EXAM BASED TRACKING VERSUS DISTRICT BASED MIXING: IMPLICATIONS ON HOUSEHOLD SORTING AND INTERGENERATIONAL EDUCATION MOBILITY*

Yong Suk Lee**
Brown University

Mar 21, 2011

Abstract

In societies where parental income and ability are not highly correlated due to institutional or historical reasons, an education system that better matches high ability students to quality education regardless of income level would better align individual skills with the needs of the overall economy. I examine whether the institutional change from an exam based high school entrance policy ("tracking") to a district based lottery allocation mechanism ("mixing") affects intergenerational education mobility. A simple model predicts that, under mixing, the gradient on household income relative to one's ability becomes larger in the achievement production function due to households sorting across districts. I empirically test these hypotheses utilizing data from Korea where the transition from tracking to mixing occurred during the 1970s. Using a nationally representative sample, I find that the difference in difference estimates on father's education relative to one's middle school grade increases under mixing. This effect is strongest in Seoul and the larger cities which have multiple school districts. I then directly test for household location sorting within Seoul utilizing a boundary discontinuity design on the change of housing and commercial land prices. The relocation of top-tier high schools in Seoul during this period helps the identification and I find a discrete increase of 50% in the change in housing land prices for the top school district.

Keywords: Secondary Education, Tracking, Mixing, Intergenerational Mobility, Sorting, Boundary Discontinuity Design

JEL Codes: I20, J62, O15, R23

---

* I am grateful to Vernon Henderson and Andrew Foster for advice and guidance. I thank Abhijit Banerjee, Nathaniel Baum-Snow, Kenneth Chay, Stephen Gibbons, Adrienne Lucas, Isaac Mbiti, Sriniketh Nagavarapu, David Weil and participants at the Northeast Universities Development Economics Consortium Conference, Urban Economics Association Meeting, Bureau for Research and Economic Analysis of Development Summer School, Brown University's Applied Micro Lunch and Urban Lunch for valuable comments and suggestions. All errors are mine.
**Department of Economics, Brown University, 64 Waterman St. Providence RI,02912:Yong_Suk_Lee@brown.edu

I. Introduction

As many developing countries achieve universal primary education, governments and international organizations have focused on expanding secondary education. Policy debates on the development of secondary education include, among many topics, expanding access, financing, improving quality and relevance (World Bank, 2005). Nonetheless, an important aspect of secondary education that has received less attention is how students should be matched to schools when the quality of secondary education varies across schools. Secondary education that better matches high ability students to quality education regardless of income level would better align individual skills with the needs of the overall economy, especially, in societies where parental income and ability are not highly correlated due to institutional or historical reasons, e.g., class structure, race bias, gender bias, war or ethnic conflict.

This paper examines two major student allocation mechanisms for high school. One is a competitive exam based high school entrance policy ("tracking") where the school administers an entrance exam and chooses its students based on test results. The other is a district based allocation mechanism ("mixing") where a household's choice of school is directly based on its residential location. The transition from tracking to mixing is not uncommon. The UK and Scandinavian countries which used to track students shifted to comprehensive education during the second half of the 20th century (Maurin and McNally 2008). South Korea, which this paper examines, also shifted from tracking to mixing in the 1970s. More recently, urban secondary schools in China which used to administer exams have recently shifted to district based mixing. On the other hand, top-tier public secondary schools in Kenya continue to administer exams while admission to less prestigious schools is location based (Lucas and Mbiti 2010). Whether or not to track students is a highly debated and politicized policy but as developing countries expand compulsory secondary education it is likely that more countries will follow the precedents of shifting from tracking to mixing in secondary education. In this paper, I ask how the change in the institutional arrangement from tracking to mixing affects intergenerational mobility by examining the impact of one's father's education level and own ability on one's high school achievement. To empirically examine this question, I utilize a set of unique events from South Korea during the 1970s where the government changed the high school admissions policy from tracking to mixing across several cities over multiple years as well as the relocation of top-tier high schools in Seoul.

Figure 1 provides a first look at the evolution of intergenerational education mobility by plotting the coefficients in the regression of one's college score on father's education by high school entrance year. Coefficients on cohorts that enter high school before the mid 1970s hover around 0.05 and then steadily increase reaching above 0.15 for the late 1990s cohort. The coefficients imply that a 10% increase in father's education increased one's college entrance exam score by around 5% for the early cohorts but the effect rises more than threefold to above 15% for the later cohorts. What is noticeable is that this rise starts in the mid 1970s which is when the high school student allocation mechanism shifted from tracking to mixing. The policy reform took place in Seoul in 1974 and then rolled out to other cities until 1980. This reform to mixing had a considerable impact on households since entering a prestigious high school had been decisive in entering a top tier college. The main focus of this study is to examine whether the change in education policy from tracking to mixing caused the increase in intergenerational education persistence as depicted in Figure 1.

To help with the empirical estimation, I first construct a simple model of household educational choice and achievement production. The theory predicts that under mixing, the gradient on household income relative to one's ability becomes larger in the achievement production function due to households sorting across districts. Also, housing prices would differentially increase in better school districts as wealthier households bid up the price to move there. In the empirical analysis, I test two main hypotheses. I first test whether father's education, which proxies for income, becomes relatively more important in the reduced form estimation of achievement when the allocation mechanism changes to mixing. Next, I test whether the housing price increases more in the better school districts because of people sorting under the mixing regime.

I utilize the regional staggering and timing of the policy and employ difference in difference methods to empirically show that the shift to district based mixing increases the coefficient on father's education in the achievement production. The effects are strongest in Seoul and the largest cities but weaker in smaller cities. The smaller cities are less likely to be affected because they do not have multiple school districts and people are more likely to have pre-sorted in smaller geographic areas. Also, the interaction based analyses indicate that the policy effect is strongest after 4 to 5 years into the reform. This suggests that location sorting occurs as a response to the change in policy but it takes several years to move to a new equilibrium.

In the second part of the empirical analysis I directly test for location sorting within Seoul utilizing a boundary discontinuity design on the change of housing and commercial land prices. A unique event occurred in Seoul soon after the shift to mixing. Several of the most prestigious high schools were exogenously relocated from the old city center to the periphery area in order to reduce central city congestion. This incident helps the identification and statistical estimation of location sorting and I find a discrete increase of about 50% in the change in housing land prices in the top school district.

This paper provides evidence that district based secondary school allocation generates location sorting and decreases intergenerational education mobility compared to a competitive entrance exam system. There have been studies that examine tracking within schools and how it affects individual achievement and teacher incentives (Duflo et al. 2008). However, this paper is different in that it focuses on tracking at the school level and examines how that can affect society level mobility and sorting. Such studies have been done in the European context (Maurin and McNally 2008) but I believe this is the first study to examine this issue in the context more relevant for developing countries. Discussions in the economics of education literature focus on expanding household choice either through vouchers, charter schools, or mobility programs (Epple et al. 1998, Hoxby 2000, Katz 2000, Ladd 2002). However, the choice for schools to choose students, the core characteristic of a tracking regime is often left out in the educational choice debate. This paper suggests that tracking at the school level may better match ability to quality education in economies where income and ability are not highly correlated.

The paper proceeds as follows. Section II presents a model of household educational choice and simulates economies to draw out prediction and set up frameworks for empirical work. Section III examines the background on the historical and institutional context of the two education regimes in Korea and then discusses the data I use to estimate achievement. Section IV discusses the identification strategy to estimate changes in bias due to the different matching between students and schools and Section V shows the reduced form results. Next I shift to household location sorting in Seoul. Section VI discusses the policy background in Seoul and the identification strategy. Section VII provides the empirical results on sorting utilizing a border discontinuity design. Section VII concludes.

II. Theoretical Examination

It is not straightforward which regime would lead to higher intergenerational mobility. Under tracking wealthier households can provide private tutoring to their children so they can enter the better high schools. On the other hand, under mixing wealthier household can pay higher housing prices and move to better school districts. It could simply be that the margin of competition shifts from private tutoring to housing prices without much effect on intergenerational mobility. Then again, one channel may lead household income to have a stronger affect on achievement. In order to understand how households respond to the regime shift and to obtain a better sense of what to measure in the empirical analysis, I first examine a simple model of household preferences and achievement production.

II.1 A Model of Household Sorting under Tracking and Mixing

II.1.1 Household preferences

Consider a simplified world where there are N neighborhoods with each neighborhood having a high school of quality θ_H. Quality is teaching or facility aspects that affect student achievement. I further assume that quality can be ordered so that there is a ranking of high schools. Under tracking students take entrance exams and apply to the school of choice[1] and schools choose students based on the high school entrance exam results. Once the regime shifts from tracking to mixing, neighborhoods become school districts. Each district has the same number of houses and each household consumes 1 unit of housing and pays a housing cost (rent) of *r* depending on the district.

Each household has one adult and one child and is identified by its *(y,a)*-type, where *y* denotes household income and *a* is the child's ability. The household's utility function, *U(·)*, increases with numeraire consumption *c* and the educational achievement of the child *t*, and it is continuous and twice differentiable in both variables. Achievement $t=t(x,a,\theta)$, is a continuous and increasing function of child's ability *a*, school quality θ, and amount of private supplementary education, e.g., tutoring $x\geq 0$. School quality is assumed to be exogenous in the model. In Korea's context, school quality at the change of regimes was more or less historically determined with older establishments being more high quality. Though school quality can change under the mixing regime as a result of sorting, at the onset of policy change it is fair to assume school quality θ to be constant[2]. Also, I abstract away from peer effects as peer quality directly maps with school quality in a more complex version of this model and does not change the main implications. A model with peer effects is included in the Appendix [3].

