Estimation using HPICHG
VARIABLES	log(CHONSEI)		log(CHONSEI)		
CREDITSCORE	**-8.96e-05*** **	**-8.64e-05** **			
	(3.40e-05)	**(3.40e-05)**			
CREDITGRADE			**-0.0178*** **	**-0.0175*** **	
			(0.00235)	**(0.00236)**	
HPICHG		**0.00316*** **		**0.00308*** **	**0.00316*** **
		(0.000939)		**(0.000936)**	**(0.000939)**
log(VALUE)	0.839***	0.838***	0.843***	0.842***	0.837***
	(0.00862)	(0.00863)	(0.00865)	(0.00865)	(0.00860)
SIZE	3.86e-05	3.87e-05	3.90e-05	3.91e-05	3.86e-05
	(4.74e-05)	(4.73e-05)	(4.72e-05)	(4.70e-05)	(4.74e-05)
SEX	0.0230***	0.0222***	0.0238***	0.0230***	0.0220***
	(0.00840)	(0.00840)	(0.00839)	(0.00839)	(0.00841)
AGE	0.00211***	0.00215***	0.00200***	0.00205***	0.00217***
	(0.000335)	(0.000335)	(0.000334)	(0.000335)	(0.000335)
INCOME	-2.30e-10***	-2.33e-10***	-2.21e-10***	-2.24e-10***	-2.40e-10***
	(7.62e-11)	(7.69e-11)	(7.45e-11)	(7.52e-11)	(7.71e-11)
JOBDUMMY	0.00926	0.0101	0.0130	0.0137	0.00862
	(0.00944)	(0.00945)	(0.00942)	(0.00942)	(0.00944)
JOBTYPE	-0.00980	-0.0101	-0.0150	-0.0152	-0.00846
	(0.0122)	(0.0122)	(0.0121)	(0.0122)	(0.0121)
Observations	13,588	13,542	13,588	13,542	13,542
R-squared	0.645	0.645	0.646	0.646	0.645

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1
All regressions include a constant, year fixed-effects and region fixed-effects

1.10.2 Estimation results using the survey data

The estimation results using the survey data are shown in Tables 4 and 5. Our main interest is the effect of the surveyed house price expectations on the chonsei deposit. This approach complements the indirect method of using the past housing appreciation rate from Brueckner, Calem, and Nakamura (2012). Columns (1) and (2) show the results of the empirical analysis with and without household characteristics. The estimation results show that if the house price is expected to rise in the future, the chonsei deposit will be 11% greater than if prices are expected to fall or stay the same. The age of respondent has a small but significantly negative relationship with the chonsei deposit while the effect of the house value on the chonsei deposit is again positive. Unlike estimation results using bank data, income has an insignificant relationship with chonsei. Also, the age of the landlord is negatively related to chonsei. This discrepancy presumably can be attributed to the difference in the composition of observations in the survey data, which were randomly selected from the population, although a full explanation is not clear.

Table 1.4: The relationship between the chonsei deposit and expectations of future house prices

	(1)	(2)
VARIABLES	log(CHONSEI)	log(CHONSEI)
EXPECTATION	**0.108****	**0.107****
	(0.0521)	**(0.0521)**
log(VALUE)	0.618***	0.609***
	(0.0369)	(0.0387)
SEX		0.0281
		(0.0681)
AGE		-0.00972***
		(0.00305)
INCOME		0.000193
		(0.000107)
Observations	1,863	1,863
R-squared	0.256	0.262

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

All regressions include a constant year fixed-effects and region fixed-effects

Table 5 shows the estimation results showing the effect of the surveyed house price expectations on the contract type. Columns (1) and (2) show the results of the linear probability model with and without household characteristics, while columns (3) and (4) show the estimation results using the probit model in the same specification. Looking at the coefficients of $EXPECTATION$, we can see that if the house price is expected to rise in the future, pure chonsei becomes more likely than mixed chonsei. This finding confirms the implication of the model, which shows that the contract move toward pure chonsei as δ increases. Also, the probability that the lease contract type is pure chonsei is positively related to the house value.[23]

Table 1.5: The relationship between the contract type and expectations of future house prices

	(1)	(2)	(3)	(4)
	Linear probability model		Probit model	
VARIABLES	Contract Type		Contract Type	
	(1=Pure chonsei, 0=Mixed chonsei)		(1=Pure chonsei, 0=Mixed chonsei)	
EXPECTATION	**0.0507****	**0.0500****	**0.158****	**0.156****
	(0.0215)	**(0.0216)**	**(0.0657)**	**(0.0658)**
log(VALUE)	0.0810***	0.0794***	0.235***	0.230***
	(0.0151)	(0.0160)	(0.0436)	(0.0461)
SEX		0.0311		0.0911
		(0.0302)		(0.0864)
AGE		-0.00187		-0.00539
		(0.00133)		(0.00394)
INCOME		2.86e-05		8.53e-05
		(4.19e-05)		(0.000127)
Observations	1,863	1,863	1,863	1,863
R-squared	0.070	0.072	0.055	0.057

Robust standard errors in parentheses in the linear probability model

T statistics in parentheses in the probit model

Pseudo R-squared is presented for the Probit model.

*** p<0.01, ** p<0.05, * p<0.1

All regressions include a constant year fixed-effects and region fixed-effects

1.11 Conclusion

This study examines chonsei as a means of housing investment by simultaneously incorporating the mortgage nature of chonsei and its character as a lease contract. To this end, we present a theoretical model and derive an equilibrium chonsei contract while empirically verifying the predictions of the model.

This model shows that the higher is the landlord's default cost, the smaller is the chonsei deposit, while the more favorable are expectations for future housing prices, the greater

[23] Appendix 4 shows the regression results using the ratio of chonsei and house price (CTV) as a dependent variable. The results are consistent with the results using chonsei as a dependent variable.

is the chonsei deposit. We confirm the validity of the theoretical predictions by showing empirically how the chonsei deposit changes as the landlord's credit rating and house price expectations change. The paper also complements existing research methodology that indirectly proxies price expectations through the housing price appreciation of the previous period. We empirically verify the predictions of the theoretical model using survey data that directly asks for expectations about future housing prices.

By presenting a solid theoretical model that incorporates the mortgage characteristics of the chonsei system and verifying it empirically, we are able to extend the existing studies on housing finance. Furthermore, this study is meaningful in that it predicts the change of the housing rental market according to the change of the economic environment by analyzing the effects of the change of the economic parameters on the chonsei deposit and lease contract type.

Bibliography

[1] Ambrose, B.W., and Kim, S.W., 2003. Modeling the Korean chonsei lease contract. Real Estate Economics 31, 53-74.

[2] Brueckner, J.K., 2000. Mortgage Default with Asymmetric Information. Journal of Real Estate Finance and Economics, 20, 251–74

[3] Brueckner, J.K., Paul S. Calem, and Leonard I. Nakamura., 2012. Subprime Mortgages and the Housing Bubble. Journal of Urban Economics, 71, 230–43

[4] Brueckner, J.K., Paul S. Calem, and Leonard I. Nakamura., 2016. House-price expectations, alternative mortgage products, and default. Journal of Money, Credit and Banking, 48, 81–112.

[5] Campbell, T. S., and J. K. Dietrich., 1983. The Determinants of Default on Conventional Residential Mortgage Loans, Journal of Finance 38, 1569-1579.

[6] Cho, S., 2010. Household wealth accumulation and portfolio choices in. Korea. Journal of Housing Economics 19, 13-25.

[7] Gyourko, J., and Han, J.K., 1989. Housing wealth, housing finance, and tenure in Korea. Regional Science and Urban Economics 19, 211-234.

[8] D. M. Harrison, T. G. Noordewier, and A. Yavas., 2004. Do riskier borrowers borrow more? Real Estate Economics, 32:385–411.

[9] Hendershott, P. H., and W. R. Schultz., 1993. Equity and Nonequity Determinants of FHA Single-Family Mortgage Foreclosures in the 1980s. Journal of the American Real Estate and Urban Economics Association 21, 405-430

[10] Kim, J., 2013. Financial repression and housing investment: an analysis of the Korean chonsei. Journal of Housing Economics 22, 338–358.

[11] Kim, K.H., 1990. An analysis of inefficiency due to inadequate mortgage financing: The case of Seoul, Korea. Journal of Urban Economics 28, 371-390.

[12] Lee, C.M., and Chung, E.C., 2010. Monthly rent with variable deposit: A new form of rental contract in Korea. Journal of Housing Economics 19, 315-323.

[13] Lekkas, V., J. M. Quigley, and R. Van Order.,1993. Loan Loss Severity and Optimal Mortgage Default. Journal of the American Real Estate and Urban Economics Association 21, 353-371.

[14] Navarro, I., and Turnbull, G.K., 2010. Antichresis leases: Theory and empirical evidence from the Bolivian experience. Regional Science and Urban Economics 40, 33-44.

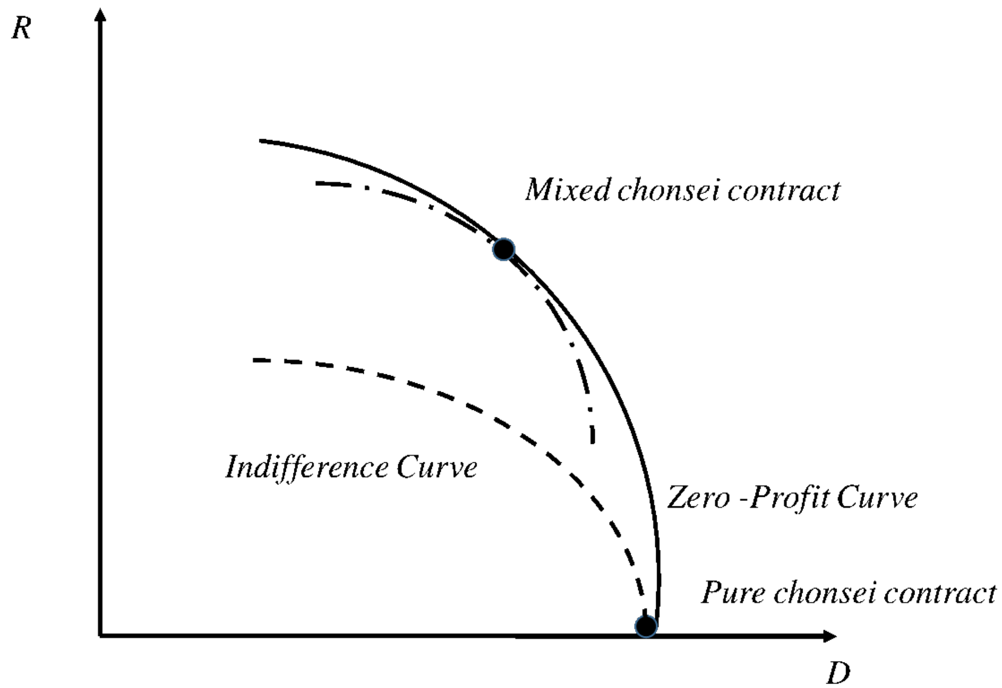
[15] Quigley, J. M., and R. Van Order., 1995. Explicit Tests of Contingent Claim Models of Mortgage Default. Journal of Real Estate Finance and Economics 11, 99-117.

[16] Renaud, B., 1989. Compounding financial repression with rigid urban regulations: Lessons of the Korean housing market. Review of Urban and Regional Development Studies 1, 3-22.

[17] Son, J.Y., 1997. A review of the Korean housing market and related policies. Review of Urban and Regional Studies 9, 80-99.

Appendix -1

Figure 1.3: Pure chonsei and mixed chonsei contract in $R-D$ plane



Appendix -2

Using (3), the default probability is obtained as

$$P(default) = \int_{\underline{P}}^{\frac{D-C}{\alpha}} f(P,\delta)dP = \frac{\frac{(D-C)}{\alpha} - \underline{p}}{\overline{p} - \underline{p}} \tag{1.23}$$

Using (16), we can check the change of the default probability according to the change of δ and c. Applying D^* to (23) gives the equilibrium likelihood default as

$$P^*(default) = \frac{2(\overline{p} - \underline{p}) + \delta - \frac{\theta(1+\eta)-\eta+\theta}{\alpha(\eta-\theta)}C}{\overline{p} - \underline{p}} \tag{1.24}$$

Differentiating P^* with respect to C and δ yields

$$\frac{\partial P^*(default)}{\partial C} = -\frac{\frac{\theta(1+\eta)-\eta+\theta}{\alpha(\eta-\theta)}}{\overline{p} - \underline{p}} \tag{1.25}$$

$$\frac{\partial P^*(default)}{\partial \delta} = \frac{1}{\overline{p} - \underline{p}} \tag{1.26}$$

As can be seen, a favorable shift in the future house price distribution (an increase in δ) raises P^*. This conclusion makes sense because the higher δ makes D^* increase and it will make the likelihood of default increase. However, the effect of C on the likelihood of default is ambiguous, being determined by the relative sizes of η and θ.

Appendix -3

Substituting D^* in (3) yields

$$(1+\theta)R^* = P_0 - D^* - \frac{1}{2}\theta(\overline{P}+\underline{P}) + \theta C - \theta(\frac{\frac{1}{2\alpha}(D^*-C)(D^*-C-2\alpha\overline{P})}{\overline{P}-\underline{P}}) \tag{1.27}$$

Let $\Omega \equiv -\frac{\theta(1+\eta)}{(\eta-\theta)}$ where $\Omega < 0$ by previous discussion. Then, after substituting D^* from (16), in (20), we can rearrange the expression for R^* to yield

$$(1+\theta)R^* = P_0 - (2\alpha + \frac{1}{2}\theta)\overline{P} + (\alpha - \frac{1}{2}\theta + \alpha\theta)\underline{P} \tag{1.28}$$

$$-\alpha(1+\theta)\delta + (2\theta - \Omega(1+\theta))C - \theta(\frac{\frac{(\alpha\delta+(\Omega-1)C)^2}{2\alpha} - \frac{2\alpha-1}{4}\underline{P}^2)}{\overline{P}-\underline{P}})$$

Let $\Theta \equiv \frac{(\alpha\delta+(\Omega-1)C)^2}{2\alpha}$. Then, we can rewrite (28) as

$$(1+\theta)R^* = P_0 - (2\alpha + \frac{1}{2}\theta)\overline{P} + (\alpha - \frac{1}{2}\theta + \alpha\theta)\underline{P} \tag{1.29}$$

$$-\alpha(1+\theta)\delta + (2\theta - \Omega(1+\theta))C - \theta\frac{\Theta}{\overline{P}-\underline{P}} + \theta\frac{\frac{2\alpha-1}{4}\underline{P}^2}{\overline{P}-\underline{P}}$$

Our interest is in how R^* changes with changes in C and δ. As can be seen from (29), when C increases, R^* increases because factors multiplying C are positive ($(2\theta-\Omega(1+\theta)) > 0$, and $-\theta\Theta_C^{'} > 0$ since $(\Omega-1) < 0$). Also, factors multiplying δ are negative ($-\alpha(1+\theta) < 0$, and $-\theta\Theta_\delta^{'} < 0$ since $\alpha > 0$). Therefore, R^* decreases as δ increases.[24]

[24] $-\theta\Theta_C^{'} > 0$ and $-\theta\Theta_\delta^{'} < 0$ hold when $\alpha\delta+(\Omega-1)C > 0$ and $\alpha\delta+(\Omega-1)C$ is less than 0 but its absolute

value is not too large.