The budget constraint faced by a *(y,a)*-type household is $c + px + r = y + \delta a$,where wealth depends not only on present income, *y*, but also some discounted factor, δ, of child's ability. Having some factor of ability in the budget constraint implies (1) that parents implicitly know that they will be compensated by their children when they grow up and (2) that families can extend credit utilizing its network of families, relatives and friends when it does not have access to formal credit institutions. These two assumptions are not unusual in developing economies. Banerjee (2002) discusses the various channels of how altruism, intergenerational contracting and credit constraints affect investments in human capital. Munshi and Rosenzweig (2009) show that families in rural India utilize jati networks to smooth consumption when there is a dearth of formal credit markets. I do not model intergenerational contracting explicitly but instead simplify the exposition by allowing household wealth to include a factor of the child's ability. This allows me to concisely convey the main points when comparing equilibrium properties between the two education regimes.

Maximization of $U(c, t(x,a,\theta))$ subject to the budget constraint yields the indirect utility function:

---

[1] It could be the other way around. Students apply first and take exams later. This does not matter if we assume everyone goes to the other school if one is not accepted by their first choice.

[2] Even if school quality changed the change would happen gradually. In the empirics, I examine policy effects by year in addition to difference in difference estimates to examine this possibility. Also, the theory would predict that over time all schools at the high quality schools get better as more able and higher income households move in to the district. The schools in the low quality district would become worse in a closed city model or ambiguous in an open city model.

[3] My initial model included school's location decision and showed that school quality to be a one to one function of peer quality. I provide in the discussion section how peer effects would affect the model's predictions.

(1) $$V(p,r,\theta;y,a)=U(y+\delta a-r-px(y+\delta a-r,a,\theta,p),t(x(y+\delta a-r,a,\theta,p),a,\theta)),$$

where $x(y+\delta a-r,a,\theta,p)$ is the demand for tutoring by household *(y, a)*.

Similar to Epple and Pratt(1998), I use a fundamental property and two additional properties of the *(y,a)*-type household's indifference curve in the *(r, θ)* plane.

Property 1- Increasing bid-rent: $$\left.\frac{dr}{d\theta}\right|_{V=\bar{V}}=-\frac{\partial V/\partial\theta}{\partial V/\partial r}=\frac{u_2}{u_1}\frac{\partial t}{\partial\theta}>0.$$

Property 2 - Single-crossing in income: $$\frac{\partial r_\theta}{\partial y}>0.$$

Property 3 - Single-crossing in ability: $$\frac{\partial r_\theta}{\partial a}>0.$$

Property 1 is a fundamental property that characterizes the bid rend function in terms of school quality *θ*. This implies that the indifference curve in the *(r, θ)* plane slopes up and illustrates the natural feature that people are willing to pay higher rents for better schools. Property 2 and 3 are additional assumptions. Property 2 states that all else equal, higher income households are willing to pay more for school quality. Likewise due to Property 3, all else equal, higher ability households are willing to pay more for school quality. Figure 2 depicts the *(y,a)*-type household's indifference curve in the *(r, θ)* plane and its single crossing properties.

II.1.2 Preferences of schools under the two regimes

I do not explicitly model the school's objective function but discuss the main points relevant for understanding how students are matched to schools under the two regimes. Under the tracking regime, each high school administers an exam and admits their students based on the exam results. Schools care about reputation, either for prestige or alumni support, and aim to obtain the brightest students. However, the test score on the high school entrance exam does not necessarily reflect one's true ability, *a*. Instead, the high school entrance exam score is also a product of one's ability and household resources expended towards tutoring in middle school. Middle school quality is assumed to be homogenous across schools since this is after the middle school equalization policy. In reality middle school quality would vary but it is safe to assume that the quality variation was much smaller than high schools after the elite middle schools were shut down. Hence, one's test score in the high school entrance exam can be expressed as $t_m=t_m(x_m, a)$ where x_m is the amount of tutoring invested during middle school. Even though schools want the higher ability students they eventually select students both on the margins of ability and income due to tutoring effects in the test score production. The parameters in this production function would determine how relevant each components are in the selection process.

On the other hand, under district based mixing, schools cannot choose students but simply admit those who reside in its district. The choices schools have are quite limited because they are not even allowed to charge differential tuition. Even the private schools are subsidized by the government and are not allowed to charge higher tuition than the public schools. This has been the case even under the tracking and regime and persists even till today. Schools, especially historically prestigious ones, may still care about reputation and in the long run relocate to different districts but we abstract away from such behavior in the model. The important point under mixing is that there are still differential qualities across schools and that households can access those schools by living in those districts.

II.1.3 Equilibrium under the two regimes

Equilibrium under tracking

First I define an equilibrium under tracking: an allocation of students across schools such that each student attends one and only one school is an equilibrium under tracking if and only if (i) all households maximize $V(p,r,\theta;y,a)$ and (ii) student allocation clears in all schools. Since there are no school districts and students apply to schools and enter where they are accepted, there is no need for households to move (assuming no transportation cost or spillovers from the better quality school) and no one pays any premium for being in the neighborhood close to school H. Since the schools know their quality levels and know that any *(y-a)*-type household prefers better quality, a perfect *sorting by observed ability* occurs. That is, students who score higher in the high school entrance exam, are matched with higher quality schools. Since slots are limited in each school, there is a cut off score for each high school and the cut off increases with school quality in equilibrium.

Equilibrium under mixing

Once again a definition: An allocation of households across districts such that each household lives in one and only one district is an equilibrium under mixing if an only if (i) all households choose their residence to maximize $V(p,r,\theta;y,a)$, and (ii) housing market clears in all districts. Now there are school districts and schools do not have any power to choose students in this regime. School quality is determined historically so everyone knows which district is better in terms of school quality. Because students are randomly allocated to the school within their district, a household's location choice determines the school its child will attend. Everyone would want to live in the better school district but the trade off of moving to the better district is that one has to pay a higher housing price, *r*. Hence, households reshuffle across districts with each household willing to pay some higher rent to live in the better school districts. An equilibrium outcome unique to the mixing regime is the existence of a housing price premium on quality, r_θ, which also serves as evidence of location sorting under mixing later in the empirics. I next describe how r_θ is determined under the mixing equilibrium.

For clarity, I illustrate the case where there are only two districts, districts H and L, and two schools with school quality θ_H and θ_L ($\theta_H > \theta_L$). Every household depending on its income and ability, *(y,a),* has a willingness to pay for the higher school quality, θ_H, so that it is indifferent between living in the two districts. If a household can pay less than its willingness to pay for θ_H and live in district *H* it is better off and will move to district H. Each household's willingness to pay for θ_H can be ordered and the household with a higher willingness to pay will overcut the next type's willingness to pay by epsilon and live in district *H*. Because of the fixed number of houses in district H and market clearing, there is a marginal household who will be able to live in the last available house in district H. Its willingness pay determines the market rent at district *H*. This marginal household's boundary indifference condition,

(2) $$V(y,a;\theta_L,r_L) = V(y,a;\theta_H,\tilde{r}),$$

not only determines the rent premium $\tilde{r}$ for district H but also the locus of *(y,a)*-types that are indifferent between the two districts. Any household above the locus sorts into district H and those below remains in district L.

II.2 Simulation and Graphical Illustrations of Equilibrium under Tracking and Mixing

In this section I simulate equilibrium properties with five schools each comprising a school district. I draw 500 households from a joint normal distribution with a correlation of 0.3 in the (ability, income) space. Consider a simple Cobb-Douglas utility function, *u(c,t)=c·t*, and the achievement production function, $t = t(a,x,\theta) = (x+k)^{\alpha} a^{\beta} \theta^{\gamma}$. All households have the same base level of home input, *k*, and can choose the corner solution of no tutoring, *x*=0. Household solves:

(3) $$\max_{x} (y + \delta a - r - px)(x+k)^{\alpha} a^{\beta} \theta^{\gamma}, \quad s.t.\ x \geq 0.$$

Households will choose to tutor (*x>0*) if household wealth is large enough, i.e., if

(4) $$y + \delta a > \alpha^{-1} pk + r$$

and choose $x^* = \frac{\alpha}{p(1+\alpha)}(y+\delta a - r) - \frac{1}{1+\alpha}k$ and attain

(5) $$t^* = \left(\frac{\alpha}{p(1+\alpha)}(y+\delta a - r) - \frac{\alpha}{1+\alpha}k \right)^{\alpha} a^{\beta} \theta^{\gamma}.$$

Otherwise, $x^* = 0$ and $t^* = k^{\alpha} a^{\beta} \theta^{\gamma}$.

Under tracking everyone pays the same housing cost r and students are sorted into high schools based on middle school achievement t_m which would be analogous to (5) with potentially slightly different parameters (α_m, β_m). Recall that middle school quality θ_m is constant. Test score on the high school entrance exam, which determines the matching between students and schools under tracking, is primarily determined by (α_m, β_m) and to a lesser degree δ. Figure 3A illustrates the sorting of households into schools when $\alpha_m = 0.2, \beta_m = 0.6$ and $\delta = 0.5$. Note that the gradient is steep in ability implying that even with tutoring ability is the prime determinant of achievement. The shaded area in Panel C shows which households choose to tutor in high school based on (4) and Panel E depicts the iso-achievement lines in a tracking equilibrium. The gradient between middle school achievement in Panel A and high school achievement in Panel E are similar because the parameters of test score production in high school (α, β) are the same for that of middle school. Trying different parameter values create small kinks without affecting the steep gradient on ability in the (ability, income) space.

Under mixing each district's marginal household's willingness to pay for higher quality school determines the sorting process. Sorting starts from the lowest quality district and households move up as higher (y,a) households outbid the marginal household for higher school quality. This process continues until the housing market clears. Based on the above functional forms each household's willingness to pay to move from a district with (r_L, θ_L) to a district with θ_H is

(6) $$r_H = (y + \delta a + pk)\left(1 - \left(\frac{\theta_L}{\theta_H}\right)^{\gamma/1+\alpha}\right) + r_L\left(\frac{\theta_L}{\theta_H}\right)^{\gamma/1+\alpha}$$

when households decide to tutor even with the higher rent in district H. The market clearing process in the 5 districts results in households sorting as in Figure 3 Panel B. The gradient is entirely determined by δ in this case. Also, households may prefer to move to the district H and receive the higher quality education but no longer choose to tutor. Panel D shows this equilibrium result where within each district those slightly above the marginal household choose not to tutor but households with higher wealth nonetheless choose tutoring in the better school districts. I provides more detail on sorting with tutoring choice decisions in the Appendix. Panel F illustrates what the iso-achievement looks post high school. Since household competition on housing price is mostly dominated by income, higher income and low ability households are able to benefit from higher school quality and thus obtain higher achievement. The gradient on income is now much larger than before in the (ability, income) space.

The achievement results in Figure 3 Panels E and F are obviously highly dependent on parameter values. However, any reasonable or even extreme parameter specifications still support the result that mixing results in a higher income gradient than tracking. Figure 4 Panels A and B put the same weight of 0.4 on α and β and twice as large a value on δ. The gradient under tracking becomes less steeper on ability and the gradient under mixing becomes steeper. Nonetheless, income is still a bigger factor under mixing compared to tracking. In Panels C and D more extreme and unlikely parameter values are used. α is now 0.6 and β 0.2, which implies that achievement has less to do with ability and much more to do with household earnings. Also, I assign δ=2 so that household wealth is determined more by child's ability than own earnings. Obviously, this is an unlikely scenario even with perfect intergenerational contracting and no credit constraints. The gradients are now quite similar but still the gradient on income is relatively steeper under mixing. These two alternative illustrations support the idea that income plays a larger role in child's achievement under mixing. Finally, Panels E and F shows results with the parameter values as in Figure 3 but now with a higher price on tutoring. This is actually a more realistic depiction in the sense that less than 10% of high school students received tutoring during my sample period of 1970-1985. Since, tutoring plays almost no role in this scenario, ability dominates sorting and achievement under tracking where as income dominates under mixing.

II.3 Predictions from the Model

The above model and simulations generate the following predictions on location sorting, tutoring, and student achievement. First on location sorting: Under tracking the sorting that occurs is a matching of high school entrance test scores to high quality schools. There is no location sorting or rent premium for higher quality schools. However, under mixing households sort into better school districts by income and ability and a rent premium arises in the better districts. Second on tutoring: Tutoring choice is determined by household wealth. The gradient on income and ability do not change between tracking and mixing because school quality per se does not affect tutoring choice at both the external and internal margin. Tutoring level may change as some households who move to the better school districts by paying higher rent may consider it optimal not to expend on private tutoring. Lastly, on achievement: As illustrated in Figures 3 and 4, the gradient on income in the achievement production function becomes relatively steeper when the regime shifts from tracking to mixing. The main reason for such change is the strong sorting towards higher quality schools based on income. Based on these prediction, I test in the empirical sections (1) whether the gradient in the reduced form achievement production changes, (2) whether the gradient on tutoring choice changes, and (3) whether a rent premium arises in the better school districts as the institutional setup changes from tracking to mixing.

The above equilibrium properties and predictions rely heavily on the assumption that the initial distribution of households is random over space. That is, in order for households to sort across districts when the policy shifts to mixing households should have not been pre-sorted. However, in reality cities are formed over a long period of time and the places where affluent people congregate over generations are likely to have historic and prestigious high schools. If households were already pre-sorted, further sorting to the better school districts would be minimal when the policy shifts from tracking to mixing. I can not observe the degrees of pre-sorting in each city. However, places with more school districts and a wider distribution of income and ability dispersed over a larger area will likely see further household sorting even if some pre-sorting had been there prior to the policy change. Hence, I perform separate estimations on cities of different sizes to see if there is evidence that household sorting and thus the decrease in intergenerational mobility is more evident in larger cities.

III. The Education Regime Shift and the Data

III.1 Background: Tracking to Mixing

After the Japanese Occupation ended in 1945 there was a surge in demand for education in Korea. This surge was due to the combination of pent up demand that had been contained under colonial rule and the traditional belief that education is the most important means for prestige and success. Most jobs that Koreans consider prestigious continue to be determined by competitive exams that emphasize education and testing. Even through the tumultuous periods of the Korean War people desperately sought out education and by 1959 elementary school entrance rate reached 96%. The large pool of elementary school graduates combined with the limited number of secondary schools made admission to middle school, which had been determined by competitive exams, more competitive. As a result, by the mid 1960s, wealthier families were seeking out private tutoring to supplement their children's exam preparation for top-tier middle schools. Excessive exam pressure on young 6th graders as well as the financial burden of private tutoring on middle class families became a large social problem. The military government which took over in 1961 understood these general concerns and as a means to quell public anger and gain popularity, drastically implemented the Middle School No Exam Policy (MSNEP) which abolished middle school entrance exams and aimed to normalize all middle schools in terms of quality. Starting in Seoul and Busan in 1969, middle school became accessible to all elementary graduates. By 1971, middle school entrance exams were abolished around the whole country. What made this policy so effective in achieving its goal of "normalizing" middle school education was the literal closing down of the most prestigious middle schools in Seoul. The schools that parents aspired to send their kids disappeared in a matter of years and hence the need for cut throat competition disappeared.

Once middle school entrance became exam free, the problem of over-competition and private tutoring simply shifted to high school entrance which was still exam based. Students applied to schools of their choice, took exams offered at that school, and the school admitted students with the highest grades. This system naturally generated a "tracked" system of high schools determined by prestige and history. Once again, the competition and tutoring among middle school students became a social problem. To deal with this similar situation, the government announced the High School Equalization Policy (HSEP) with the goal of eliminating competitive high school entrance exams and equalizing the quality of high school education in 1973. However, unlike the middle school policy, HSEP was more gradual in its implementation across regions and fundamentally the elite high schools remained as is, still generating a non uniform distribution of high school quality. What made MSNEP so effective and universal were its nationwide implementation and the closing down of elite middle schools. However, this did not happen with HSEP. What HSEP did was form high school districts and assign students randomly to schools within the district based on residence, hence "mixing". Since the number of high schools was limited, the high school entrance pool was regulated by a newly administered high school qualifying exam. This was a test that simply identified the new pool of high school students. Only those who passed the test would be eligible for high school education but the school would be randomly assigned. As of today over 70% of all high school students in Korea are under the mixing regime. The regions that still maintain tracking are more or less smaller urban areas with not many schools.

Appendix Table 1 describes the role out of HSEP by region and year. The shift to mixing occurred first in Seoul and Busan in 1974 (Group 1), then the three other large cities Daegu, Incheon, Gwangju in 1975 (Group 2), and then to the 9 provincial capitals in 1979 (Group 3), and then to other various regioinal cities in 1980 (Group 4). I denote the rest of the regions that did not see this shift to mixing as Group 5. In reality, several cities in Group 5 shifted to mixing after 2000. However, starting in the 1980s elite Special Purpose Schools that administered their own competitive exam were being established and by 2000 had become a large part of the general education. Hence, the comparison between tracking and mixing becomes less clear after 1985. Therefore, I focus on the policy change that occur during the mid to late 1970s and analyze individuals who entered high school between 1970 and 1985. The number of observations that enter high school before 1970 in the data drops considerably in my sample and people who entered high school before 1970 would have attended middle school before the MSNEP, being subject to differential middle school quality. Since the focus of this paper is the shift from mixing to tracking when prior education is not tracked, I focus on post 1970 entrants.[4]

Another institutional detail concerns private schools. Private schools supply a large portion of secondary education in Korea but they are regulated under the central government and operate in the same manner as public high schools. Private schools do not have the autonomy to charge their own tuition and are subsidized by the government. Also, during this period they could not set their own admission rules. Wealthy landlords established private schools as a means to maintain their estates during the land reforms enacted by the central government soon after the Japanese Occupation in 1945. Hence, the distinction between public and private high school is in general negligible in Korea, especially during this period.

[4] People who enter high school in 1970 and 71 in Seoul and Busan were likely under a tracked middle school system. I could drop these people from the observation and do robustness checks later. For this version those are also included. Nonetheless including those observations would work against finding an effect. Hence, the estimates I report here are likely to be lower bounds and will likely become stronger if I drop those observations.