Appendix -4

Table 1.6: The relationship between the CTV, landlord's default cost, and house price appreciation

	(1)	(2)	(3)	(4)	(4)
	Estimation using Credit Score		Estimation using Credit Grade		Estimation using HPICHG
VARIABLES	CTV = (CHONSEI / VALUE)*100				
CREDITSCORE	**-0.00172****	**-0.00162***			
	(0.000868)	**(0.000866)**			
CREDITGRADE			**-0.482*****	**-0.472*****	
			(0.0689)	**(0.0690)**	
HPICHG		**0.0760*****		**0.0738*****	**0.0760*****
		(0.0250)		**(0.0249)**	**(0.0250)**
log(VALUE)	-4.383***	-4.392***	-4.267***	-4.277***	-4.420***
	(0.382)	(0.382)	(0.386)	(0.386)	(0.379)
SIZE	0.000889	0.000890	0.000901	0.000902	0.000888
	(0.00117)	(0.00116)	(0.00116)	(0.00115)	(0.00116)
SEX	0.662***	0.645***	0.686***	0.668***	0.640**
	(0.249)	(0.249)	(0.248)	(0.248)	(0.250)
AGE	0.0549***	0.0559***	0.0520***	0.0530***	0.0562***
	(0.00977)	(0.00979)	(0.00981)	(0.00983)	(0.00978)
INCOME	-4.61e-09***	-4.68e-09***	-4.31e-09***	-4.37e-09***	-4.81e-09***
	(1.57e-09)	(1.59e-09)	(1.52e-09)	(1.54e-09)	(1.59e-09)
JOBDUMMY	0.280	0.299	0.393	0.410	0.272
	(0.276)	(0.276)	(0.277)	(0.277)	(0.275)
JOBTYPE	-0.116	-0.125	-0.269	-0.276	-0.0954
	(0.329)	(0.330)	(0.330)	(0.331)	(0.329)
Observations	13,588	13,542	13,588	13,542	13,542
R-squared	0.151	0.151	0.154	0.153	0.150

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

All regressions include a constant, year fixed-effects and region fixed-effects

Table 1.7: The relationship between the CTV and expectations of future house prices

	(1)	(2)
VARIABLES	CTV = (CHONSEI / VALUE)*100	CTV = (CHONSEI / VALUE)*100
EXPECTATION	**5.073****	**4.939****
	(2.407)	**(2.403)**
log(VALUE)	-29.32***	-31.30***
	(3.833)	(4.146)
SEX		3.128
		(2.876)
AGE		-0.155
		(0.151)
INCOME		0.00181***
		(0.00435)
Observations	1,863	1,863
R-squared	0.183	0.191

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

All regressions include a constant year fixed-effects and region fixed-effects

Chapter 2

The Effect of FAR (Floor Area Ratio) Regulations on Land Values: The Case of New York

2.1 Introduction

Land use regulations play an important role in determining the aesthetics of a city, the price of a property, and the economic choices of inhabitants. Each community has a unique set of such regulations which have different impacts in the economy. In particular, zoning laws, which set the guidelines for the type of usage for a parcel of land, or density regulations such as building-height restrictions or minimum lot size rules, are popular methods of segregating incompatible uses of land and controlling the rapid overcrowding of cities.

Being able to gauge the stringency of such regulations, namely, the extent to which they cause land use outcomes to diverge from free market outcomes, is crucial in understanding their impact. Building-height regulation, for example, is stringent when the regulated height is far below the free market height. As Bertaud and Brueckner (2005) argue, a height or density regulation limits the supply of housing within a given land area and thus contributes to urban sprawl. Furthermore, Barr (2016) shows that density regulations in high density cities like New York lower the future supply of housing and create a burden on the residents through higher housing costs. Although density regulations have the benefit of decreasing the cost of public infrastructure by lowering population density, they increase commuting costs and decrease social welfare, as in the study of India by Brueckner and Sridhar (2012). Knowing the stringency of a density regulation therefore allows the policy maker to gauge its impact on housing costs and urban form.

One of the most popular ways of regulating city density is through the regulation of maximum-allowed floor area ratio ("FAR"), which equals building's floor square divided by lot size and is thus a measure of building height. Empirical research has developed methods for measuring the stringency of regulation, including FAR regulation. Glaeser, Gyourko & Saks (2005) uses the difference between the price of housing and its construction cost as a measure of stringency of all regulations, including FAR limits. Under a competitive market, the price of housing should be close to its construction cost, and any positive difference between the two prices can be interpreted as a "regulatory tax" - a result of regulations restricting the supply of housing. Using their model, Glaeser et al. (2005) calculate this regulatory tax in New York City.

Brueckner et al. (2017) estimate the relationship between FAR regulation and land values in China, showing that the regression results can be used to measure the stringency of FAR regulation. The elasticity coefficient from the regression is a measure of stringency, with a large coefficient implying that the ratio of the free market FAR to the maximum-allowed FAR is large.

We make several important contributions to the literature on urban land use and density regulation. First, using the method of Brueckner et al. (2017), we estimate the stringency of FAR regulation in New York City, which is one of the most developed cities in the world. New York City is special in that, while each borough has a unique environment, each with its own subcenters and different characteristics, the entire greater metropolitan area is organized around Manhattan as its CBD. Thus it is useful to not only study the stringency of FAR in New York City as a whole, but to also compare the differences in stringency among all the five boroughs. Furthermore, we are able to demonstrate robustness of the method not only in developing countries like China, but also in a highly developed city like New York.

Second, we contribute to the literature on the estimation of vacant land values. There is a myriad of challenges involved in trying to create a data set consisting of vacant land values. In a highly developed city, there may not be very many vacant plots of land available for sale, and even if there are, the data may not be available. Moreover, in the case where vacant land is bought and held for speculation, the sale price at which there was a sale may not accurately reflect the value of the land under its optimal use. Because of such difficulties in estimating vacant land values, most research dealing with vacant land values involves purchasing data from real estate companies or using private internal data[1].

We overcome these data difficulties by using publically available New York transactions data to capture land value. We borrow the method of Dye & McMillen (2007) and Gedal & Ellen (2018) by relying on demolition permit data to calculate vacant land values. Using estimated land values, we then apply the methodology of Brueckner et al. (2017) to measure the stringency of FAR regulation in New York.

We organize the next sections as follows. In Section 2, we present the theoretical model describing the relationship between FAR regulation and land value. In Section 3, we discuss our method of constructing vacant land value data. In Section 4, we present the results of our empirical exercise along with a discussion of our findings regarding stringency of FAR regulation.

2.2 The model

This section derives the relationship between FAR regulation and land values. We follow Brueckner et al. (2017), reproducing their derivation here for clarity and completness.

In a competitive urban land market, the problem facing a land developer is to choose

[1] Colwell & Munneke (1997) uses data from Real Estate Data, Inc. and Haughwout, Orr, & Bedoll (2008) uses data from CoStar Group

the optimal density of development. The price of housing is determined competitively and is dependent on its locational characteristics, which capture the attractiveness of the site. The housing developer's profit-maximization problem can be stated as

$$\max_{S} \pi = ph\left(S\right) - iS - r$$

where p is the price of floor space (either residential or commercial), $h\left(S\right)$ is square feet of floor space output per unit of land for S units of capital invested per unit of land, i is the capital price, and r is the land value per square foot.[2] Note that $h(S)$ represents the floor area ratio (FAR) since it equals floor space per unit of land. The first-order condition for the choice of capital per unit land is

$$ph'\left(S\right) = i$$

with S^{*} satisfying the optimality condition. A structure will be built up to the density where the marginal cost of capital is equal to the marginal revenue from the investment of capital.

Because the development market is competitive, land rent can be found by the zero profit condition at the optimal density:

$$r = ph\left(S^{*}\right) - iS^{*}$$

If the FAR is limited by regulation, then there is a maximum-allowed density, $\overline{h}$, which in turn limits the amount of capital that can be invested to $\overline{S}$. Then the land rent is given by

$$r = ph\left(\overline{S}\right) - i\overline{S}$$

when the FAR restriction is binding, the derivative of land rent with respect to the maximum-

[2] The production function for residential and commercial floor space may exhibit some differences, but assuming a common function is reasonable as an approximation.

allowed FAR is

$$\frac{\partial r}{\partial \overline{S}} = ph'\left(\overline{S}\right) - i > 0$$

Note that since $ph'\left(\overline{S}\right) = i$, $\frac{\partial r}{\partial \overline{S}} > 0$ when $\overline{S} < S^*$.[3] Now consider the elasticity of land rent with respect to the maximum-allowed capital investment:

$$\begin{aligned} E_{r,\overline{S}} \equiv \frac{\partial r}{\partial \overline{S}} \frac{\overline{S}}{r} &= \frac{\left[h'\left(\overline{S}\right)p - i\right]\overline{S}}{ph\left(\overline{S}\right) - i\overline{S}} \\ &= \frac{\left[h'\left(\overline{S}\right)p - h'\left(S^*\right)p\right]\overline{S}_j}{ph\left(\overline{S}\right) - \overline{S}h'\left(S^*\right)p} \\ &= \frac{\left[h'\left(\overline{S}\right) - h'\left(S^*\right)\right]\overline{S}}{h\left(\overline{S}\right) - \overline{S}h'\left(S^*\right)} \end{aligned}$$

where the second equality comes from eliminating i by using the first-order condition for the choice of housing density. Because of the concavity of h, we know that $h'\left(\overline{S}\right)\overline{S} < h\left(\overline{S}\right)$, so that the elasticity of land value with respect to the capital investment limit is less than one. Imposing the functional form $h(S) = S^\beta$ for the production function (where $\beta < 1$) yields

$$E_{r,\overline{S}} = \frac{\left(S^*/\overline{S}\right)^{1-\beta} - 1}{\frac{1}{\beta}\left(S^*/\overline{S}\right)^{1-\beta} - 1} \tag{2.1}$$

Therefore, the elasticity depends on the ratio of the optimal capital investment to the maximum-allowed capital investment, $S^*/\overline{S}$. Furthermore, differentiating (1), it is easy to show that

$$\frac{\partial E_{r,\overline{S}}}{\partial\left(S^*/\overline{S}\right)} > 0, \tag{2.2}$$

so that the elasticity $E_{r,\overline{S}}$ is large when the stringency of regulation is high, with $S^*/\overline{S}$ large. Thus the percentage increase in land value from increasing the capital investment limit in

[3] Similarly, Thrall (1987) shows that land values are greatest when planning density equals to the market density. Otherwise, land values are lower.

an area where the density regulation is relatively stringent is greater than that in an area where the density regulation is relatively loose.

A few additional steps give the elasticity of land value with respect to the maximum-allowed FAR, $\overline{h}$, which (unlike $\overline{S}$) is empirically observable. Since $\overline{h} = \overline{S}^{\beta}$, we have $\overline{S} = \overline{h}^{1/\beta}$, which implies

$$E_{r,\overline{h}} \equiv \frac{\partial r}{\partial \overline{h}} \frac{\overline{h}}{r} = \frac{1}{\beta} E_{r,\overline{S}}$$

and

$$\frac{\partial E_{r,\overline{h}}}{\partial \left(S^{*}/\overline{S} \right)} > 0 \tag{2.3}$$

As in (3), the above equation shows that the elasticity of land value with respect to FAR is higher in areas where $\frac{S^{*}}{\overline{S}}$ is higher - that is, where density regulation is more stringent. Since land value and the maximum-allowed FAR are both observed variables, it is straightforward to estimate this elasticity.

Suppose that we employ an empirical model to estimate $E_{r,\overline{h}}$ and thus obtain an elasticity coefficient, denoted θ. Then, assuming a value for the production function parameter β, we can set the estimated θ equal to $\frac{1}{\beta}$ times (1) and then solve for $\frac{S^{*}}{\overline{S}}$, subsequently solving for $\frac{h(S^{*})}{h(\overline{S})} = (\frac{S^{*}}{\overline{S}})^{\beta}$. This calculation yields

$$\frac{h \left(\overline{S} \right)}{h \left(S^{*} \right)} = \left(\frac{1 - \theta\beta}{1 - \theta} \right)^{-\frac{\beta}{1-\beta}} \tag{2.4}$$

Using this formula along with the empirical results, the relationship between the free market and optimal densities can be computed.

2.3 Empirical Analysis

2.3.1 Background

New York City has one of the most complicated sets of land use regulations in the country, a behemoth of 4,338 pages (and still counting). The first version of zoning regulations in New York appeared in 1916 and was called the New York City Zoning Resolution (the "1916 Zoning Resolution"). The main goal of this law was to stop skyscrapers from blocking light and air from reaching the streets, and this goal was primarily accomplished by restricting building heights and setting standards for the shape of skyscrapers. Throughout the years, the 1916 Zoning Resolution was continuously amended to match the rapid change in technology and society, and by the 1950s it "resembled a torn 'patchwork' " (Marcus, 1992) which required change.

In July 1956, the Chairman of the New York's City Planning Commission hired the architectural firm, Voorhees, Walker, Smith, and Smith ("Voorhees") to begin a new zoning study. By 1961, the revised version of the Zoning Resolution proposed by Voorhees was approved by the New York City Board of Estimate and became the 1961 New York City Zoning Resolution, which is still in effect today. One of the most significant changes between the 1916 Zoning Resolution and the 1961 Zoning Resolution was that in the latter version, density regulation was accomplished not by limiting building height, but by limiting the maximum-allowed FAR.

While there have been countless amendments to the 1961 Zoning Resolution, these amendments are mostly exceptions to the traditional 1961 Zoning Resolution and do not affect our study. At the onset, the 1961 Zoning Resolution allowed for a variety of uses of land that were compatible with the character of the neighborhood as long as they did not over-strain public infrastructure. And because some uses were more profitable than others, the unprofitable uses became obsolete unless the municipality proposed that the area be designated as a Special District. For example, in 1967 the New York City Planning Commission sought to preserve New York City as the theater capital of the nation, and proposed a plan for incentive zoning. This plan would provide developers building in the Special District designated in the area enclosed by 57th Street on the north, 40th Street on the south, Eight Avenue on the west, and the Avenue of the Americas on the east a floor area bonus of up to 44% if they included a legitimate theater as a part of the project. In a similar manner, other Special Districts were created, such as public areas and arcades along Broadway near Lincoln Center to preserve the area as an international center for the performing arts; extra retail space in Fifth Avenue to stimulate economic growth; and rehabilitation of residential areas in Little Italy to preserve and enhance the historic and commercial character of the community.[4]

During the 1950s and the 1960s, city planners sought to increase open space, and the 1961 Zoning Resolution included an incentive zoning device in the form of bonus floor area if a plaza or an arcade was included in the project. However, with a steady rise in crime, these open spaces were increasingly perceived to be a haven for drug-related and criminal activity. By the 1980s, plazas and arcades became obsolete and avoided by the public, and most districts either reduced or eliminated the incentives. Since the teardown and reconstruction of buildings that received bonus floor areas through such incentive zoning in the past would mean a lower FAR today, there is no reason for any developer to wish to completely tear down such buildings rather than renovating.

2.3.2 Data

The estimation of vacant land values is challenging yet crucial for the study of the role of land in economics. There have been various attempts to estimate the value of vacant land thus far.

[4]As in the case of the Special Midtown District discussed above, each Special District provides bonus floor area to developers who include certain features in their project. In the case of schools or hospitals, the developers are able to build up to a higher density designated by the maximum-allowed facility FAR, which is usually higher than the maximum-allowed residential or commercial FAR. In receiving bonus floor area in Special Districts, however, the bonus is in the form of increasing the maximum-allowed residential or commercial FAR, and not the facility FAR.