III.2 Individual Level Data

In this section I discuss the individual level data used to estimate achievement and tutoring choices. Data used for the border discontinuity design is discussed in Section VI. I obtain detailed individual education history and family background information using the Korea Labor and Income Panel Survey (KLIPS). KLIPS is an annual panel survey of around 5000 households and 11000 individuals. The base survey is conducted annually and asks various questions on labor market and income dynamics of households and individuals. The first wave was conducted in 1998 and continues to be administered since then. There are special supplementary surveys conducted each year focusing on a specific topic. For instance, health and retirement in the 4^{th} wave, labor unions and relations in the 8^{th} wave, etc. The 11^{th} wave which was conducted in 2008 had a supplemental education survey, which asks detailed questions on the educational history and retrospective school life since middle school. I focus on the set of people who answer the education supplemental survey, and since my study examines the effect of secondary school policies on achievement, I restrict my sample to those who went to high school between 1970 and 1985. Summary statistics for this sample is presented in Table 1 Panel A. There are 2,456 individuals in the sample ranging from people born in 1952 to 1972. KLIPS collects data on the names, location, enrollment and graduation years of the middle school, high school, and college. Also, it identifies whether the high school was a general high school or vocational high school and whether the college was a four year university or a two year technical college. KLIPS does not have individual college entrance exam test scores and moreover the college entrance exam changed several times over the course of this period.

Since I have no data on individual test scores, I construct achievement measures utilizing the average entrance exam test scores by college. I utilize two sources that provide such information. One is an article in the daily newspaper Joongang-Ilbo which provides 1976 average test scores by college. The other is a college application reference book that provides average scores for 1994. Since the exams were different for the two periods, I normalize each test score to a 100 scale and then take the average of the two years to get a score for each college listed in the two sources. For all other schools not listed in the two sources, I categorize each school (e.g., 4 year college in Seoul, regional 2 year technical college, etc.), and assign an average score based on the information in the two sources. Appendix Table 2 and Appendix Figure 1 describe in detail how the scores were constructed and graphs its distribution. I match each individuals college to the scores I construct above. The mean college score for the sample is 47.3 and ranges from 41 to 81. Those who do not enter college are assigned a score of 41. Also, based on the ranking of the scores, I generate dummy variables, i.e., whether one attended a top 3, 10, or 20 college to later check how robust the findings are with different achievement measures.

KLIPS asks enrollment year and location of one's high school. I link this information with the High School Equalization Policy by region and year to identify each individual's exposure to tracking or mixing. Another valuable aspect of the KLIPS supplemental survey is that it asks one's achievement level in middle school. Specifically, it asks everyone to report one's middle school Korean, English, math grade in a 1 to 5 scale with 5 being the highest. I add all grades to construct an index ranging from 3 to 15 and use this as a proxy for pre high school ability. The sample mean is 10.5 with a standard deviation of 2.7. Controlling pre high-school ability is essential in extending empirics out from the theory. Though the self reporting feature may be prone to bias, this measure none the less provides valuable information of one's ability prior to high school.

IV. Identification Strategy for Achievement

IV.1 Estimating Biased Coefficients without Further Bias

Though theory is based on a structural equation of achievement production, the empirics will estimate a reduced form equilibrium outcome on household income and ability. This is because intergenerational mobility or persistence is inherently a reduced form notion and the quality of secondary education one receives, an unobserved component, is non-randomly matched to students. Furthermore, the location sorting that arises with the shift to mixing changes the mechanism of how school quality is matched to students and thus results in the change in reduced form estimation of intergenerational mobility. In other words, the reduced form estimates are biased because income and ability are correlated with the quality of one's high school education, which in general is unobserved. Specifically, the theory points to the following structural equation of achievement for individual *i* attending school *s*:

$$y_{is} = c + \beta x_i + \gamma a_i + \pi \theta_{is} + \varepsilon_{is} \tag{7}$$

where x is log household income, a is log ability, θ is log school quality, ε is white noise, and y is a log of one's achievement.

There are several points to note concerning the right hand side variables. First, since θ is not random and not observed in general, θ is the source of bias in both the tracking and mixing regimes and the degree of bias changes when households non-randomly sort across district borders. Second, household income x is noisily measured due to several reasons. Some of the concerns raised in the literature are (1) surveys report occupation but not income, (2) father's income can be inferred but not the income from other household members, and (3) data only allows for short term income measures based on one or a few entries. Some solutions to this problem in the mobility literature have been proposed, such as, to use father's education as an instrument for father's income, utilize father-son pairs with multiple panel entries (Solon 1992), and to utilize individual tax data rather than surveys (Saez et al. 2010). The data I use suffers from all these problems and in particular is problematic in inferring father's income based on occupation because the individuals in the panel span over a 20 year period during which income by occupation will have changed considerably.[5] Hence, I use father's years of education rather than some income measure for x. The education measure is mostly noise free and is also a good measure of income and hence has been used to directly measure intergenerational education mobility (Kremer 1997). Lastly, ability, a, is not measured in general. In particular, I need each individual's pre-high school ability not some ability measure that is inferred from post-high school outcomes like earnings to test the implications of the model. Fortunately, the KLIPS supplemental education survey asks respondents to answer his or her middle school grade in three subjects in a 1 to 5 scale. This provides me with a proxy, though noisy, of pre-high school ability. However, the measurement error associated with this variable is likely to be a source of bias. If people's report on this variable is biased toward their achievement, in the sense that, those who eventually went to a good college or got a good job report higher than their true middle school grade and vice versa, then γ will be biased upwards. On the other hand, there could be classical measurement error resulting in attenuation bias. Nonetheless, I include this self-reported middle school grade as a proxy variable for one's pre-high school ability in the empirical analyses with such caveats in mind.[6] [7]

Another layer of complexity comes from the fact that the dependent variable that proxies achievement is not a true measure of one's performance but an average test score measure inferred from which college one attends. In other words, the dependent variable in essence is a fixed rank measure that is confined between 41 and 81. Fixing the range of the LHS variable can be problematic if the variance of the true outcome changes over time. A mechanical increase in the variance of the RHS variables leading to an increase in the variance of true outcome will return smaller coefficient estimates when the range of the proxy dependent variable is fixed. Hence, the change in one coefficient over time captures not only a policy affect but also the change in the variance in achievement resulting from the variability in the factors of achievement. Fortunately, comparing ratios, which is the focus of this analysis, circumvents this problem. Appendix C clarifies this point.

IV.2 Sources of Variation and Model Specification

The goal of estimation is to find biased coefficients where the only source of bias is the non-random matching of students to different quality high schools. In order to control for all other potential

[5] The problem is that I do not have income by occupation for all years. Individuals report father's occupation when they were 14. Hence for those who enter high school between 1970 and 1985 are reporting father's occupation during the periods of around 1968 and 1983. Linking 2000 income to these periods is likely to be misleading.
[6] Estimating the coefficient on father's education without including ability in the reduced form estimate would also provide a intuitive measure of intergenerational mobility. However, as I discuss next and in Appendix C, examining only one coefficient can be problematic when the proxy dependent variable is fixed in range but the distribution of true outcome varies by year.
[7] Comparing coefficient ratios help cancel out some the biases associated with using a proxy instead of true ability. Appendix B provides an illustration.

sources of endogeneity, I include city, entrance year, and age fixed effects in the reduced form estimation. The assumption needed for identification is that the fixed effects capture all other sources of biases. City fixed effects will capture city specific variations in the education environment, such as, the number of colleges and admission size as well as the administrative capacity of the city education system. Also, it can capture city specific externalities or city specific peer effects that affect individual achievement, such as, the occupational composition or education composition of household heads. Year fixed effects capture overall time trends and age fixed effects allow for the fact that the high school entrance age varies from 14 years old to 20 years old in the data.

With fixed effects included, the underlying achievement of individual *i* who went to high school in city *j* in year *m* at age *k* can be determined as:

(8) $$y_{ijkm} = \beta \; x_{ijkm} + \gamma \; a_{ijkm} + \mu_j + \eta_k + \lambda_m + \nu_{jm} + (\theta_{is} + u_{ijkm})$$

where x is log of father's education, a is log of middle school grade, μ_j, η_k, λ_m each denote the city, age, and high school entrance year fixed effects. These fixed effects can be correlated with household income and/or own ability. Also, the error term is composed of the school quality component θ_{is} and a random shock component, so that $\varepsilon_{ijkm} = \theta_{is} + u_{ijkm}$. If equation (8) captures the true association of achievement production so that there are no other endogenous variables other than that is captured by θ_{is}, estimating the following reduced form equations

(9) $$y_{ijkm} = c^T + \beta^T x_{ijkm} + \gamma^T a_{ijkm} + \mu_j^T + \eta_k + \lambda_m + \varepsilon_{ijkm} \quad \text{under tracking and}$$

(10) $$y_{ijkm} = c^M + \beta^M x_{ijkm} + \gamma^M a_{ijkm} + \mu_j^M + \eta_k + \lambda_m + \varepsilon_{ijkm} \quad \text{under mixing.}$$

will return the correctly biased set of coefficients $(\hat{\beta}, \hat{\gamma})$. Note that I allow the city fixed effects to be different between the two regimes. The shift from tracking to mixing implies that households living in different cities could simply move to Seoul and reside in the good school districts in hope of getting a better education. In other words, the shift to mixing could have generated inter-city migration of a non-random set of households. Particularly, wealthier households from rural counties could have moved to cities with better schools altering the city specific peer effects. I capture the potential inter-city migration effect by allowing the city fixed effects to change under the different regimes. Appendix D provides more detail on changing city fixed effects.

In addition to estimating achievement, I use the same framework to predict whether one receives tutoring in high school based on (4). I can't measure how much tutoring one receives but can measure the external margin of tutoring choice and perform both linear probability models and logit models.

As additional controls, I include gender as a base covariate and in some cases tutoring choice in middle school and whether one attends a general high school. Tutoring choice in middle school and attending a general high school are also choice variables but such choice are likely to be determined by *x* and *a*. Hence, including these two variables would not further raise endogeneity concerns but may reduce bias at the expense of precision.

Before turning to the main empirical analysis I first examine patterns from the raw data. Figure 5 show scatter plots of students by achievement in the (middle school grade, father's education) space by education regime. The solid circles are individuals with a college score of 72 and hollow circles are those with a score of 50. One can notice that the distribution is more rectangular under tracking for both types of circles. This suggests that middle school grade played a larger role in determining future outcome under tracking relative to mixing. Figure 6 simply fits linear lines based on the scatter plot. I first regress achievement on father's education and middle school grade with no further controls for the sample with a score above 50 or 72 and then plot iso-achievement lines using the coefficient estimates. The dashed lines are for tracking and solid lines for mixing. The slopes are clearly steeper for the dashed lines. In the following section I test whether the difference in gradients maintains under more rigorous analyses and if the difference is statistically significant.

V. Empirical Results on Achievement and Tutoring Choice

V.1 Results on Achievement

Pooling equations (9) and (10), in practice, I run the following regression

$$(11)\; y_{ijkm} = c + \beta_1 x_{ijkm} + \gamma_1 a_{ijkm} + \beta_2 x_{ijkm} T_{jm} + \gamma_2 a_{ijkm} T_{jm} + \delta T_{jm} + \mu_j + \mu_{2j} T_{jm} + \eta_k + \lambda_m + \varepsilon_{ijkm}$$

where y_{ijkm} is the log average college test score of individual *i* who went to high school in city *j* in year *m* at age *k*. T_{jm} is a dummy indicating whether the individual belongs to the "mixing" regime. I am primarily interested in whether $\frac{\beta_1 + \beta_2}{\gamma_1 + \gamma_2} > \frac{\beta_1}{\gamma_1}$. Column (2) provides estimates of equation (11) and column (1) provides estimates without changing city fixed effects. β_2 is positive and significant in both specifications. The effect of log father's education on one's log college score more than doubles when the regime shifts to mixing. Holding one's ability constant a 10% increase in one's father's education increases one's college score by 2.6% under tracking and over 7% under mixing. The interaction term on one's middle school grade γ_2 is also positive although the relative increase is much smaller. The bottom panel in Table 2 provides the main empirical test of whether the gradient on father's education relative to ability increases with mixing, i.e., it estimates the non-linear specification $t = \hat{\beta}_{mix}\hat{\gamma}_{track} - \hat{\gamma}_{mix}\hat{\beta}_{track}$

and reports the p-value of the one sided hypothesis test $H_0 : \frac{\beta_{Traking}}{\gamma_{Tracking}} = \frac{\beta_{Mixing}}{\gamma_{Mixing}}$ and $H_1 : \frac{\beta_{Traking}}{\gamma_{Tracking}} < \frac{\beta_{Mixing}}{\gamma_{Mixing}}$ by testing t>0. The test statistic is around 0.008 or 0.009 and the one sided test is statistically significant at the 1% level. Including whether one received tutoring in middle school in column (3) does not do much to any of the coefficients, which is supportive of the idea that tutoring in middle school is a redundant variable with prior tutoring decision being made with variables already included in the equation. Including general high school choice (GHS) in column (4) reduces the coefficients' magnitude a little, which is indicative of GHS capturing other omitted variables that were not captured in equation (11). Nonetheless the p-value of the hypothesis test is very small and virtually the same as before. Column (5) reduces the sample to students who enter general high school students only. The test statistic is the same at 0.0072 but standard errors increase a bit returning a p-value of 0.06. Including GHS as a control variable would be problematic if GHS choice changes the sample in terms of father's education and middle school grade pre and post policy change. Table 4 columns (5) and (6) report linear probability estimates of GHS choice. I test whether the gradients change between tracking and mixing. The two sided hypothesis test reports extremely high p-values at almost 0.9. Based on these results I keep GHS as a control variable in following specifications as well.[8] Appendix Table 4 performs triple difference estimation using gender or high school status as an additional layer of variation. Estimates using vocational high school students as the control group for general high school students within the treatment group continues to return a statistically significant difference in the gradients.
Using the coefficients in column (4), I find that $\frac{\beta_1}{\gamma_1} = 0.08 < 0.24 = \frac{\beta_1 + \beta_2}{\gamma_1 + \gamma_2}$ which is a three fold increase. Estimates using other columns return similar increases.

Results from Table 2 support the hypothesis that father's education became more important relative to own ability in determining one's high school achievement under mixing. However, one might be concerned that this effect may be driven by only a few cities and sorting responses may be different across city size. This indeed could be the case. It is very likely that when the districts were set up at the onset of regime change, the district with the high quality school may already have been occupied by higher income families due to historical reasons. Suppose the prestigious high schools were schools with longer history that were located in the old city centers where there tends to be higher income households. Then we would expect to see less sorting after the shift to mixing. Hence, the magnitude of the

[8] Appendix Table 4 (not reported yet) performs triple difference estimation using gender or high school status as an additional layer of variation. Estimates using vocational high school students as the control group for general high school students within the treatment cities continue to return statistically significant differences in the gradient.

coefficients may differ across different cities based on the initial spatial context of each city. In short, the results should be highly sensitive to whether or not households were already pre-sorted and whether the city was spatially diverse enough for further sorting to occur after the policy change.

Therefore, I next examine the above by breaking the sample into cities of different size. The fact that the policy itself rolled out starting with the biggest cities and then to the smaller cities makes the stratified analysis more natural. I first do this exercise separately for Seoul and then to the cities that change policy in 1974 and 1975, the Group 1 and 2 cities, and lastly for cities that change policy in 1979 and 1980, the Group 3 and 4 cities. Table 3 provides the results for the city group specific DD results. As evident from the coefficient on the interaction term of father's education with MIX and the p-value of the hypothesis test, the regime shift impacted the larger Group1 and 2 cities more than the smaller Group 3 and 4 cities. Column (6) indicates that in the smaller sample of general high school students only, the one sided hypothesis test does not return a significant p-value in the 1979 and 80 cities. The impact of the policy change on the smaller cities are weak, suggesting the possibility that the historical distribution of households were already aligned so that more affluent households lived near the better schools. On the other hand, larger cities in Group 1 and 2 (Seoul, Busan, Daegu, Incheon, and Gwangju) which have multiple districts and more diverse households spread out over a larger area likely had more room for further sorting, and as a result saw a decrease in intergenerational education mobility when the policy changed to district based mixing. The ratio of coefficients for the 1974 and 75 cities changes from 0.06 to 0.23 in column (5) whereas for the 1979 and 80 cities in column (6) the change is from 0.08 to 0.14.

Interaction Terms Analysis

The above DD results confirm that the regime shift to mixing did change the equilibrium gradient of income relative to ability with stronger change coming from larger cities. In this subsection I examine how the gradient evolves by cohort and city groups. I estimate the following relationships which is the interaction term analog of the DD regression in equation (11) :

$$
\begin{aligned}
(12)\qquad y_{ijkm} = c &+ \sum_{l=-10}^{15}(x_{ijkm}\cdot d_l)\beta_{1l} + \sum_{l=-10}^{15}(a_{ijkm}\cdot d_l)\gamma_{1l} + \sum_{l=-10}^{15}(x_{ijkm}\cdot d_l\cdot T_i)\beta_{2l} \\
&+ \sum_{l=-10}^{15}(a_{ijkm}\cdot d_l\cdot T_i)\gamma_{2l} + \mu_j^{(T)} + \eta_k + \lambda_m + \varepsilon_{ijkm}
\end{aligned}
$$

where d_l is a dummy that indicates whether individual i is in the l th policy cohort into the mixing regime. I define each policy cohort by 2 year intervals because one year intervals provide cell sizes that are too small. The 1st policy cohort is defined by individuals either in the first or second cohort that experiences the policy change to mixing. Since, each city shifted to mixing in different years, the 1st policy cohort is composed of individuals who entered high school in Seoul in 1974 or 1975 as well as those who entered high school in Incheon, Daegu, Gwangju in 1975 or 1976, and so forth. A negative number implies that the individual was under the tracking regime. Changing fixed effects are used in all specifications. In order to perform the above interaction analysis I need to assign hypothetical policy cohorts for the control region which actually never see the policy being implemented. Thus, I hypothetically assume that for the Group 5 cities the policy change occurs in 1980 like the Group 4 cities. Appendix Table 5 provides the coefficient estimates. I will focus on Figure 7 which plots the evolution of the ratio of coefficients by policy cohorts. The solid line represents cities that shift to mixing and the dashed line represents the control group that maintain the tracking system. What is evident is that after a few years into mixing the solid line and dashed line decouple quite systematically. Even though the dashed lines are arbitrarily constructed using a hypothetical policy year, the graph does support the notion that intergenerational mobility decreased after 4 to 5 years of policy change.