The paper that is most relevant to our method in land value estimation is Dye & McMillen (2007), which estimated vacant land values by merging property sales data and teardown permit data in the Chicago metropolitan area. This was done under the assumption that the sales price of teardown property was close to its land value. However, this method fails to account for the fact that some of these properties may stay vacant for a long time after the building was torn down. Such properties may have different characteristics or development potentials from properties where a new building is erected soon after a teardown, and the inclusion of such properties may lead to heterogeneity problems within the data set.

Our study overcomes this problem by using the rich data that is available through New York City's Open Data, a joint project between the Mayor's Office of Data Analytics (MODA) and the Department of Information Technology and Telecommunications (DoITT). The main data comes from Primary Land Use Tax Lot Output ("PLUTO"), an extensive and a very detailed data set with information about every plot of land in the five boroughs of New York. PLUTO is organized by New York City's Department of City Planning and has been available for free to the public since July 2013. Our extract of the data is from September 2016 and contains information on 834,182 plots of land. Some information found in PLUTO that is used in our study includes lot area, maximum allowable FAR, and built FAR. Furthermore, we also have information on the borough in which the property is located, address, type of property (residential or commercial), assessed value of property, the year that a property was built, and BBL (borough, tax block, & lot) code. In the data, properties are divided into three classes of maximum-allowed FAR: commercial FAR, residential FAR, and facility FAR. Among these, the only classes of FAR that are relevant to the property types that we study are the commercial FAR and the residential FAR. Facility FAR determines the maximum-allowed FAR for public buildings such as schools or hospitals. In examining the relationship between land value and FAR, we are interested in minimizing unobserved characteristics of properties. As such, we matched each property's use type with one of the two classes (commercial or residential), using their maximum-allowed FARs.[5]

We combine the data from PLUTO with demolition permits and new building permits from New York City's Department of Buildings. In order to rehabilitate, tear down, or reconstruct a building, the property owner must first apply for a permit, and the permit is granted only after an inspection of the property. The types of permits range from teardown, new building, addition, and remodeling, but we focus on teardown and new building permits to estimate the value of vacant land. Next, we extract the 2003-2016 rolling sales data for properties from New York City's Department of Finance, using data from Google to measure the distance from the property to the Empire State Building (ESB). One issue with the rolling sales data is that it includes inheritance and gifting of property. In these cases, the sales price listed is simply for record and is too low to be counted as a legitimate sale of property. Because the sales price of these properties are not representative of the actual property value, we remove these data points.

To construct our data, we first match all the extracted data with the BBL (borough, tax block, and lot) code for each property. We find all properties that were issued teardown permits, then isolate properties that were also issued new building permits. Then, following Dye & McMillen (2007), we match these buildings with the rolling sales data by looking at sales that occurred in the interval starting two years a teardown permit was issued and ending at the time a new building permit was issued. Those restrictions yield 2,720 observations.

Our method of estimating vacant land values makes several contributions to the field. First, using both teardown data and new building permit data allows us to overcome the heterogeneity problem associated with only using teardown data. Also, the vacant land values that we estimate are more accurate than vacant land values estimated using only demolition permit data in the sense that all property included in our data has real development potential at the time of sale. Finally, our data consist of information that is publicly available, thus

[5]The bonus floor area that can be rewarded in Special Districts discussed in the previous section is determined after the sale and demolition of a property and is thus not observed in our data separately as bonus FAR. Since vacant land value should depend solely on the stated maximum-allowed FAR and not on the potential bonus floor area it could receive, the possibility of a bonus can be ignored.

Table 2.1: Summary Statistics - 2016 PLUTO Matched with Rolling Sales Data (2006-2016)

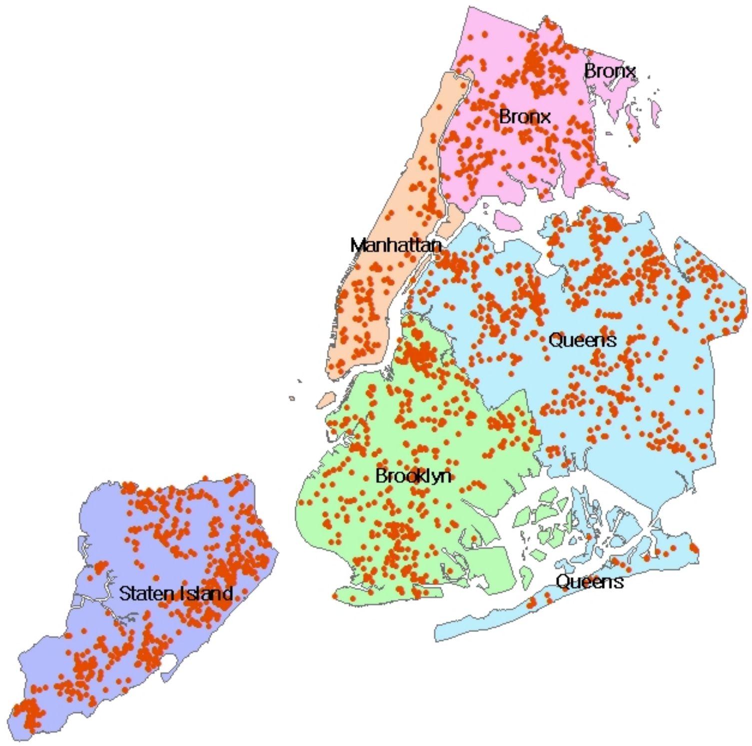
	No	Mean	S.D
Lot Area(sqft)	2720	35386.706	328609.064
Max FAR	2720	2.376	2.127
Built FAR	2720	1.732	2.595
Distance from Empire State Building	2720	9.810	4.418
Zone Type (Commercial=1, Residential=0)	2720	0.023	0.148
Sale Price($)	2720	1405280.348	5260531.644
Price Per square foot($)	2720	310.304	1052.636
N	2720		

allowing any researcher to replicate our steps without the burden of high costs associated with private commercial data.

Table 1 shows the summary statistics for the data constructed using the steps outlined above. Our data contains key information such as sales price, built FAR, and maximum-allowed FAR. The mean of the maximum allowable FAR in the sample is 2.376 and built FAR, which represents the total building area built in the lot, is 1.732 on average. The difference in these values shows that teardown properties tend not to be built to the maximum allowed FAR. The mean price per square foot for the sample is $310. Average distance from the Empire State Building is 9.8 miles.

In order for our constructed data to be representative of the population data, the sample must be chosen at random from the population. In particular, 1) our sample must not be clustered heavily in a particular borough, and 2) the distribution of the sample must not deviate significantly from that of the population. That our sample meets the first criterion can be seen from Figure 1, which shows that the vacant land values we estimate are quite evenly distributed among all of the boroughs.

Figure 2.1: Location Data of Matched Sample



2.4 Empirical strategy

Our primary interest is in the effect of FAR regulation on land values. There are two issues to consider in such a study: 1) endogeneity problems arising from unobserved characteristics of land and 2) the differences in the elasticity of land value with respect to maximum-allowed FAR in each of the boroughs.

To identify the effect of FAR on land value as given by (1), we regress the log of land value on the log of maximum-allowed FAR along with sales year fixed effects (dummy variables for individual sale years). Using the entire sample, the regression is given by

$$log(r_{it}) = \alpha_t + \theta log(FAR_{it}) + \epsilon_{it}, \tag{2.5}$$

where i denotes the parcel of land, t denotes the year of sale, and ϵ is the error term for the individual parcel. θ is the elasticity of land value with respect to the maximum-allowed FAR and α_t is the year fixed effect.

Equation (5) may lead to an omitted variable bias, as we are unable to filter out unobserved building characteristics that affect land value and are correlated with FAR. To reduce this bias, we use zip code fixed effects (dummy variables for individual zip codes) under the assumption that properties within a zip code have relatively similar characteristics. Then changes (5) to

$$log(r_{ict}) = \alpha_t + \beta_c + \theta log(FAR_{ict}) + \epsilon_{ict}, \tag{2.6}$$

where c is the zip code where the property is located and β_c is the zip code fixed effect.

Additionally, it is possible to examine how the effect of FAR restriction on land value varies across space. Barr (2014) shows that in New York, the height of buildings decreases with distance from the ESB, and Brueckner et al. (2017) shows that in Beijing, the FAR limit decreases with distance to the center of the city. To incorporate distance, we estimate

$$log(r_{ict}) = \alpha_t + \beta_c + \theta log(FAR_{ict}) + \gamma x_i + \eta(x_i * log(FAR_{ict})) + \epsilon_{ict} \tag{2.7}$$

where x is the distance from the ESB. The effect of maximum-allowed log FAR on log land value (the partial derivative of the right hand side of (7)) depends on the distance from the ESB and is given by $\theta + \eta x_i$. In the same way, the effect of distance from the ESB on land value depends on the maximum-allowed FAR and is given by $\gamma + \eta \log(FAR_{ict})$. In addition to the above, we take the approach of Barr (2014) by introducing borough subcenters.

Finally, we predict that the effect of FAR regulation on land values will manifest differently in each borough. For example, a unit increase in the maximum-allowed FAR in Manhattan may lead to higher increases in land values compared to other boroughs. One explanation for this prediction is that in New York, the characteristics of each of the borough are very distinct, which would affect the average profit-maximizing FAR (S^*). For example, Manhattan is considered to be the center of the greater New York metropolitan area, and the average profit-maximizing FAR is likely to be much greater than in other boroughs. As a result, the stringency of the FAR regulation in Manhattan may be different from that of the other boroughs. To test this theory, we create dummy variables for each borough and interact these variables with the log of FAR limit to compare the effect of FAR regulations in each borough. The resulting regression is

$$\begin{aligned} log(r_{it}) =& \alpha_t + Manhattan\, dummy * \theta_1 log(FAR_{it}) + Bronx\, dummy * \theta_2 log(FAR_{it}) \\ &+ Brooklyn\, dummy * \theta_3 log(FAR_{it}) + Queens\, dummy * \theta_4 log(FAR_{it}) \\ &+ StatenIsland\, dummy * \theta_5 log(FAR_{it}) + \epsilon_{it} \end{aligned} \quad (2.8)$$

2.4.1 Basic Empirical results

Tables 2 shows the estimation results, with column (1) corresponding to the model in (5) and column (2) adding zip code fixed effects, as in (6). The estimated coefficients in columns 1 and 2 show that an increase in maximum-allowed FAR increases land value. There is evidence that the estimation of (5) leads to an upward bias, in that with the inclusion of zip code fixed effects, there is a decrease of more than 50% in the coefficient, from 0.801 to 0.398. The zip code fixed effects allows us to alleviate the problem of omitted variable bias by controlling for many unobserved property characteristics. The estimation result is consistent with the findings in Brueckner et al. (2017), showing that it is possible to estimate the effects of land use controls like FAR regulation on land values even in developed cities like New York.

We calculate $\frac{h(\overline{S})}{h(S^*)}$ in equation (4) using the results in Table 2. We assume that β values from the function $h(S) = S^\beta$ are 0.6 and 0.8 and use the estimated θ value of 0.398 in Table 2. Using the formula in (4), $\frac{h(\overline{S})}{h(S^*)}$ then equals 0.70 and 0.61 when $\beta = 0.6$ and 0.8, respectively. The implication is that the regulated height of buildings in New York is about 61 to 70% of the free-market height.

Table 2.2: Regression Results Using Matched Data; Rolling Sales Data (2006-2016)

	(1)	(2)
VARIABLES	LN(Land Price/lot area)	
Ln(FAR)	0.801***	0.398***
	(0.0515)	(0.0710)
Zipcode Fixed effects	No	Yes
Zone Type (Commercial=1, Residential=0) Fixed Effects	Yes	Yes
Sales Year Fixed Effects	Yes	Yes
Observations	2,720	2,720
R-squared	0.260	0.586

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

2.4.2 Empirical results with distance

Next, we estimate the effect of distance from the ESB on the maximum-allowed FAR using equation (7). Table 3 shows the results, with column (1) and (2) including a distance term (with and without zip code fixed effects) and columns (3) and (4) adding a distance-log FAR interaction term. We see that the coefficients of the interaction term in columns (3) and (4) are negative, indicating that the stringency of FAR regulation becomes looser as properties are farther away from the ESB. Thus, moving away from the CBD, height regulation becomes less stringent.

While the signs switch on the pure distance coefficients when we include the interaction term, the overall effect of distance on land value remains negative and is consistent with the results without the interaction term. Note also that when we include zip code fixed effects, pure distance coefficient becomes insignificant, reflecting the fact that properties within a zip code all have similar distance from the ESB. The exclusion of zip code fixed effects, however, fails to account for unobserved property characteristics. Therefore, we turn to the utilization of subcenters to separate properties into different boroughs in order to account for these unobserved property characteristics, under the assumption that properties within a borough are similar to each other.

Table 2.3: Regression Results Using Matched Data; Rolling Sales Data (2006-2016) with Zip Code Fixed Effect and Interaction Term

VARIABLES	(1)	(2)	(3)	(4)
Ln(FAR)	0.443***	0.392***	1.444***	0.833***
	(0.0611)	(0.0709)	(0.0771)	(0.187)
Distance (Distance from Empire State Building)	-0.0550***	-0.0294	0.0921***	0.0150
	(0.00440)	(0.0272)	(0.00877)	(0.0333)
Interaction(Distance*Ln(FAR))			-0.159***	-0.0505**
			(0.00914)	(0.0218)
Zipcode Fixed effects	No	Yes	No	Yes
Zone Type Fixed effects	Yes	Yes	Yes	Yes
Sales Year Fixed effects	Yes	Yes	Yes	Yes
Observations	2,720	2,720	2,720	2,720
R-squared	0.296	0.586	0.400	0.587

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

2.4.3 Empirical results by borough

Our next interest is to explore the differences in the effects of the FAR regulation in each borough without controlling for distance. To explore this issue, we estimate equation (8), whose results are shown in Table 4. Column (1) shows the effect of the FAR restriction on land value in each borough. The key result here is that land values increase more for every unit increase in the maximum-allowed FAR in Manhattan than in any other borough, which shows that Manhattan has the most stringent FAR regulation.

The results in Table 4 are consistent with those in Table 3. Because stringency decreases with distance to the ESB in a regression without borough effects, it makes sense that Manhattan exhibits the highest stringency when borough effects are introduced. While the Bronx FAR coefficient in column (1) is negative, in contradiction to the theory, addition of zipcode fixed effects (column 2) flips the sign to positive and significant. The Staten Island coefficient, however, becomes insignificant in column (2), suggesting that FAR restrictions are not binding in that borough.