I next extend the interaction analysis to the city level. The estimating equation is analogous to (12). The only difference is that instead of assigning arbitrary policy years for the control cities (Group 5), I can use exact high school entrance years because the cities in each group shift to mixing in the same year. Table 4 presents results for Seoul, the 1974 and 75 cities and the 1979 and 80 cities. The coefficients on column (1), (3) and (5) represent the treatment effect for the 2 year intervals and the coefficients on column (2), (4) and (6) are for the control group. The coefficients of Panel A in columns (1) and (3) become significant starting around 1978 or 79 years, about 4 years after the regime shift to mixing. On the other hand, we can't find any significant coefficients in column (5) post policy years. Figure 8 graphically represents the evolution of the coefficient ratios and confirms the findings in Table 4. Seoul separately and the Group 1 and 2 cities all exhibit a decoupling with the control group starting in 1978 or 1979, which is about 4 years after the policy change in 1974 or 1975. Panel B shows no persistent trends in the evolution for the coefficient ratios of Group 3 and 4 cities.

V.2 Results on Tutoring Choice

I next examine whether we find significant changes in which households decide to provide tutoring due to the regime shift to mixing. Equation (4) provides the theoretical starting point for examining this question. Tutoring choice is determined by present earnings and some factor of ability. Also higher rent lowers the likelihood to choose tutoring without changing weights on income and ability. On the other hand, school quality per se does not enter the tutoring decision. This suggests that the gradient on father's education relative to one's middle school grade should not change whether the regime is tracking or mixing. I estimate equation (11) with the binary dependent variable indicating tutoring in high school. Covariates are no longer in log form as implied by equation (4). Table 5 presents OLS results and Appendix Table 5 presents logit results. Due to the small cell size in many cities the logit specification drops many observations. Also, the non-linearity of logit makes the gradient comparison more difficult than OLS. Hence, I focus on results in Table 5 and supplement the discussion with the logit results in Appendix Table 5. Columns (1) and (2) are for all students and columns (3) and (4) are for general high school students only. None of the coefficients on the interaction terms are significant and the two-sided hypothesis test return extremely large p-values. Hence, we do not find evidence that the gradient on father's education and middle school grade changed in tutoring choice. The logit results in Appendix Table 5 also find the same results. None of the interaction terms are significant.

Table 6 then examines whether we find differential effects across different groups of cities. The two sided hypothesis test all have large p-values unable to reject the null that gradients did not change. However, there is some hint that the gradient on ability relative to income becomes steeper with the shift to mixing based on columns (4) and (5). However, none of the change in gradient is significant at the 10% level.

V.3 Robustness Checks

Instead of using the average exam score of one's college as a proxy for achievement, I use binary dependent variables indicating whether one enters a top 3, 10, or 20 college and run linear probability regressions with the same specification as Table 2 column (4). Table 7 columns (1) to (3) provide results depending on the number of top schools used. The coefficients on the interaction terms are all significant and the gradient on income relative to ability increases with the shift to mixing. The one sided hypothesis test is significant at the 10% level when the dependent variable is a top 10 or top 20 college.

When calculating average college exam score I arbitrarily assign a lower bound of 41 to those who do not go to college at all. However, there could have been other reasons (e.g., credit constraint, health etc.) other than achievement in high school that inhibited one from advancing to college. In column (4) of Table 7, I assign the highest college score of 87.9 to those who were initially assigned the lower bound of 41. Note that the coefficients on father's education and middle school grade are now both negative. Also, the gradient becomes steeper on father's education under mixing. This suggests that under tracking lower achievement was also associated with ability more and under mixing ability plays less of a role.

The last two columns of Table 7 attempts to control for school quality by including high school fixed effects. Note that the sample drops from 2267 to 2168. This is because I can not identify some of the high schools names or the high schools were outside of South Korea. I manually go through the list of high school names to clean the matches and provide unique codes to each high school. Recall that the shift in gradient occurred between the unobserved matching between household types and school quality. Therefore, once we control for high school fixed effects, the policy should have no impact on the ratio of coefficients. As expected, none of the interaction terms in columns (5) and (6) are significant and there is no statistically significant evidence that the ratio of coefficients increased post policy.

VI. Identification of Location Sorting Using a Boundary Discontinuity Design

VI.1 The Regime Shift in Seoul and Sources of Variation

The above reduced form results are consistent with the theoretical prediction that the shift to mixing causes households to sort across districts which in turn decreases intergenerational mobility. In this section, I directly test for location sorting by employing a border discontinuity design that examines the change in housing land prices across school districts in Seoul. I focus on Seoul for several reasons. First, the coefficient plots in the previous sections show that the increase is strongest for Seoul after the shift to mixing. Second, Seoul is the dominantly larger city which implies that I will have a large enough number of observations for neighborhood land prices in the border discontinuity design. Lastly and most importantly, as discussed previously, household sorting depends on the historical context of each city and Seoul is unique in the sense that schools actually moved soon after the regime change to mixing. The fact that top-tier schools relocated helps the identification and generates an exogenous event that was unrelated to the historical pre-sorting.

How the policy rolled out in Seoul was more than a simple shift from tracking to mixing. When the regime shifted to mixing in 1974, five school districts were created based on administrative districts. In general, two to five administrative districts formed one school district. However, to minimize the sudden transition from tracking to mixing the city created a Unified Central District (UCD) in the old city center. Hence, all the top-tier high schools were included in this district at the time of policy change. What was unique about the UCD was that anyone living in the city could apply to a high school in the UCD and could attend if he or she was selected through a lottery. If one was not selected then one would attend one's own district high school which was also determined through a lottery. Hence, there was not much incentives for households to move to the UCD, especially considering the fact that housing prices were already high in the city center.

On the other hand, in the October of 1972 the Education Minister announced that Gyeonggi High School, the ranking one school, would be relocated to the southeast part of Seoul, which later becomes school District 8. The announcement was initially met with strong disagreement by the alumni but eventually the relocation happened and the new campus opened south of the Han River to Gangnam-gu in 1976. Also, Hweemun High School, another top-tier high school, finished construction of its new campus in 1978 in District 8 and sold its old campus in the city center to Hyundai Group. Several other high schools followed suit. Figure 9 illustrates the movement of 3 of the top tier high schools between 1975 and 1980. Figure 9 also illustrates how the share of college educated adults increased between 1975 and

1980 in this region. All of the top-tier high schools were originally located in the historic city center but by 1990 only 4 out of the 10 schools were still in the city center. Appendix Table 3 provides the list of the top 10 high schools in Seoul and their location changes. What these relocations imply in terms of timing is that when the policy shifted from tracking to mixing in 1974, people already knew that the most prestigious high school and other top-tier high schools would relocate to District 8 in years to come. This exogenous relocation of high schools generated variation in school quality by districts and provided incentives to relocate even though the Unified Central District existed. The UCD was eventually abolished in 1980 and a full-fledged mixing policy would stay in place until the late 1980s when special purpose exam schools started to emerge. In the following boundary discontinuity analysis I utilize the relocation of these high schools to statistically examine sorting. The boundary of the Unified Central District and District 8 along with the neighborhoods(dong) of Seoul are depicted in Figure 10.

VI.2 Neighborhood Level Housing Land Price and Distance Data

The main prediction of the model in terms of location sorting is that the shift to mixing increases housing prices at the better school district. I use a border discontinuity design to find evidences of differential changes in housing prices along district boundaries. Since housing price data are not available uniformly across neighborhoods for the 1970s, I use dong-level (neighborhood area administrative level) housing and commercial land price appraisal data assessed by the Korea Appraisal Board. Representative lands for high, medium, and low quality housing and commercial building in each neighborhood area were assessed annually and created in reports. I copied the reports on Seoul for 1971 to 1979 and entered the data manually for all odd number years for this analysis. Also, in order to calculate distance I generate a center point for each neighborhood and calculate the distance to the borders of two districts I focus on, the Unified Central District and the 8th District, utilizing GIS tools. I use integer distance measures in kilometers rather than the calculated distance in GIS. That is, if the neighborhood center is less than 1km I assign a distance value of 1, and if the center is between 1 and 2 km I assign a distance of 2. I use this integer distance because the continuous distance is also an approximation that is dependent on the generated neighborhood center point. Table 1 Panel B provides the summary statistics of the main variables I use in the border discontinuity design. Note that I have 464 distance observations but the land price observations are smaller. This is because some the neighborhoods did not have housing or commercial buildings and the periphery regions of Seoul at this time period were not developed. A positive distance measure implies that the neighborhood is within the district of interest and zero identifies the border.

VI.3 The Boundary Discontinuity Design

Theory predicts that the school districts with high quality schools would see a rent premium once the policy shifts to mixing. In other words, we should see housing prices increase more in higher quality school districts after the policy change. A cross sectional comparison of prices between high and low quality school districts would incorrectly identifying school quality effects as many observable and unobservable attributes that affect housing price is also correlated with school quality. However, this unique set up in Seoul where the policy change from tracking to mixing is incidentally accompanied by the exogenous relocation of the top school generates the variations to identify the rent premium caused by wealthier households sorting towards better school districts. Hypothetically, a difference in difference framework where one compares housing price between the high quality school district to a low quality school district before and after the policy change could identify the policy affect on housing price. However, there are many things that affect housing prices especially in school districts where the geographic area is quite large and price differences within each district is extremely large. Focusing on the boundary, while mapping the geographic characteristics of each location into a distance measure, enables one to more precisely estimate the price change that arises from household sorting caused by the regime shift. The following provides a more formal discussion.