Table 2.4: Regression Results Using Matched Data; Rolling Sales Data (2006-2016)

	(1)	(2)
VARIABLES	LN(Land Price/lot area)	
Manhattan Dummy*Ln(Far)	1.172***	0.533*
	(0.0447)	(0.323)
Bronx Dummy*Ln(Far)	-0.310***	0.185**
	(0.0463)	(0.0906)
Brooklyn Dummy*Ln(Far)	0.256***	0.0941
	(0.0535)	(0.0713)
Queens Dummy*Ln(Far)	0.433***	0.423***
	(0.0351)	(0.0599)
Staten Island Dummy*Ln(Far)	0.493***	-0.142
	(0.0639)	(0.147)
Zipcode Fixed effects	No	Yes
Sales Year Fixed Effects	Yes	Yes
Observations	2,696	2,696
R-squared	0.455	0.596

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

2.4.4 Empirical results with distance by borough

One feature of New York is that every borough constitutes a unique culture and environment with its own subcenters. As a result, it is useful to consider how the distance from a subcenter affects the land value of properties and the stringency of regulation within a particular borough. We use the subcenters in each borough identified by Barr (2014) to estimate how the interaction between maximum-allowed FAR and distance from a subcenter affects land values in each borough. To accomplish this, we transform (7) as follows:

$$\begin{aligned} log(r_{ict}) =& \alpha_t + \beta_c + \theta log\left(FAR_{ict}\right) + \gamma x_i^1 + \eta_1 Manhattandummy * \left(x_i^1 * \theta_1 log(FAR_{ict})\right) \\ &+ \eta_2 Bronxdummy * \left(x_i^1 * \theta_2 log(FAR_{ict})\right) + \eta_3 Brooklyndummy * \left(x_i^1 * \theta_3 log(FAR_{ict})\right) \\ &+ \eta_4 Queensdummy * \left(x_i^1 * \theta_4 log(FAR_{ict})\right) + \eta_5 StatenIslanddummy * \left(x_i^1 * \theta_5 log(FAR_{ict})\right) + \epsilon_{ict} \end{aligned} \quad (2.9)$$

where x_i^1 is the distance from a borough's subcenter to a property within the borough. Barr (2014) shows that Manhattan has two subcenters - Wall Street and Downtown Manhattan;

Table 2.5: Regression Results Using Matched Data; Rolling Sales Data (2006-2016)

VARIABLES	(1) LN(Land Price/lot area)	(2) LN(Land Price/lot area)
Ln(FAR)	1.201***	0.814***
	(0.0756)	(0.150)
Distance (Distance from each subcenter)	0.152***	0.0391
	(0.0178)	(0.0364)
Manhattan Dummy*(Distance*Ln(FAR))	-0.0760***	-0.182
	(0.0253)	(0.124)
Bronx Dummy*(Distance*Ln(FAR))	-0.256***	-0.112**
	(0.0177)	(0.0474)
Brooklyn Dummy*(Distance*Ln(FAR))	-0.165***	-0.128***
	(0.0169)	(0.0337)
Queens Dummy*(Distance*Ln(FAR))	-0.137***	-0.0625
	(0.0185)	(0.0416)
Staten Island Dummy*(Distance*Ln(FAR))	-0.214***	-0.0855***
	(0.0226)	(0.0255)
Zipcode Fixed effects	No	Yes
Zone Type Fixed effects	Yes	Yes
Sales Year Fixed effects	Yes	Yes
Observations	2,720	2,720
R-squared	0.367	0.591

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

the subcenter in the Bronx is The Hub Neighborhood (at East 149th and Third Avenue); the subcenter in Brooklyn is downtown Brooklyn; the subcenter in Queens is Flushing and Jamaica; and there was no subcenter in Staten Island. In our empirical strategy, we used the shorter of the distance from the two subcenters in the case of Manhattan and used the ESB in the case of Staten Island. Our primary interest is in the estimation of η_j for $j = 1, ..., 5$ and we include zip code fixed effects.

As in our previous discussion, column (1) in Table 5 shows that the stringency of FAR regulation decreases with distance from subcenters in each of the five boroughs. In Manhattan and Queens, the inclusion of zip code fixed effects (column (2)) diminishes the significance of the interaction term, which may indicate that, when we control for unobserved variations, the stringency of FAR regulation is fairly uniform across space. The interaction coefficient for Brooklyn and Staten Island is negative and significant regardless of what fixed effects are included, and the coefficient for the Bronx depends on the specification. Because Manhattan as a whole is the center of the greater New York Metropolitan Area, distance from a subcenter does not affect the stringency of density regulation, whereas, in Brooklyn the subcenter plays a more important role. On the other hand, as discussed in Barr (2014), it is possible that Queens has multiple subcenters, and therefore that the distance from the chosen subcenter may not mean much for properties located near other subcenters.[6]

2.5 Conclusion

This paper studies the stringency of land-use regulation in New York City with a focus on FAR. An FAR regulation is said to be stringent when the unregulated, profit-maximizing FAR is significantly higher than the maximum-allowed FAR. Theory predicts that the elasticity of land value with respect to maximum-allowed FAR is higher in areas where the density restriction is more stringent. We estimate this elasticity for New York City, finding that the stringency of building height regulation is greatest in the borough of Manhattan. Even though Manhattan already has the tallest building in New York, the results suggest that regulations restrict heights there more that in other boroughs.

These conclusions are derived making use of publically available data, and we introduce a method to resolve some of the issues in calculating vacant land values through the use of both demolition permits and new building permits. The research helps to advance the state of knowledge regarding land use regulation.

[6]It is difficult to interpret the results for Staten Island, as we used the ESB as the subcenter. However, we are able to confirm that as we get farther away from the ESB, the stringency of density regulation does become looser.

Bibliography

[1] Barr, J. (2014). "The Floor Area Ratio Gradient: New York City, 1890-2009." *Regional Science and Urban Economics*, 48:110-119.

[2] Barr, J. (2016). *Building the Skyline: The Birth and Growth of Manhattan's Skyscrapers.* New York: Oxford University Press.

[3] Bertaud, A. & J.K. Brueckner. (2005). "Analyzing Building-Height Restrictions: Predicted Impacts and Welfare Costs." *Regional Science and Urban Economics*, 35(2):109-125.

[4] Brueckner, J.K. (1987). "The Structure of Urban Equilibria: A Unified Treatment of the Muth-Mills Model." *Handbook of Regional and Urban Economics*, 2:821-845.

[5] Brueckner, J.K. (1980). "A Vintage Model of Urban Growth." *Journal of Urban Economics*, 8:389-402.

[6] Brueckner, J.K. (1981). "A Dynamic Model of Housing Production." *Journal of Urban Economics*, 10:1-14.

[7] Brueckner, J.K., S. Fu, Y. Gu, and J. Zhang. (2017). "Measuring the Stringency of Land Use Regulation: The Case of China's Building Height Limits." *The Review of Economics and Statistics*, 99(4):663-677.

[8] Brueckner, J.K. & K.S. Sridhar. (2012). "Measuring Welfare Gains from Relaxation of Land-Use Restrictions: The Case of India's Building-Height Limits." *Regional Science and Urban Economics*, 42(6):1061-1067.

[9] Dye, R.F. & D.P. McMillen. (2007). "Teardowns and Land Values in the Chicago Metropolitan Area." *Journal of Urban Economics*, 61(1):45-63.

[10] Gedal, M. & I.G. Ellen. (2018). "Valuing Urban Land: Comparing the Use of Teardowns and Vacant Land Sales." *Regional Science and Urban Economics*, 70:190-203.

[11] Glaeser, E.L., J. Gyourko, & R.E. Saks. (2005). "Why Have Housing Prices Gone Up?" *American Economic Review*, 95(2):329-333.

[12] Haughwout, Andrew, James Orr, and David Bedoll. 2008. "The Price of Land in the New York Metropolitan Area." *Federal Reserve Bank of New York Current Issues in Economics and Finance*, 14(3): 1–7

[13] Marcus, N. (1992). "New York City Zoning – 1961-1991: Turning Back the Clock – But with an Up-to-the-Minute Social Agenda." *Fordham Urban Law Journal*, 19(3):707-726.

Thrall G I (1987). "Land use and urban form: The consumption theory of land rent." *Methuen London, UK*

Chapter 3

The Impact Of R&D Grant On Venture Capital Financing: Focusing On R&D Grant Program of Korea

3.1 Introduction

As noted by the classical literature of Nelson (1959) and Arrow (1962), research and development (hereafter R&D) activities are under-invested by the market due to their non-rival nature as public good. These problems provide an important basis for the public sector to support private R&D activities.

R&D grant is a typical support method used by many countries. The literature on the financial side of the impact of private sector R&D support has focused on whether support crowds in or crowds out firms' R&D expenditure. Wallsten (2000) finds that firms that received SBIR programs reduced their internal R&D expenditures compared to those who did not, while Busom (2000) also showed that in a sample of Spanish firms that received R&D support, grants completely crowded out internal investment in 30% of the cases. On the other hand, Klette and Moen (2000) find that targeted R&D support increased firms' R&D expenditure using a sample of Norwegian firms. As a comparative analysis of these studies, Zuniga-Vicente, Alonso-Borrego, Forcadell, and Galan-Zazo (2014) compared 77 firm-level studies. They confirm that existing empirical studies show mixed results as to whether the subsidies provided by the public sector to the firm have crowded out private R&D investments.

R&D grants can also impact firms' incentives to seek outside investment for their R&D activities, but not many studies have examined this channel. In a theoretical and conceptual approach, Czarnitzki (2006) and Takalo and Tanayama (2010) argue that there is a '*resource effect*' that reduces R&D investment from outside, as the R&D grant provides the firm with the necessary liquidity. Lach (2002) argues that the government reduces demand for investment from external sources by selectively allocating grants to the highest-potential firms, who would likely attract the most private investors. On the other hand, Feldman and Kelley (2006) and Meuleman and De Maeseneire (2012) argue that an R&D grant has a positive effect on external investment by sending a signal to the market on the quality of the firm's technology. In particular, Howell (2017) provides an empirical analysis of energy-related firms supported by the SBIR program and their connection to investment from venture capital. Thus she shows that the probability of receiving investment from a venture capital firm is twice as high for firms supported by the SBIR program than for firms with no support from the initial phase of SBIR program.

In the empirical analysis of the effects of R&D support, sample selection and endogeneity issues are the key identification problems. In general, sample selection issues can be addressed by the use of the matching method. However, not many studies have used the matching method in the context of R&D support, since this method requires data on the firms that are selected by the government and those that are not, and these data are generally treated as confidential. Almus and Czarnitzki (2003) investigate the effects of R&D subsidies in East Germany using the matching method. Jaffe and Le (2017) also use the method to study the effect of R&D grants on innovation outcomes in New Zealand. However, as David, Hall and Toole (2000) point out, there still remains the problem of endogeneity in the selection process itself, even after accounting for sample selection.

Our research makes some important contributions to previous literature. First, we analyze the effect of R&D grants on VC investment receipt, which has not been sufficiently studied. This lends empirical support for existing hypotheses about the impact of R&D on firms' external financing. In particular, Korea has the highest per capita R&D grant expenditure among OECD countries, and analyzing the effects of the program on an industry-wide scale will make an important contribution to the existing literature. We also mitigate key identification issues in empirical studies using both matching and instrumental variable methods. In particular, Korea's unique program rules provide us with the opportunity to identify the effects of public support through the use of an effective instrument.

Our research proceeds as follows. In Section 2, we provide a general description of Korea's R&D system and a detailed description of the program that we examine. Section 3 describes the data, empirical strategy, and the results. Section 4 concludes with a summary of our findings and their policy implications.

3.2 Overview of Korean R&D Policies

Korea places great emphasis on R&D; its Constitution states that one of the most important roles of the public sector is the promotion of technology through support for R&D. The government currently allocates more than 5% of the total annual budget to supporting private-sector R&D activities, the highest per capita expenditure not only in Asia but also among OECD countries. The largest portion of Korea's R&D subsidies is composed of direct R&D grants, which, as of 2017, totaled about $ 19 billion. The government has a set of support regulations and project selection rules that apply to all departments, but the detailed operation is left to each ministry. All firms that have passed the minimum qualification requirements (such as default risk) can apply for R&D grants. Firms submit an application containing the purpose of the project, the size of the project, and the operational plan. The government then undertakes a qualitative evaluation of the applicant's eligibility which takes into consideration the purpose of the project, its growth potential, and innovativeness.

The Korean program features a unique rule called *'Government grant pay-back system'*. A firm that receives R&D grant from the government to conduct R&D projects must return a certain portion of the grant to the government if they want to commercialize the project output at its conclusion. In the case of a large firm, 40% of the amount supported must be returned to the government. In the case of a small firm, the reimbursement rate is 10%. Firms are categorized into large and small firms by the number of employees. A firm with more than 300 employees is classified as a large firm, while a company with less than 300 employees is classified as small.[1]

This regulation reflects a compromise between the need to publicly support R&D and the government's budgetary constraints.[2]

3.3 Empirical approach

In this section, we analyze the impact of government subsidies on the future likelihood of receiving investment from venture capital using internal data provided by the Korean government for the first time. In particular, we use two empirical approaches to address selection bias and endogeneity issues that are often present in estimating the effects of government R&D policies. As Jaffe (2002) notes, the unobserved characteristics of firms that receive R&D grants will form a strong positive correlation with innovation performance. The unobserved variables of firms that receive grants and firms that do not receive grants may be different, and this difference may result in different outcomes. Therefore, identifying

[1]Number of employees is a main criterion that distinguishes between a large and a small firm. However, the capital owned by the firm is also considered in exceptional cases. For example, if the number of employees is more than 300, but the capital is less than $ 8 million, it is classified as a small firm, and in the opposite case, it is classified as a large firm. However, the number of employee criterion applies to over 95 % of firms in the sample.

[2]We interviewed relevant experts on the rationale behind this system. There were various explanations on the origins of this system, but the main theme is that the government tried to strengthen support for R&D and reign in government budget deficits in the 1970s and 1980s at the same time. There is also a similar rule in Israel, but an important difference is that in Israel, if a company makes a profit using developed technology, it must repay a certain percentage of its profit to the government for a certain period of time.

biases is important for accurate estimation of policy effects.

As a way to mitigate potential selection bias, we use the propensity marching score (PSM) method. The PSM method compares the treatment group with a control group constructed to have the closest characteristics to the treatment group, and estimates the treatment effect of receiving government grant. The first step is to obtain the propensity score, which is the probability of receiving government grant based on the characteristics of firms. Then, the treatment group and the non-treatment group are matched using the propensity score. The average treatment effect is then estimated by the weighted mean outcome difference between the treatment group and the non-treatment group.

Next, we address the endogeneity issue using the instrumental variable approach through a bivariate probit regression proposed by Woodridge (2010), by exploiting a special feature of the Korean R&D subsidy system. As described in the previous section, the program classifies firms into large and small based on a cutoff of 300 employees, and different pay back rates are applied to each firm. This difference in payback ratios leads to the prediction that the government will choose large firms over small firms so as to increase the payback income even if two firms of different sizes are otherwise identical. So we use a dummy for firm size (=1 if large firm) as an instrument to address the endogeneity issue. We use these two methods to test for the presence of potential biases and to estimate the policy effects controlling for these biases.

3.3.1 Data sources

The data used for empirical analysis is composed by merging three sets of data. The main data set is confidential government data, which is a list of all firms applying for R&D grant to the Ministry of Industry of Korea. This data includes general information such as the names and identifiers of all firms applying for R&D grant from 2009 to 2013, the number of employees, industry, and whether the firms is large or small.

The firm's financial information comes from one of the largest credit rating companies in Korea. Using the firm's identification number, we match information such as sales, debt ratio, and R&D intensity (ratio of internal R&D investment to sales).

Firms' venture capital receipt data uses the internal information of '*Venture Investment Disclosure System*' of the Ministry of Finance of Korea. This system is similar to the disclosure system in the stock market. Through this system, firms inform the government that they have received investment from venture capital (including direct investment and loan), and the government reviews and discloses this information to the public.[3] This system has a major purpose in providing information to the public that the firm has sufficient capabilities and skills to receive investment from the market. In order to induce the provision of such information, the government provides firms disclosing receipt of venture capital financing with fast-tracked listing in the stock market and patent examination, and offers tax benefits to research and development activities funded through venture investment. We use the firm identifier to match firms that received government grants to their venture investment receipt status in the subsequent three years (2012-2016).

The construction of this data set holds significance for a number of reasons. First, comprehensive data on R&D support programs in Korea, where government support for R&D plays a significant role, was used for the first time. In 2017, Korea's R&D budget is about 19 billion dollars, and the amount supported by the Ministry of Industry is about 5 billion dollars, which accounts for more than 30% of the total budget. The fact that most of the support for firms is done by the Ministry of Industry shows that this data is suitable for analyzing the effects of government grant on the industry. In addition, data of all firms applying for government grant (including all selected or not selected firms) were built at the firm level for the first time, and the nature of these data can be used to estimate the effect of government grant at the firm and industry level.

In addition, we can analyze the effect of the government grant on firms' receipt of external investment by using the characteristics of Korea's venture investment disclosure system.

[3]Venture capital invested in firms here refers to associations, private equity funds, etc. whose purpose is to invest in firms based on an assessment of firms' technical capabilities.

Thus far, little research has been done on the effect of government R&D grant on venture capital receipt. Howell (2017) matched data on firms that were supported by the US energy SBIR program with data that was manually collected for private financing deals, identifying the impact of government grants on subsequent VC investments. On the other hand, our data takes advantage of an institutional idiosyncrasy to overcome the limitation of existing data. Given the incentives of the disclosure system used by our research, firms with VC investments are expected to release the information, which can be effective in reducing the problem of missing data. Based on these data, we can contribute to the existing literature on the impact of government grants on firms' financing constraints and also test whether government grants crowd out or crowd in private investment.

Table 1 reports summary statistics of key variables. The data contains information on 21,606 firms applying for government grants from 2009 to 2013 and has information on whether they have been venture capital invested within three years after applying for a government grant. The firm's industry is classified on the basis of seven technologies; IT (Information Technology), BT (Bio Technology), NT (Nano Technology), ST (Space Technology), ET (Environment Technology), CT (Culture Technology) and other industries.

Table 3.1: Summary statistics

	Mean	S.D	Description
VC	0.1		Venture Capital Recipient (1=Receive investment from Venture Capital, 0=otherwise)
SELECT	0.3		Grant Recipient (1=Receive Grant, 0=otherwise)
GRANT	147.4	390.0	Grant amount
EMP	388.8	3646.1	Number Of Employees
SALES	409347.3	4285021.8	Revenues from the year before the government grant application
DEBT	296.8	10145.0	Debt Ratio(%)
RDINVEST	94.3	5254.3	R&D Intensity (Internal R&D Investment/Revenue)(%)
N	21606		

R&D grant sample is comprised of 5 years, which span from 2009 to 2013
VC recipient sample is comprised of 5 years, which span from 2012 to 2016
Variables with missing Std. Dev. are dummy variables.
Monetary units are in 1,000,000 Korean won, which is approximately equivalent to 1 thousand US dollars

3.3.2 Empirical strategy

Propensity Score Matching

Policy analysis is frequently complicated by the issue of sample selection. In other words, receiving a grant from the government may be correlated with other firm characteristics that also affect the outcome. Therefore, we use the propensity score matching method to reduce the potential selection bias in the estimation of policy effects. The first step is to obtain the propensity score, which is the expected probability of receiving an R&D grant based on firms' observed characteristics. In the next step, the treated and untreated firms are matched based on the propensity score. The average treatment effect is then estimated by the weighted average difference of the outcomes (VC investment receipt) of the two matched groups.

Denoting the outcome in the presence and absence of the policy treatment by VC_1 and VC_0, where VC is the outcome status (VC=1: received investment from venture capital ; VC=0: otherwise) , respectively, and with $SELECT$ the treatment status ($SELECT$ =1: treated (got government grant); $SELECT = 0$: untreated), the average treatment effects for the treated (ATT) can be defined as

$$ATT = E(VC_1 - VC_0|SELECT = 1) = E(VC_1|SELECT = 1) - E(VC_0|SELECT = 1) \tag{3.1}$$

In Equation (1), $E(VC_1|SELECT = 1)$ can be estimated with a simple mean of the outcome (VC) in the group of firms that are subsidized, but $E(VC_0|SELECT = 1)$ is by definition non-observable. In order to overcome this problem, $E(VC_0|SELECT = 1)$ needs to be substituted by referring to a suitable "counterfactual" of untreated firms. More precisely, in order to control for selection bias on observables, the difference in outcome between the two groups needs to be exclusively due to the policy intervention. One way to achieve this is by choosing untreated firms in such a way that they match treated firms in terms of their propensity score, $Pr(SELECT = 1|X(orP(X))$. In other words, untreated firms are to have the same probability of being funded than treated ones, given the set of pre-treatment characteristics, X, which are supposed to affect both the treatment and the outcome. The PSM estimate of the ATT is given by

$$ATT_{PSM} = E_{P(X)|SELECT=1}\{E[VC_1|SELECT = 1, P(X)] - E[VC_0|SELECT = 0, P(X)\} \tag{3.2}$$

We use the Kernel matching and the Calliper matching method for matching. The first approach is to generate a synthetic counterfactual for a particular treated firm by a kernel-weighted average of the characteristics of all matched untreated firms. The closer the propensity score of the untreated firm to that of the treated firm, the higher the weight assigned to that untreated firm in constructing the counterfactual case of the treated firm. The Caliper method matches the treated firm with a maximum of n-nearest untreated firms and applies the same weighting. We apply $n = 5$ in this study.

The propensity score can be a good predictor of treatment because the explanatory variables we use are key variables that the government use to screen R&D grants. By pairing similarly situated firms as treatments and controls in this manner, the PSM model reduces the selection bias due to unobserved variables.

As a first step, we estimate the impact of the firm's observed characteristics on the probability of receiving government R&D grants. The predicted probability is used as the propensity score in the PSM method. We estimate using the LPM and probit models, with the results shown in Table 2. As can be seen from this table, the probability of receiving a grant from the government increases as the firm's number of employees, sales, and internal R&D intensity increase. However, the higher the debt ratio, the lower the probability. In other words, the government has a tendency to choose firms with larger size, internal R&D investment, and less debt.

Table 3.2: Effects on the probability of receiving a R&D grant

	(1)	(2)
VARIABLES	LPM	probit
		Marg.Effects
log(EMP)	0.0366***	0.0420***
	(0.00422)	(0.00459)
log(SALES)	0.0544***	0.0625***
	(0.00305)	(0.00328)
log(DEBT)	-0.0410***	-0.0431***
	(0.00273)	(0.00276)
log(RDINVEST)	0.0382***	0.0444***
INDUSTRY DUMMY(BT)	(0.00241)	(0.00251)
	-0.167***	-0.0650***
	(0.0386)	(0.0111)
INDUSTRY DUMMY(CT)	-0.201***	-0.0934***
	(0.0416)	(0.0191)
INDUSTRY DUMMY(ET)	-0.185***	-0.0811***
	(0.0382)	(0.00912)
INDUSTRY DUMMY(IT)	-0.198***	-0.0955***
	(0.0380)	(0.00858)
INDUSTRY DUMMY(NT)	-0.0749*	0.0277*
	(0.0397)	(0.0149)
INDUSTRY DUMMY(ST)		0.0948**
		(0.0412)
INDUSTRY DUMMY(Etc)	-0.101***	
	(0.0382)	
Constant	0.516***	
	(0.0476)	
Observations	21,606	21,606
R-squared	0.215	

Standard errors in parentheses,Year fixed effects included

*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 3 shows our main estimation results. The PSM approach assumes that the treated and untreated firms are sufficiently similar after matching based on the Propensity score. As described above, we estimated the Kernel matching method and the Caliper matching method using bandwidths of 0.05 and 0.01, respectively. The bandwidths represent the difference that can be tolerated in the matching process. The lower the bandwidth, the more conservative the matching between treated and untreated firms. The graph in Appendix 1 shows whether PSM matched samples meet the common support condition that any firm characteristics do not completely predict treatment status. In other words, this condition means that for any specific characteristic that a treated firm possesses, the control group must also have at least one firm that has the same characteristic. This can be confirmed by the fact that the density function of the propensity score of the treated firm and that of the controlled firm are sufficiently overlapped.[4]

Table 3.3: Estimation using PSM- Effects of R&D grant receipt on Venture Capital Investment Receipt

		(1)	(2)	(3)	(4)
		All sample			
Specification		Kernel, bandwidth 0.05	Kernel, bandwidth 0.01	Calliper, bandwidth 0.05	Calliper, bandwidth 0.01
	Mean of control	0.1424	0.1416	0.1425	0.1424
VC investment Receipt	Treatment effect	**-0.0228*** **	**-0.0220*** **	**-0.0228*** **	**-0.0227*** **
		(0.00501)	**(0.00554)**	**(0.00762)**	**(0.00728)**
	Margianl effect	-0.160	-0.155	-0.160	-0.159

Standard errors in parentheses, Standard errors are bootstrapped with 100 replications.
*** p<0.01, ** p<0.05, * p<0.1

From the estimation results in Table 3, we can confirm that R&D grant receipt reduces the probability of receiving VC investment. Although results vary slightly with the estimation method used, receipt of an R&D grant lowers the probability of VC investment receipt by 16 percent as compared to firms without R&D grant (treatment effect of 2.3 percentage points, relative to the VC investment receipt rate in the control group of 14 percent). In general, the standard deviation of estimates are larger when a small bandwidth is used, but all estimates are statistically significant.

The estimation results offer empirical support for the views of Czarnitzki (2006) and Takalo and Tanayama (2010), who claim that government R&D grants reduce recipient firms' demand for external financing. In addition, this can be interpreted as evidence that the government, as Lach (2002) argues, tends to subsidize firms that are already capable of attracting private investment, as a result of applying too rigorous a screening process.

In order to check the robustness of the estimation results, we estimate the effects using treatment and control firms that are matched within the same industry category and year.

[4]Appendix 4 shows how the covariates before and after matching are balanced. Through this, we can confirm that the characteristics of firms become similar after matching.

This is to prevent inaccuracies that may arise from matching firms in different industry categories and years. Table 4 shows the estimation results, and it can be confirmed that matching within the same industry and year does not qualitatively alter our results. All estimates remain statistically significant, albeit somewhat smaller in magnitude.

Table 3.4: Estimation using PSM- Effects of R&D grant receipt on Venture Capital Investment Receipt (Same Industry-Year Match Sample)

	(1)	(2)	(3)	(4)
	All sample			
Specification	Kernel, bandwidth 0.05	Kernel, bandwidth 0.01	Calliper, bandwidth 0.05	Calliper, bandwidth 0.01
	0.1399	0.1412	0.1392	0.1425
VC investment Receipt	-0.0190***	-0.0165***	-0.0178***	-0.0178***
	(0.00557)	(0.00594)	(0.00636)	(0.00643)
	-0.136	-0.117	-0.128	-0.125

Standard errors in parentheses, Standard errors are bootstrapped with 100 replications.
*** p<0.01, ** p<0.05, * p<0.1

Bivariate Probit Model using Instrument Variable

A way to alleviate endogeneity problems in this set up is to use a bivariate probit regression taking into account the instrumental variable approach. We may reasonably suspect the presence of endogeneity in the form of unobserved variables affecting a firm's receipt of both VC investment and government grants. In order to mitigate this problem, we utilize the characteristics of the grant repayment policy of the government R&D system as an instrumental variable. The instrumental variable for selection is whether or not a firm is large, i.e. employs more than 300 people. As explained in the previous section, a large firm must pay back 40% of its grant to the government from the outcome of the project, while a small firm can repay only 10% of the grant and still benefit from the project. This difference in payback rates allows us to expect that the government will favor large firms in the selection process. So we use Z (1 if large firm, 0 otherwise) as the instrument variable for $SELECT$.

An additional consideration is that the main variables in our model are binary. The outcome variable (VC), treatment variable ($SELECT$), and instrumental variable (Z) are all binary, so this environment is suitable for using the bivariate probit model proposed by

Woodridge (2010). In this case, we can set the following model

$$VC_i = 1[\alpha_i + \beta SELECT_i + \gamma X_i + \varepsilon_i > 0]$$

$$SELECT_i = 1[\delta_i + \theta X_i + \phi Z_i + \nu_i > 0] \tag{3.3}$$

where the unobserved disturbance (ε_i , ν_i) is independent of X and assumed to be distributed as bivariate normal with mean zero. Also, each has unit variance and $\rho = Corr(\varepsilon_i\ , \nu_i)$. If $\rho = 0$ in this model, the likelihood of a bivariate probit model is simply equal to the sum of the likelihood of two univariate probit model

In our setting,VC is an outcome variable indicating whether the firm has received venture capital investment ($VC = 1$ if receive investment from VC, otherwise 0). $SELECT$ indicates whether the firm has received government grant ($SELECT = 1$ if granted, otherwise 0), Z is an instrument variable ($Z = 1$ if large firm, otherwise 0), and X is other covariates that can affect selection as well as VC investment receipt. Using this model, we can estimate the causal effect of having government grant on receiving investment from venture capital by allowing a correlation between the error terms of the two probit models.

First, we check that the instrumental variable has a strong correlation with the treatment variable. We examine the relationship between the two variables using the Linear Probability Model and the probit model respectively. Table 5 confirms that there is a statistically significant relationship between the instrument and the treatment variable, which confirms that the government leans toward large firms to provide grants.

Next, to confirm the validity of the instrument and the endogeneity of the treatment variable, we test the null hypotheses of weak instrumental variables through the F-statistic in the LPM model using two stage least squares. We then formally test endogeneity through the Wald test for ρ in the bivariate probit model. The high *F statistics* value in the LPM model using 2SLS in Appendix 3 reject the null hypotheses of weak instrumental variables, indicating that the instrumental variable is a strong predictor of $SELECT$. Also, the positive and significant correlation (ρ) between the unobserved factors of the two probit models in Table 6 confirms that there is a potential endogeniety issue.

Table 3.5: Estimation using probit and LPM using IV approach- First stage regression

	(1)	(2)
VARIABLES	LPM First stage	probit First stage
	T(1=Receive grant, 0=otherwise)	
Z(1 if large firm, otherwise 0)	**0.148*** **	**0.971*** **
	(0.0139)	**(0.0705)**
log(EMP)	0.0165***	0.0412***
	(0.00457)	(0.0144)
log(SALES)	0.0649***	0.216***
	(0.00291)	(0.00997)
log(DEBT)	-0.0444***	-0.136***
	(0.00284)	(0.00842)
log(RDINVEST)	0.0431***	0.144***
	(0.00248)	(0.00770)
Observations	21,606	21,606
R-squared	0.166	

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 6 shows the results of estimating the treatment effect to be identified through Equation (1). Our main estimation results are shown in column (3), (4) and (5). We estimated using Linear Probability Model to check robustness according to Woodridge (2010) and the results are shown in Appendix 2. Columns (3) and (4) are the results of the two probit models shown in Equation (1) and based on this, the marginal effects are obtained in column (5). First, by checking columns (1), (2) and (5), we can confirm that government grant receipt has a negative and significant impact on receiving VC investment. Also, by comparing columns (2) and (5), it can be seen that there is an upward bias due to unobserved variables.[5]

[5] We can reasonably assume that the size of grants received from the government can affect the investment from VC. To confirm this, we estimate the effect through the following simultaneous equation model in Appendix 3. In other words, since only the firm selected by the government observes the size of the grant, we estimate the treatment effect by including the amount received from the government in the probit model for VC, assuming that $SELECT$ is only affected by Z. Estimation results are shown in columns (1), (2) and (3). The impact of receiving a grant from the government on receiving investment from a VC remains unchanged, even after including the size of the grant, and still shows a negative and significant relationship.

When comparing the estimation results using the PSM method in the previous section and the results using the IV method, the results do not represent a meaningful difference (16% and 10%, respectively). We can confirm that the R&D grant has a negative and statistically significant effect on VC investment receipt, and we can check the robustness of these results using two different methods.

In addition, the greater the number of employees, the greater the marginal probability of receiving an investment from VC, while the greater the sales, the less likely the marginal probability of receiving an investment from VC. Also, it can be seen that the debt ratio has a negative and significant relationship with the investment from the VC, and the internal R&D investment ratio has a positive but insignificant relationship with the investment from the VC.