Suppose land price y_{dt} at location *d,* measured as distance from a boundary, in period *t* is determined by some smooth function $y_{dt} = f_t(X(d)_t) + \rho_t D_{d>0} + \varepsilon_{dt}$ where *X(d)* is a vector of *N* factors ($x^1(d)_t$ $x^2(d)_t$ $x^3(d)_t$ ... $x^N(d)$) that affect housing land price, i.e., housing quality, school quality, neighborhood amenity, etc. Each $x^n(d)$ is smooth in *d* and thus *f(X(d))* maps a multitude of factors into a one dimensional value that is smooth in *d*. *d* is denoted so that *d*=0 identifies the boundary of interest and $D_{d>0}$ is a dummy variable indicating the district of interest, i.e., the higher quality school district. Even at the cross section, e.g., *t*=1, some components of *X* may not be continuous across the boundary causing a discrete jump in prices. ρ_t captures this discrete jump around the boundary at time t. ε_{dt} is white noise. Now suppose the policy changes from tracking to mixing and affects household sorting in *t*=2. Even with none of the components in X changing, in t=2 we should see a discrete rent premium determined by the marginal household, and hence a jump in ρ_t. Allowing the functional form to be different in period 2 and taking the difference between the two periods we get

(13) $$y_{d2} - y_{d1} = f_2(X(d)_2) - f_1(X(d)_1) + (\rho_2 - \rho_1)D_{d>0} + \varepsilon_{d2} - \varepsilon_{d1}$$

Since, $f_2(X(d)_2) - f_1(X(d)_1)$ is a smooth function in d it can be fitted with a non-parametric or polynomial function *m(d)*. Rewriting (13) we get

(14) $$y_{d2} - y_{d1} = m(d) + \rho D_{d>0} + e_d \,.$$

where ρ identifies the rent premium generated by households sorting in the mixing regime. The exogenous change in school quality by the relocation of top-tier schools help identify this effect in case households were already pre-sorted into neighborhoods with good schools. The identification of ρ relies on the fact that *m(d)* is smooth along the borders. That is, no other factors other than school quality in ($x^1(d)_t$ $x^2(d)_t$ $x^3(d)_t$ ... $x^N(d)$) should generate any *additional* discrete jump between the two time periods of policy change.[9] In other words, no other factors that affect housing price should have changed discretely along the borders exactly when the education policy changes. One potential concern may be that this was a period of rapid development with many new construction projects. So a discrete difference in housing supply near the boundary could happen and hamper the identification. I discuss such possibility in the discussion section.

VII. Empirical Results on the Change in Housing Land Prices

In practice, I specify a polynomial function on distance and estimate the following regression:

(15) $$\begin{aligned} y_{i,t_2-t_1} = c \; &+ \beta_{01} d_i + \beta_{02} d_i^2 + \ldots + \beta_{0p} d_i^p \\ &+ \rho \; D_i + \beta_{11} D_i d_i + \beta_{12} D_i d_i^2 + \ldots + \beta_{1p} D_i d_i^p + \varepsilon_d \end{aligned}$$

where *y* is the change in log housing or commercial land price and D_i =1 if the neighborhood is inside the UCD or District 8 depending on the border being examined. Since all households were eligible to apply to schools in the UCD regardless of residential location, I expect no sorting or weak sorting across the UCD border. On the other hand, families were aware that several top-tier high schools would relocate into District 8. Hence, I expect to find a differential increase in housing land prices across the District 8 border when the shift to mixing occurs. Before running the above regression, I first visually inspect the changes in prices along the district borders to see if there are any apparent discontinuities. Also, the visual inspections help specify the degree of the polynomials to use in the regression.

Figure 11 Panel A depicts the change in the log of medium quality housing land price by distance to the District 8 border for 1971-1973, 1973-1975, 1975-1977, and 1977-1979. The scatter plots clearly show a drastic jump along the boundaries between 1973 and 75. I then fit 4th order polynomials using coefficient estimates from Table 8. The change in log prices between 1973 and 1975 exhibit the sharpest discontinuity which is evidence of people sorting when the regime shifted in 1974. In other words, households knew that District 8, which was less developed compared to the city center at that time, would receive top-tier high schools and started to sort into that district bidding up the housing prices once the

[9] Maybe should add a graphical illustration later

regime shifted to mixing.[10] I do the same plots for the UCD with distance measures originating at the UCD border. Figure 11 Panel B describes the change in housing land prices. I find no drastic jumps at the border which is consistent with the predictions.

Table 8 reports the regression results when I fit equation (15) with 4th order polynomials for housing land prices and 3rd order polynomials for commercial land prices. For the 1971-73 land prices I fit a quadratic because of lower observation counts. I retain the same functional form on each side of the border because there are not that many running variables once the data is inside the district of interest (the d>0 region). Fitting a separate polynomial on the positive distance side imposes too much curvature when fitting the data. The magnitudes and standard errors of ρ are reported in Table 8. I use robust standard errors that are clustered at the gu (larger administrative district) level. The results on District 8 indicate a significant and strong increase in the change in log prices between 1973 and 1975 which amounts to a 58%, 49%, and 59% increase for medium, high, and low quality housing land, respectively. The change in housing land price between 1975 and 1977 also shows a 29% increase. On the other hand, the 1975 to 1977 change in housing and commercial land price exhibit a significant decrease in the UCD. This is the period already into the mixing regime with Gyeonggi High School moving out to District 8 and many of the top-tier high schools in the UCD planning to move out or constructing new buildings in District 8. Hence, this drop can be associated with the drop or expected drop in school quality in the UCD. The incentive to sort into the UCD was much smaller in 1974 since anyone in the city could apply to the UCD high schools. On the other hand, once the top-tier schools started to move out there were incentives on part of the households to sort out of the UCD.

One potential problem with the above discontinuity design is the imposing of a certain polynomial function that may not describe the true association. The 4th order polynomial may be picking up something that is in fact continuous and vice versa. To overcome this problem I focusing on data points near the border rather than fitting a 4th order polynomials over the whole support. This helps identify the jump with less structure but at the cost of losing observations. For District 8, I trim the sample and focus on the data within 5km and 3km of the boundary. For the UCD, I focus on data within 5km and 2km of the boundary. The windows are shorter for the UCD because the UCD was much denser than District 8, as can be seen in Figure 10. Also, I allow the functional form to be either quadratic or linear and allow the coefficients to change across borders. Figure 12 illustrates the plots and fit the change in medium quality housing land and commercial land prices around the 5km window for District 8 and UCD. Table 9 reports regression results using the 5km band and Appendix Table 6 reports results for the 3km or

[10] I also plot and fit 3rd order polynomial for the change in medium quality commercial land price by distance to the 8th district border. There are not many commercial land data points in the 8th district (distances greater than zero) which is reflective of the less developed situation at that time. There is no evident jump for the 1973 to 75 change.

2km windows. The results are consistent with that of Table 8. The increase in the change in 1973-1975 housing land prices becomes stronger when I use the narrower window for District 8. The increase in medium quality housing land price is 73% as indicated in column (2) of Table 9. For the UCD I find a 22% decrease in the 1975-1977 change in housing land prices which is of similar magnitude as in Table 8.

In short, the change in housing land prices in the border discontinuity design show statistical evidence that districts that were exogenously receiving or losing top-tier high schools see a significant jump up or down in the change in housing land prices. The above results are consistent with the theoretical predictions on location sorting. I discuss alternative explanations in the final section of the paper.

VIII. Discussion

As more students graduate from primary schools and slots for secondary education are limited, allocating students and providing a decent secondary education are a becoming a greater challenge in developing countries. If the allocation mechanism affects intergenerational mobility as shown in this paper, the type of policy developing countries pursue will ultimately affect how its human resources are distributed throughout its economy. This paper's empirical results show that the policy change from exam based tracking to district based mixing increases intergenerational education persistence while generating household location sorting. However, these results are driven by household demand while assuming supply side factors, i.e., school quality and housing supply constant remain constant. I discuss each of these factors next.

I have argued that the driving forces of my results are the household side sorting on income and ability. However, schools could change quality because as a response to the reform. In reality schools in the better school districts became better as more able and high income households moved in. The difference in difference estimates assume that school quality remains the same. If school quality actually improves in the better districts then that would have generated more sorting and my estimates would be upward biased. First, it seems unlikely that school quality would change in such a short period. But none the less, even if there were further sorting due to the improvement of school quality in better school districts, the reduced form estimates is what we are interested in. My strategy in the empirical work was to identify estimates that include biases that arise from sorting. We are mainly interested in how the shift to mixing increases intergenerational persistence regardless of whether households sorting occurred from the first order policy changes or the second order sorting that arises as school quality improves along with peer quality in the better school districts.

Second, another supply side factor that may be relevant is the change in housing supply. In the case of Seoul, District 8, the district that had the top-tier schools relocated into, also experienced large increases in housing supply. In this case, the boundary discontinuity estimates are likely to be downward biased as the marginal household that determines the rent can come from a lower *(y,a)*-type then when housing supply is fixed.