Table 3.6: Estimation using probit,Bivariate probit, and LPM using IV approach- Effects of R&D Grant receipt on Venture Capital Receipt

	(1)	(2)	(3)	(4)	(5)
Dependent Variable			VC	SELECT	VC
	Simple probit		Bivariate probit-Use Z as IV		
	Marg. Effects	Marg. Effects	Coefficients	Coefficients	Marg. Effects
SELECT	**-0.0229*****	**-0.00729**	**-0.459*****		**-0.102*****
	(0.00539)	**(0.00661)**	**(0.124)**		**(0.0287)**
Z(1 if large firm, otherwise 0)				1.061***	
				(0.0735)	
GRANT		-4.72e-05***			
		(1.18e-05)			
log(EMP)	0.00538	0.00748**	0.0369**	0.0614***	0.00819**
	(0.00336)	(0.00340)	(0.0157)	(0.0150)	(0.00351)
log(SALES)	-0.0251***	-0.0244***	-0.0940***	0.191***	-0.0209***
	(0.00230)	(0.00231)	(0.0132)	(0.0104)	(0.00278)
log(DEBT)	-0.00137	-0.00134	-0.0211*	-0.133***	-0.00468*
	(0.00212)	(0.00212)	(0.0109)	(0.00869)	(0.00246)
log(RDINVEST)	-0.00186	-0.00104	0.00572	0.132***	0.00127
	(0.00191)	(0.00192)	(0.0100)	(0.00796)	(0.00224)
INDUSTRY DUMMY(BT)	0.0125	0.0137	0.0323	-0.202***	0.00717
	(0.00864)	(0.00865)	(0.0403)	(0.0344)	(0.00894)
INDUSTRY DUMMY(CT)	-0.0196	-0.0195	-0.124*	-0.284***	-0.0276*
	(0.0152)	(0.0152)	(0.0701)	(0.0592)	(0.0156)
INDUSTRY DUMMY(ET)	0.0133*	0.0142*	0.0296	-0.257***	0.00657
	(0.00736)	(0.00737)	(0.0353)	(0.0286)	(0.00783)
INDUSTRY DUMMY(IT)	0.0243***	0.0255***	0.0744**	-0.299***	0.0165**
	(0.00684)	(0.00685)	(0.0339)	(0.0269)	(0.00746)
INDUSTRY DUMMY(NT)	0.0140	0.0168	0.0720	0.0737	0.0160
	(0.0118)	(0.0118)	(0.0535)	(0.0464)	(0.0119)
INDUSTRY DUMMY(ST)	0.00248	0.00574	0.0467	0.293**	0.0104
	(0.0326)	(0.0328)	(0.149)	(0.130)	(0.0330)
INDUSTRY DUMMY(Etc)					
Constant			-0.181**	-1.596***	
			(0.0839)	(0.0782)	
Observations	21,606	21,606	21,606	21,606	21,606
Rho (ρ)				0.223***	
				(0.0799)	

Standard errors in parentheses, Year fixed effects included

*** p<0.01, ** p<0.05, * p<0.1

3.4 Conclusion

R&D grants have an important role in society in that they address the problem of the under-investment in innovation activities. However, such public support should be designed to increase recipients' innovative capacity and to facilitate investment activities in the market. We have estimated the impact of a Korean R&D grant program on firm's VC investment receipt based on an internal data that has not been used in previous literature. In particular, we addressed the issues of sample selection and endogeneity, which are key issues in the estimation of the effects of R&D policies, by using the PSM method and the instrumental variable method respectively, and by exploiting the idiosyncratic features of the Korean program.

The results of the empirical analysis suggest that a firm receiving the R&D grant has about 15% less chance of receiving VC investment receipt than a firm without the R&D grant. This result can be regarded as an empirical validation of Lach's (2002) claim that government subsidies lead recipient firms to substitute away from private investment even when they are capable of attracting such investment. It is also consistent with the claims of Czarnitzki (2006) and Takalo and Tanayama (2010) that grants themselves reduce the financial constraints of the firm and reduce the demand for external investment. Based on these results, we can derive some policy implications. First, targeted support based on the nature of the firm can be more effective than general R&D support. Howell (2017) asserts that the effect of support for younger and smaller firms is greater than for firms that are not. Therefore, it may be worth considering designing a targeted support policy that takes this analysis into consideration. Also, the system should be designed in such a way that it prevents the firms receiving the grant from becoming uncompetitive in the market.

Bibliography

[1] Aghion, P. and Howitt, P. (1992), “A model of growth through creative destruction.” Econometrica, 60(2) 323–351.

[2] Almeida, H., P. Hsu, and D. Li, (2012), “When less is more: Financial constraints and innovative efficiency”, Working paper, University of Illinois at Urbana-Champaign.

[3] Almus, M., Czarnitzki, D. (2003). “The effects of public R&D subsidies on firms’ innovation activities: The case of East Germany”. Journal of Business and Economic Statistics 21 (2), 226–236

[4] Arora, A., Ceccagnoli, M. and Cohen, W.M. (2008), “R&D and the patent premium.” International Journal of Industrial Organization, 26(6) 1153-1179.

[5] Arrow, K. J. (1962), “Economic Welfare and the Allocation of Resources for Invention.” In The Rate and Direction of Inventive Activity, edited by R. Nelson, 609–626. Princeton, NJ: Princeton University Press.

[6] Bronzini, R. and Iachini E. (2014), “Are incentives for R&D effective? Evidence from a regression discontinuity approach.” American Economic Journal: Economic Policy, 6(4) 100–134.

[7] Busom, I., 2000. An empirical evaluation of the effects of R&D subsidies. Econ.Innov. New Technol. 9 (2), 111–148.

[8] Cappelen, Å, Raknerud, A. and Rybalka, M. (2012) "The Effects of R&D Tax Credits on Patenting and Innovations." Research Policy 41 (2) 334–45. Cohen, L. R., Roger Noll (2002), The technology pork barrel. Brookings Institution Press.

[9] Czarnitzki, D. (2006). Research and development in small and medium-sized enterprises: The role of financial constraints and public funding. Scottish Journal of Political Economy, 53(3), 335–357

[10] Czarnitzki, D., Hanel, P. and Rosa, J.M. (2011) "Evaluating the impact of R&D tax credit on innovation: A micro-econometric study on Canadian firms." Research Policy, 40(2011) 217-229.

[11] David, P.A., Hall, B., Toole, A., 2000. Is public R&D a complement or substitute for private R&D? A review of the econometric evidence. Res. Policy 29 (4–5), 497–529

[12] Dechezleprêtre, Antoine, Elias Einiö, Ralf Martin, Kieu-Trang Nguyen, and John Van Reenen. (2016). "Do Tax Incentives for Research Increase Firm Innovation? An RD Design for R&D." CEP Discussion Paper dp1413. Centre for Economic Performance, LSE.

[13] Feldman, M. P., & Kelley, M. R. (2006). The ex-ante assessment of knowledge spillovers: Government R&D policy, economic incentives and private firm behavior. Research Policy, 35(10), 1509–1521.

[14] Gurmu, S. and Pérez-Sebastián, F. (2008) "Patents, R&D and lag effects: Evidence from flexible methods for count panel data on manufacturing firms." Empirical Economics, 35(3) 507-526. Hall, B. H.(2002), "The financing of research and development. " Oxford review of economic policy, 18(1), 35-51.

[15] Hall, B. H. (2010), "The financing of innovative firms. " Review of Economics and Institutions, 1(1). Hall, B. H., & Lerner, J. (2009), "The financing of R&D and innovation. " (No. w15325). National Bureau of Economic Research.

[16] Hall, Bronwyn H., Pietro Moncada-Paternò-Castello, Sandro Montresor & Antonio Vezzani. (2016), "Financing constraints, R&D investments and innovative performances: new empirical evidence at the firm level for Europe", Economics of Innovation and New Technology, 25:3, 183-196.

[17] Hall, B. and Ziedonis, R. (2001) "An empirical study of patenting in the U.S. semiconductor in- dustry, 1979-1995." RAND Journal of Economics, 32(1) 101-128.

[18] Hao, K. Y., and A. B. Jaffe. (1993), "Effect of Liquidity on Firms' R&D Spending." Economics of Innovation and New Technology 2 (4): 275–282. Howell, Sabrina T. 2017. "Financing Innovation: Evidence from R&D Grants." American Economic Review 107 (4): 1136–64

[19] Jacob, Brian and Lars Lefgren (2010) "The impact of research grant funding on scientific productivity", Journal of Public Economics, 95(9-10), 1168-1177.

[20] Klette, T. J., Møen, J., & Griliches, Z. (2000). Do subsidies to commercial R&D reduce market failures? Microeconometric evaluation studies. Research Policy, 29(4), 471–495.

[21] Lach, Saul (2002), "Do R&D subsidies stimulate or displace private R&D? Evidence from Israel" Journal of Industrial Economics, 50, 369-90.

[22] Meuleman, M., & De Maeseneire, W. (2012). Do R&D subsidies affect SMEs' access to external financing?. Research Policy, 41(3), 580–591. Nelson, R.R. (1959). "The simple economics of basic scientific research". Journal of Political Economy 49, 297–306.

[23] Takalo, T., & Tanayama, T. (2010). Adverse selection and financing of innovation: is there a need for R&D subsidies?. The Journal of Technology Transfer, 35(1), 16–41.

[24] Trinh Le & Adam B. Jaffe (2016) The impact of R&D subsidy on innovation: evidence from New Zealand firms, Economics of Innovation and New Technology, 26:5, 429-452

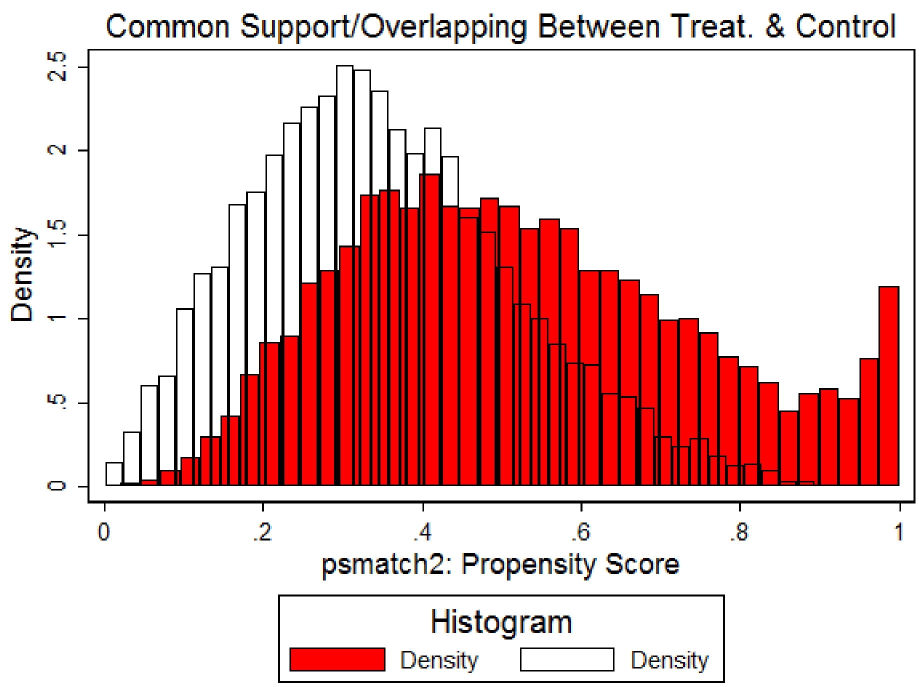
[25] Wallsten. S. (2000). "The effects of government-industry R&D programs on private R&D: the case of the Small Business Innovation Research program". RAND Journal of Economics 31, 82–100.

[26] Wooldridge, J.M. (2010) Econometric Analysis of Cross Section and Panel Data. Boston, MA: Massachusetts Institute of Technology

[27] Zúniga-Vicente, ˜ J., Alonso-Borrego, C., Forcadell, F.J., Galàn, J.I., 2014. Assessing the effect of public subsidies on firm R&D investment: a survey. J. Econ. Surv. 28 (1), 36–67.

Appendix-1

Figure 3.1: Common Support/Overlapping Between Treatment group & Control Group



Appendix-2

Table 3.7: Estimation using LPM - Effects of R&D Grant receipt on Venture Capital Receipt

	(1)	(2)	(3)	(4)
Dependent Variable				
	LPM			
	LPM	LPM	LPM-2SLS	LPM- Simultaneous
SELECT	**-0.0235*** **	**-0.0192*** **	**-0.411*** **	**-0.353*** **
	(0.00538)	**(0.00584)**	**(0.105)**	**(0.0893)**
GRANT		-1.22e-05*		-1.76e-05***
		(6.53e-06)		(6.06e-06)
log(EMP)	0.0143***	0.0151***	0.0285***	0.0261***
	(0.00334)	(0.00337)	(0.00535)	(0.00437)
log(SALES)	-0.0274***	-0.0271***	-0.00630	-0.00466
	(0.00242)	(0.00243)	(0.00632)	(0.00633)
log(DEBT)	-0.00159	-0.00160	-0.0175***	-0.0166***
	(0.00217)	(0.00217)	(0.00493)	(0.00452)
log(RDINVEST)	-0.00140	-0.00108	0.0134***	0.0142***
	(0.00192)	(0.00193)	(0.00455)	(0.00440)
Observations	21,606	21,606	21,606	21,606
Robust F-statistic for weak IV			70.196***	
R-squared	0.026	0.026		0.027

Standard errors in parentheses, Year fixed effects included

*** p<0.01, ** p<0.05, * p<0.1

Appendix-3

We can reasonably assume that the size of grants received from the government can affect the investment from VC. To confirm this, we estimate the effect through the following simultaneous model.

$$VC_i = 1[\alpha_i + \beta SELECT_i + \gamma X_i + \varepsilon_i > 0]$$

$$SELECT_i = 1[\delta_i + \phi Z_i + \nu_i > 0] \tag{3.4}$$

Table 3.8: Estimation using Bivariate using IV approach (Simultaneous equation model)-Effects of R&D Grant receipt on Venture Capital Receipt

	(1)	(2)	(3)
Dependent Variable	VC	SELECT	VC
	Bivariate probit-Simultaneous		
	Coefficients	Coefficients	Marg. Effects
SELECT	**-0.994*** **		**-0.244*** **
	(0.119)		**(0.0346)**
Z(1 if large firm, otherwise 0)		2.001***	
		(0.0581)	
GRANT	-0.000177***		-4.34e-05***
	(4.95e-05)		(1.21e-05)
log(EMP)	0.0538***		0.0132***
	(0.0139)		(0.00340)
log(SALES)	-0.0970***		-0.0238***
	(0.0102)		(0.00231)
log(DEBT)	-0.00570		-0.00140
	(0.00846)		(0.00207)
log(RDINVEST)	-0.00272		-0.000666
	(0.00770)		(0.00189)
INDUSTRY DUMMY(BT)	0.0561		0.0138
	(0.0346)		(0.00848)
INDUSTRY DUMMY(CT)	-0.0765		-0.0188
	(0.0605)		(0.0148)
INDUSTRY DUMMY(ET)	0.0592**		0.0145**
	(0.0296)		(0.00723)
INDUSTRY DUMMY(IT)	0.102***		0.0250***
	(0.0278)		(0.00674)
INDUSTRY DUMMY(NT)	0.0728		0.0179
	(0.0475)		(0.0116)
INDUSTRY DUMMY(ST)	0.0248		0.00609
	(0.133)		(0.0327)
INDUSTRY DUMMY(Etc)			
Constant	0.0886	-0.290***	
	(0.0914)	(0.00896)	
Observations	21,606	21,606	21,606
Rho		0.716***	
Robust F-statistic for weak IV		(0.124)	
R-squared			

Standard errors in parentheses, Year fixed effects included

*** p<0.01, ** p<0.05, * p<0.1

Appendix-4

Table 3.9: Test of balancing of covariates

		Mean		Mean		Mean	
		All		Small		Large	
Variable	Unmatched Matched	Treated	Control	Treated	Control	Treated	Control
log(Number Of Employees)	U	3.9779	2.9484	3.4048	2.9328	7.5666	5.964
	M	**3.6402**	**3.5503**	**3.402**	**3.3621**	**6.0868**	**6.116**
log(Sales)	U	9.2152	7.6055	8.4612	7.5875	13.951	11.088
	M	**8.7645**	**8.6311**	**8.4576**	**8.4087**	**11.716**	**11.663**
log(Debt Ratio)	U	4.6497	4.9534	4.6682	4.9542	4.5238	4.7917
	M	**4.6716**	**4.6439**	**4.6709**	**4.6695**	**4.6294**	**4.7711**
log(R&D Investment/Sales Revenue)	U	1.7147	1.9812	1.9089	1.987	0.50185	0.84762
	M	**1.8006**	**1.8482**	**1.9088**	**1.8958**	**0.82141**	**0.92939**

Chapter 4

Strategic R&D Projects Choice of the firm: Theory and Evidence from Korea

4.1 Introduction

Government grants play an important role in promoting firms' innovative activities. Many empirical studies have been conducted showing how government support affects firms' innovation output, such as the number of patents or the introduction of new products. However, there is little theoretical approach to describe the strategic choice of firms that implement innovation projects using government subsidies. Existing discussions about the innovative behavior of firms have taken into account differences in the size of firms and financial constraints. Cohen and Klepper (1996a) analyze the strategic innovation activities of large and small firms under the assumption that a large firm can better utilize the outcome of an innovation project because of its higher production capacity. Akcigit and Kerr (2015) explain using an endogenous growth model that exploitation Research and Development (hereafter R&D) activity increases more than exploration R&D as firm size increases. Mathesons et al. (2010) illustrate the possible R&D portfolio grid that firms can choose from. They use the probability of success and net present value to characterize four different types of R&D projects and investigate strategic management practices informed by the grid. Damsgaard et al. (2017) analyze the optimal project selection of an entrepreneur and an incumbent using game theory, assuming that success in innovation projects leads to a reduction in production costs, and that the incumbent faces a lower cost of market entry than the entrepreneur.

A number of empirical studies have been conducted on the impact of R&D support on innovation performance. Almus Czarnitzki (2012) investigate the effect of R&D subsidies in Eastern Germany and find a positive effect on innovative activity by using matching methods. Lach (2002) studies the effect of R&D grants policy in Israel using the Difference In Difference method. González, Jaumandreu, and Pazó (2005) study R&D policies in Spain and find a positive but small impact of public grants. Most recently, a regression discontinuity framework has been applied in estimating the policy effects. Jacob and Lefrgren (2010) study the effect of research grants on scientific productivity using a regression discontinuity design (RDD). Bronzini and Iachini (2014) also use a RDD to evaluate an Italian regional R&D subsidy program and find a positive effect that is limited to small firms. The effect of R&D incentives on innovation outcomes such as the number of patents, new products, and publication is also an important topic of previous literature. Branstetter and Sakakibara (2002) find that public grants have a positive effect on the patenting activity of Japanese firms. Bérubé and Mohnen (2009) find that Canadian firms benefiting from R&D tax credits and grants are more likely to develop new products.

This paper contributes to the existing discussion in several respects. First, we analyze the impact of a unique government R&D support system on firms' innovation activities. In South Korea's *R&D subsidy payback system*, firms use grants to carry out innovation projects. The government categorizes a firm with more than 300 employees as a large firm, and a firm with less than 300 employees as a small firm. When a firm that has implemented an innovation project using government grants tries to commercialize the project output, the firm must return a percentage of the subsidy received back to the government. In the case of large firms, 40% of the government subsidy must be returned and the small firm must return 10% of the subsidy. And the government tracks the innovation achievements commercialized by firms, assigning them into two categories: product innovation and process innovation. Here, product innovation refers to innovations that translates into new marketable (patentable) products, and process innovation means improvements in the efficiency of the production process.

We use the unique features of this program to theoretically predict the effects of R&D grant-related conditions on firms' project choice. In addition, this prediction is empirically verified through a regression discontinuity analysis using unique internal data. We compare the firms' strategic innovation activities in these grant schemes through the following steps. First, analyze two firms near the threshold of 300 employees - one with slightly more than 300, one with slightly less. We assume that other than the number of employees, these two firms have identical characteristics affecting innovation. Under these assumptions, the only difference between the two firms is the payback ratios for the subsidies they receive, as a result of having fewer or more than 300 employees. These two firms then choose an innovation project that maximizes their revenue, where the project is characterized by its success probability. A project with a low probability of success generates a large revenue at the time of success, and a project with a high probability of success generates a small revenue at the time of success. We analyze the two firms' project choices by solving maximization problem under different payback condition, The analysis shows that the firm subject to a high payback ratio select a riskier project compared to a firm that faces a low payback ratio.

Finally, we use a regression discontinuity design to estimate the effect of these conditions on the firms' choice of innovation activity using internal data from the South Korean government.

This paper proceeds as follows. Section 2 describes the *'Government grant pay-back system'* in Korea. Sections 3 analyze the optimal project choice for firms under different conditions. Section 4 analyzes the optimal project selection from the social planner's perspective and compares it to the market equilibrium. Section 6 empirically confirms the predictions of this theoretical analysis.

4.2 Institutional setting

South Korea has a unique rule for grants called the *'Government grant pay-back system'*. A firm that receives subsidies from the government to conduct R&D projects must return some of the subsidies to the government if they want to commercialize the project results upon completion. In the case of a large firm, defined as having more than 300 employees, 40% of the grant must be returned to the government. In the case of a small firm, defined by having less than 300 employees, 10% of the amount must be repaid. [1]

A small firm faces a relatively favorable tax rate compared to a large firm. Therefore it is possible that firms will artificially keep the number of employees below 300 to enjoy the tax benefits granted only to small firms. To prevent this from happening, the government allows a grace period of three years during which a small firm can continue to enjoy the tax benefits even after becoming a large firm. However, there is no such grace period in relation to the R&D grant.[2] As we will show below, this does not lead to manipulation around the threshold. Intuitively, the 3-year window of continued access to benefits other than subsidies makes it unlikely that a growing company will remain a small company by manipulating the number of employees, solely to seek preferential treatment under the R&D grant program.

The government keeps track of the type of innovation resulting from projects it has subsidized for a period of five years. The government distinguishes between two types of innovation: product innovation and process innovation. Product innovation is defined as technological development that can lead to new products. The government measures product innovation performance based on how many patents are filed and registered by a firm. Process innovation refers to innovations that improve processes, improve efficiency, or reduce costs. The government quantitatively and qualitatively assesses whether companies have reduced

[1] The number of employees is the main criterion that distinguishes between a large firm and a small firm. However, the capital owned by the firm is also considered for exceptional cases. For example, a firm employing more than 300, but with a capital of less than $8 million, would be classified as a small firm; a firm with the opposite characteristics could be classified as a large firm. However, over 95% of firms in the sample are classified either as a small firm or a large firm solely according to the number of employees.

[2] The main reason is that R&D grant paybacks are not classified as taxes and operate as a condition accompanying the contract.

costs or improved processes, and measures process innovation performance based on these evaluations. As noted by Mathesons et al. (2010), it is recognized that product innovations have low probability of success but high revenue generation potential via patenting, while process innovations have a low-risk, low-return profile.

4.3 Model

In this section, we examine the strategic choice of firm's innovation activities based on game theory within the payback system. Consider a market with two firms. We assume that the unobserved variables that affect the selection and performance of innovation projects of both firms are the same. The difference between the two firms is that one firm has slightly less than 300 employees, and the other has slightly more. As a result of this difference, the payback ratios applied to the firms are 10% and 40%, respectively. Under these assumptions, we set up a model that each firm maximizes expected revenue after repayment and analyze the optimal project choices of the firms.

4.3.1 Model Set-up

Each firm chooses from an infinite number of independent R&D projects. The outcomes of the chosen projects ultimately target the same market, and therefore the firm's choice of project can be viewed as different paths to bolstering their market positions. For example, one of the firms may choose a project to lower the cost of producing LED display panels while the other may choose to develop a new patentable display technology. Each project is characterized by the probability of success (p) of the project and the resulting expected revenue $pR(p)$, where $R^{'}(p) < 0$ and $p \in (0, 1)$.

The cost function to perform the project is given by $C(p)$, where $C'(p) < 0$. This is because the cost of an innovation project with a high probability of success is cheaper than that is not. The grant partially offset the cost, so the net cost is $C(p) - G$.

4.3.2 Optimal project choice under repayment

Therefore, a firm chooses a project that maximizes expected net revenue as follows

$$\underset{\{p\}}{Max}\ E(R) = p(R(p) - bG) - (C(p) - G) \tag{4.1}$$

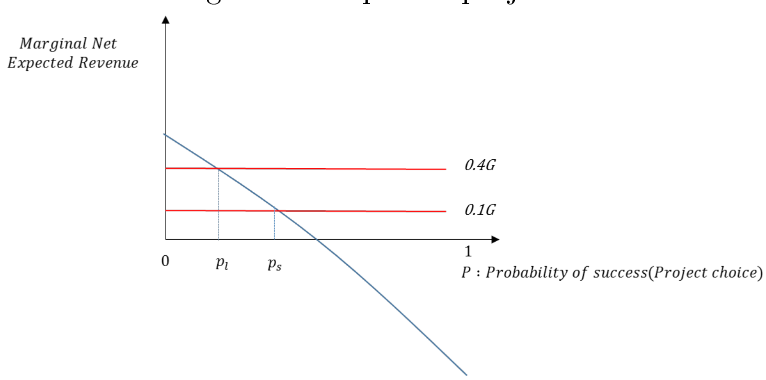
where $b = 0.4$ for large firms and 0.1 for small firms. The first-order condition for p is

$$dE(R)/dp = pR^{'} + R - C' - bG = 0 \tag{4.2}$$

and the second order condition requires $2R' + pR" - C" < 0$.

Differentiation of the f.o.c. then shows $dp/db < 0$, so that the large firm chooses a riskier project.

Figure 4.1: Optimal project choice



Based on this, we can see that the firm that repays 10% under the same conditions will choose a project that is much safer than the firm that repays 40%. And it can be predicted that these projects will bring incremental or procedural innovation, that is process innovation. Large firms, on the other hand, are expected to create a product innovation by selecting projects that are risky but generate high revenue.

Such results would run counter to the original intent of the pay-back program, which was to encourage small firms to innovative more by offering more favorable repayment terms. However, the analysis above shows that small firms would rather choose less risky projects.

In the next section, we look at the optimal project choice for the social planner and compare the result with that of the market equilibrium.

4.4 Empirical analysis

This section empirically verifies our theoretical prediction on the strategic behavior of firms using unique data. The rules of the program allow us to estimate the impacts of the different repayment conditions without having to consider compounded impacts that come from other unknown factors. Also, the evaluation of outcomes of projects by experts mitigates the problem of misclassification and missing values that can arise when using other proxies. Finally, the advantage of using both the number of product innovation and process innovation is that we can observe the firms' behavior under the condition of the grant.

4.4.1 Data description

The data used for empirical analysis is constructed by merging two data sets. The main data is the internal confidential data of the government, which is a list of all firms supported by the Ministry of Industry. This data includes general information such as the names and identifiers of all firms that were recipients of R&D subsidy programs from 2009 to 2013, the number of employees of each firm, the grant amount that the firms received, and whether the firm is a large or small firm. This data also contains information on how many patents have been filed and registered from the project, and how many process innovations were produced as outcomes.

The financial information of the firm comes from the credit rating company. The data has the most extensive financial information for companies in Korea and contains data on company sales, debt ratio, and their interal R&D investment.

To the best of our knowledge, we are the first to use rich internal data of this sort that tracks all firms that received a government R&D grant. The type of innovation that resulted from a government-sponsored project is tracked quantitatively, so we can estimate the effect of the unique terms of the grant on the type of innovation. Table 1 shows the summary statistics for small firms and large firms, respectively, grouped into two bins around the 300 employees threshold.

Table 4.1: Summary statistics

	Sample (200<<300)		Sample (300<<400)		
VARIABLES	Mean	S.D	Mean	S.D	Description
EMP	245.3	29.1	356.0	28.9	Number of Employees
GRANT	746.1	666.5	776.1	681.2	Grant amount
SALES	77483.6	55687.2	115138.1	66154.2	Revenues from the year before the government grant application
DEBT	86.3	49.6	89.9	50.3	Debt Ratio(%)
RDINVEST	4.1	7.9	3.2	6.2	R&D Intensity (Internal R&D Investment/Revenue)(%)
PRODUCT	1.5	2.4	4.4	1.8	Number of Product Innovation (Patent Granted)
PROCESS	3.4	1.9	1.3	2.0	Number of Process Innovation
N	594		336		

The sample is comprised of 5 years, which span from 2009 to 2013

Monetary units are in 1,000,000 Korean won, which is approximately equivalent to 1 thousand US dollars

4.4.2 Empirical Set-up

We test the treatment effect at the firm-size threshold (300 employees) by using a regression discontinuity design. We investigate the effects of different payback conditions on firms' choice of innovation activities as indicated by the product innovation and process innovation.

For identifying the treatment effect on firms' innovation outcomes and types, we prescribe the following model

$$Y_i = \alpha_1 + \gamma_1 Z_i + X_i + \varepsilon_i \tag{4.3}$$

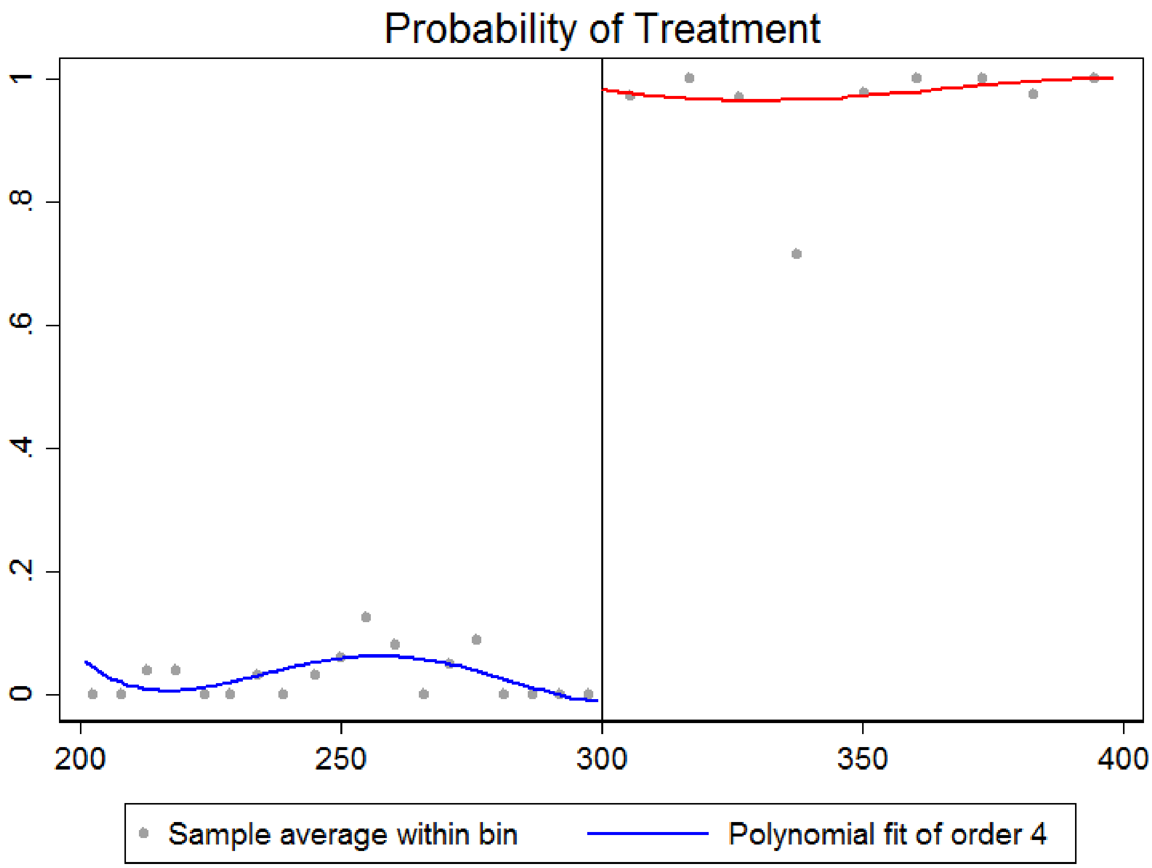
where, i indexes firms, Y_i is the innovation outcome measured by number of product and process innovations, X_i is the number of employees, and ε_i is the error term. The treatment variable $Z_i \in \{0,1\}$ is determined by the cut-off rule: $Z_i = I\ (X_i \geq 300)$ where I is the indicator function and 300 is the cut-off point. $Z_i = 0$ if X_i is less than 300 employees, i.e, the firm is small, and 1 otherwise.

Validity test

This section tests the validity of a standard RD design. In order for RD design to be a valid identification strategy, it is required that the following conditions are satisfied. First, the probability of treatment receipt around the cut off point should be discontinuous and its direction should be monotonic. Also, firms should not manipulate their running variables. And the distribution of the firm's other covariates should be smooth near the threshold. We test whether these conditions are met. Discontinuity conditions and monotonic treatment response conditions are shown in Figure. 4. When the number of employees exceeds 300, it can be discontinuously treated as a large firm.[3]

[3] As explained in the previous section, treatment is determined by the number of employees, but the amount of the firm's capital is also taken into account. Therefore, some firms have more than 300 employees but are classified as small firms and vice versa. However, as can seen from the figure, the treatment assignment rule applies to about 95% of the observations, so we use the sharp RDD to estimate the treatment effect.

Figure 4.2: Probability of being treated as a large firm against the number of employee

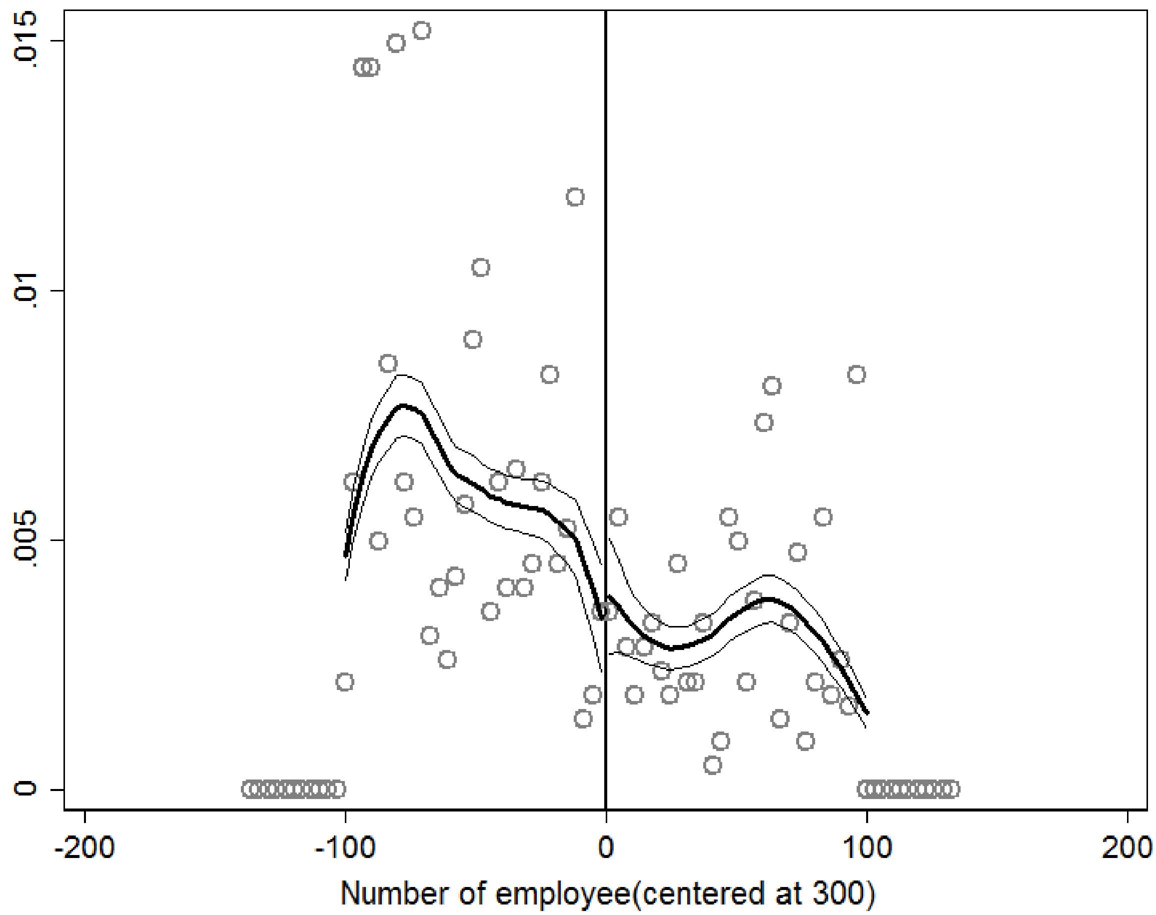


We test the validity of the running variable by examining the density of a running variable (number of employees) around threshold that is used by McCray (2008) to test the manipulation of running variables. Figure 5 shows non-significant discontinuities. The test gives a log difference in density heights at the cut off is 0.0014 with the standard error of 0.0021. This estimates shows that the difference in the density of a running variable around the threshold is not significantly different from zero. The McCrary test shows that firms do not manipulate their employment to enjoy benefit from the program. This estimation satisfies a necessary condition for valid set-up of the model.

These results are in line with institutional characteristics. As mentioned earlier, a possible reason for manipulation is the jump in the repayment rate applied to firms that cross the

300 employee threshold. However, even for small firms growing to more than 300 employees, other institutional benefits such as tax rates remain for three years. This institutional feature is likely to render the returns to manipulation, for the sole purpose of enjoying a lower repayment rate, at best marginal.

Figure 4.3: Empirical density of running variables(number of employee)



We also test the smoothness of the conditional means of other covariates that can affect innovation outcomes. The results in Table 2 show that firm covariates do not change statistically significantly according to treatment. This suggests that covariates, which can affect the innovation performance of firms near the threshold, are similar, thus indicating that treatment can identify the effect of treatment on innovation outcome by reducing bias. Appendix 2 again confirms this result by graphically showing that covariates do not show discontinuity due to treatment at the threshold. These validity test results show that the environment of our empirical analysis satisfies the required assumptions of an RD design.

Table 4.2: RD validity tests-Falsification test for covariates

	RD effects of treatment on covariates			
	(1)	(2)	(3)	(4)
VARIABLES	Grant	Sales	Debt	RD intensity
RD_Estimate	-159.4	-4227.7	9.498	0.652
	(121.9)	(9880.4)	(9.446)	(0.857)

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1
The Calonico, Cattaneo and Titiunik (2014a)
bias-corrected estimates along with standard errors

4.4.3 Estimation results

In this section, we estimate the treatment effect we want to identify through equation (20). To this end, we have applied the method of Regression Discontinuity (RD) point estimators with robust bias-corrected confidence intervals and inference procedures developed in Calonico, Cattaneo and Titiunik (2014a), and Calonico, Cattaneo, Farrell and Titiunik (2018). In other words, we estimate the optimal bandwidth by minimizing the bias according to the method proposed by them, and apply it to estimate the average outcome difference of the bandwidth as treatment effects. We validate the robustness of the results through an estimation using the treatment variable only and then with all other covariates.

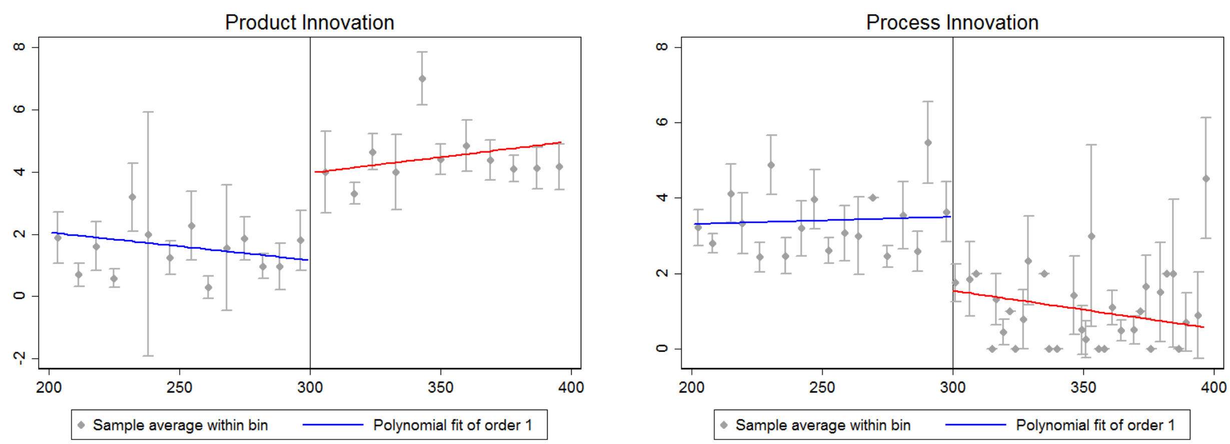
We then examine how product innovation and process innovation change discretely around the threshold. Table 3 shows the main estimation results. Our main results are presented in columns 1 and 3. To confirm the robustness of the estimation results, we use an estimation method of Calonico, Cattaneo, Farrell and Titiunik (2018) to perform an estimation including covariates that can affect the outcome, and the results are shown in columns 2 and 4. From the estimation results, we can see that firms subject to a high payback ratio will create about 2.1 to 2.5 product innovations, and conversely, about 1.6 to 1.8 less process

Table 4.3: Treatment effects on Product and Process innovation

	(1)	(2)	(3)	(4)
		(with Covariates)		(with Covariates)
VARIABLES	Product Innovation		Process Innovation	
Treatment effects	2.152***	2.577***	-1.650***	-1.829***
	(0.698)	(0.840)	(0.424)	(0.350)

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1
The Calonico, Cattaneo and Titiunik (2014a)
bias-corrected estimates along with standard errors in Column 1 and 3
The Calonico, Cattaneo, Farrell and Titiunik (2018)
bias-corrected estimates wit covariates along with standard errors in Column 2 and 4

Figure 4.4: Impact on Product Innovation and Process Innovation



innovations. This can be interpreted as the result of the strategic project choice of the firms analyzed theoretically in the previous section.

We illustrate graphically the estimation results in Figure 5. It is apparent from the figure that innovation outcomes at the threshold jump discontinuously.

These results are consistent with the firm's strategic innovation behavior, which was predicted based on game theory in the previous section. In other words, we can identify the strategic innovation behavior of firms driven by the difference in the cost of commercializing innovation outcomes, all else equal.

4.5 Conclusion

Many studies have focused on R&D policies or the financial conditions faced by firms that affect innovation performance. However, R&D policies may not only affect quantitative performance, but also firm's choice of innovation strategy or behavior. Based on the quasi-experimental case provided by the Korean government's unique institution, we identify the theoretical implications and carry out an empirical analysis to determine the effects of firms' institutional conditions on their choice of innovative activities. This study confirms another factor affecting firm's strategic innovation activities, and it can be seen that in the design of an R&D support system, implications on not just quantitative performance of innovation activity but also the type of innovation performance must be considered.

References

1. Akcigit, U. and Kerr W.R. (2015). “Growth through heterogeneous innovations”, PIER Working Paper No. 15-020

2. Bérubé, Charles, and Pierre Mohnen (2009) “Are Firms that Receive R&D Subsidies More Innovative?” Canadian Journal of Economics 42 (1): 206–225.

3. Bozeman, B., & Rogers, J. (2001). Strategic management of government-sponsored R&D portfolios. Environment and Planning C: Government and Policy, 19(3), 413-442.

4. Branstetter, Lee, G., and Mariko Sakakibara. 2002. "When Do Research Consortia Work Well and Why? Evidence from Japanese Panel Data ." American Economic Review, 92 (1): 143-159.

5. Bronzini, R. and Iachini E. (2014) “Are incentives for R&D effective? Evidence from a regression discontinuity approach.” American Economic Journal: Economic Policy, 6(4) 100–134.

6. Bronzini, R. and Piselli, P. (2014) “The Impact of R&D Subsidies on Firm Innovation” Bank of Italy Economic working papers.

7. Cameron, G. (1996) “On the measurement of real R&D: Divisia price indices for UK business enterprise R&D.” Research Evaluation, 6(3) 215-219.

8. Calonico, S., M.D. Cattaneo and R.Titiunik, (2014), “Robust Nonparametric Confidence Intervals for Regression-Discontinuity Designs,” Econometrica, 82(6), 2295–2326.

9. Calonico, S., Cattaneo, M. D., Farrell, M. H., & Titiunik, R. (2018). Regression discontinuitydesigns using covariates. Working paper, University of Michigan.

10. Cohen, W., Klepper, S. (1996) “Firm size and the nature of innovation within industries: the case of process and product R&D.” Review of Economics and Statistics, pp. 232–243

11. Färnstrand Damsgaard, E. , Hjertstrand, P. , Norbäck, P. , Persson, L. and Vasconcelos, H. (2017), “Why Entrepreneurs Choose Risky R&D Projects – But Still Not Risky Enough.” Econ J, 127: F164-F199.

12. Gonzalez, X., Jaumandreu, J., Pazo, C. (2005) “Barriers to innovation ´ and subsidy effectivenes” The Rand Journal of Economics 36 (4), 930–950.

13. Jacob, Brian and Lars Lefgren (2010) “The impact of research grant funding on scientific productivity”, Journal of Public Economics, 95(9-10), 1168-1177.

14. Jaffe, A. and Lerner, J. (2004) Innovation and its Discontents, Princeton: Princeton University Press.

15. Jaffe, A. and Le, T. (2015) “The Impact of an R&D subsidy on innovation: A study of New Zealand firms” NBER Working Paper 21479.

16. Lach, Saul (2002) “Do R&D subsidies stimulate or displace private R&D? Evidence from Israel” Journal of Industrial Economics, 50, 369-90.

17. Matheson D. & Matheson J. E. (1998) The Smart Organization: Creating Value Through Strategic R&D. Harvard University Press: Cambridge, MA.CH10

18. Matthias Almus & Dirk Czarnitzki (2012) The Effects of Public R&D Subsidies on Firms’ Innovation Activities, Journal of Business & Economic Statistics, 21:2, 226-236

Appendix-1

Figure 4.5: Conditional mean of Covariates