Lastly, peer effect, an important factor for achievement production as well as sorting is not discussed in the paper. I intentionally dropped peer effect which was included in earlier versions of the model to simplify the discussion.[11] In the initial model school quality was in a one to one relation with peer level. Under tracking, ability would be highly correlated with average peer level. But immediately after the shift to mixing there should be low correlation. However, if households value peer levels they would start to sort under mixing with peer levels increasing in the better school districts and decreasing in the worse school districts. The addition of peer effects would further increase sorting but as discussed previously on changing school quality, we are interested in the reduce form estimates of intergenerational mobility that include any sorting due to peer effects. We are not interested per se in disentangling what portion of sorting is due to school quality or peer quality. The main interest is the empirical result that tracking provides more opportunities for high ability students coming from low income households to perform better and the confirmation of that this is because household sort across districts in order to obtain higher quality or higher peers under mixing.

One final aspect concerns external validity. The setting in South Korea during the 1970s is not that different from many developing nations today. The country was recovering from war and constantly under political turmoil with limited freedom. Hence, children ability and household income was not highly correlated compared to today. Also one of the government primary agenda was expanding and improving secondary education. An important institutional aspect was that all secondary schools, regardless of public or private, charged the same tuition to households. Hence school quality was not priced by tuition. This feature of quality not being priced directly is the main reason for sorting as households competed in housing price instead to access un-priced quality. However, in countries where private schools are more expensive and better quality, then the implications of an exam based entrance would be different. Tracking would not support the level of intergenerational mobility as in this paper. However, in countries where the elite schools are still public organizations without high tuition, tracking may still be a better option in nurturing talented students in the economy.

[11] Add back in Appendix.

References

Banerjee, Abhijit (2004), "Educational Policy and the Economics of the Family," *Journal of Development Economics*, (74)1 3-32.

Bayer, Patrick, Fernando Ferrerira and Robert McMillan (2007). "A Unified Framework for Measuring Preferences for Schools and Neighborhoods," *Journal of Political Economy*, 115(4) 588-638.

Card, David and David S. Lee (2008). "Regression Discontinuity Inference with Specification Error," *Journal of Econometrics*, 142(2) 655-674.

Chay, Kenneth Y., Michael Greenstone (2003). "Air Quality, Infant Mortality, and the Clean Air Act of 1970," *NBER Working Paper No. 10053*. October.

Chay, Kenneth Y., Jonathan Guryan and Bhashkar Mazumder (2009). "Birth Cohort and the Black-White Achievement Gap: The Roles of Access and Health Soon After Birth," *NBER Working Paper No. 15078*. June.

Chetty, Raj (2009), "Sufficient Statistic for Welfare Analysis: A Bridge between Structural and Reduced-form Methods", *Annual Review of Economics*, 1 451-488.

De Bartolome, Charles A.M. (1990). "Equilibrium and Inefficiency in a Community Model with Peer Group Effects," *Journal of Political Economy*, 98(1) 110-133.

Duflo, Esther (2002). "Schooling and Labor Market Consequences of School Construction in Indonesia: Evidence From An Unusual Policy Experiment," *American Economic Review*, 91(4) 795-813.

Duflo, Esther, Pascaline Dupas and Michael Kremer (2008). "Peer Effects and the Impact of Tracking: Evidence from a Randomized Evaluation in Kenya," *NBER Working Paper No. 14475*. November.

Duflo, Esther (2009). "Re-evaluating Learning," *NBER Reporter* 1.

Epple, Dennis and Glenn J. Platt (1998). "Equilibrium and Local Redistribution in an Urban Economy when Households Differ in both Preferences and Incomes," *Journal of Urban Economics* 43, 23-51.

Epple, Dennis and Richard E. Romano (1998). "Competition Between Private and Public Schools, Vouchers, and Peer-Group Effects," *American Economic Review*, 88(1) 33-62.

Imbens, Guido and Thomas Lemiux (2008). "Regression Discontinuity Designs: A Guide to Practice," *Journal of Econometrics*, 142(2) 615-635.

Henderson, Vernon, Peter Mieszkowski and Yvon Sauvageau, (1978) "Peer Group Effects and Educational Production Functions," *Journal of Public Economics*, 10 97-106.

Hoxby, Caroline M. (2000). "Does Competition Among Public Schools Benefit Students and Taxpayers?" *American Economic Review*, 90(5) 1209-1238.

Kang, Changhui, Cheolsung Park and Myoung-Jae Lee (2007). “Effects of Ability Mixing in High School on Adult Hood Earnings: Quasiexperimental Evidence from South Korea,” *Journal of Population Economics*, 20 269-297.

Katz, Lawrence, Jeffrey Kling and Jeffrey Liebman (2001), “Moving ot Opportunity in Boston: Early Results of a Randomized Mobility Experiment”, *Quarterly Journal of Economics*, 116 607-654.

Kim, Sunwoong and Ju-Ho Lee (2010). “Private Tutoring and Demand for Education in South Korea,” *Economic Development and Cultural Change*, 58(2) 259-296.

Kim, Taejong, Ju-Ho Lee and Young Lee (2008). “Mixing versus Sorting in Schooling: Evidence from the Equalization Policy in South Korea,” *Economics of Education Review* 27, 697-711.

Kopczuk, Wojciech, Emmanuel Saez and Jae Song (2010). “Earnings Inequality and Mobility in the United States: Evidence from Social Security Data Since 1937,” *Quarterly Journal of Economics*, 125(1) 91-128.

Kremer, Michael (1997). “How Much Does Sorting Increase Inequality?” *Quarterly Journal of Economics*, 112(1) 115-139.

Ladd, Helen F. (2002). “School Vouchers: A Critical View,” *Journal of Economic Perspectives*, 16(4) 3-24.

Lucas, Adrienne and Isaac Mbiti (2010), “Effects of School Quality on Student Achievement: Discontinuity Evidence from Kenya,” *mimeo*.

Maurin, Eric and Sandra McNally (2008), “The Consequences of Ability Tracking for Future Outcomes and Social Mobility” *mimeo*.

Munshi, Kaivan and Mark Rosenzweig (2009), “Why is Mobility in India so Low? Social Insurance, Inequality, and Growth,” NBER Working Paper No. 14850.

Rothstein, Jesse (2006), “Good Principals or Good Peers? Parental Valuation of School Characteristics,” *American Economic Review*, 96(4) 282-311.

Solon, Gary (1992). “Intergenerational Income Mobility in the United States,” *American Economic Review*, 82(3) 393-408.

Solon, Gary (2002). “Cross-Country Differences in Intergenerational Earnings Mobility,” *The Journal of Economic Perspectives*, 16(3) 59-66.

Urquiola, Miguel (2005). “Does School Choice Lead to Sorting? Evidence from Tiebout Variation,” *American Economic Review*, 95(4) 1310-1326.

Urquiola, Miguel and Eric Verhoogen (2009). “Class-Size Caps, Sorting and the Regression Discontinuity Design,” *American Economic Review*, 99(1) 179-215.

World Bank (2005), “Expanding Opportunities and Building Competencies for Young People: A New Agenda for Secondary Education” World Bank Policy Paper

Appendices

Appendix A. Model with peer effects
To be added.

Appendix B. Imperfect proxies and comparing ratios
To be added.

Appendix C. Dealing with omitted variable bias and scale bias
To clarify this point it is helpful to examine the above in a regression framework. In principal, I would like to estimate each individual's true achievement production using equation (16). In reality, I am mapping one's true achievement into bins representing one's post-secondary institution's average test score and then estimating that discrete and ordinal score measure. This can abstractly expressed as the following,

(17) $$s_{it} = Int[F(y_{it}) \cdot N] \cdot b + c \ ,$$

where *s* is the institution average test score measure and *N* is the number of bins, i.e., the number of average test scores I match individuals into, which in this case is 76. *F(y)* is the cumulative distribution function of achievement and *Int()* is a function that returns the integer dropping any decimals. *b* is a scale factor and *c* is a level factor that adjusts the value of *Int()* function so that it fits into the range of test scores, which is [41,87.9] in this case. In essence, what $Int[F(y_{it}) \cdot N]$ is doing is mapping individual *i*'s achievement observed in year *t* to an ordinal number ranging from 1 to 76. Then I rescale so that these number fits into the desired test score range of [41, 87.9]. Note, that I am assuming the same number of individuals in each bin and that each bin is of equal size. Though these are simplifications, this exercise still illustrates the main point which is to see how the variance in the right hand side variable can affect the coefficients when the dependent variable is an ordinal proxy of true outcome represented in a fixed range. If I further assume that F(y) follows a uniform distribution then (17) becomes

(18) $$s_{it} = Int[\frac{y_{it} - y_{\min,t}}{y_{\max,t} - y_{\min,t}} \cdot N] \cdot b + c \ .$$

Where the denominator, y_{max} - y_{min}, measures the range of achievement for that year. Since I have dense test score bins I can approximate the above in an OLS regression framework by taking out the integer function and estimating the following equation,

(19) $$s_{it} = \phi_t \cdot y_{it} + \eta_t$$

where $\phi_t = N \cdot b \Big/ y_{\max,t} - y_{\min,t}$ is a scale factor that decreases when the variability of achievement is larger. η_t is an additive time component. Note that any one change in the variability of father's education, own ability, or school quality over time will return different ϕ_t measures over time. Substituting (16) for y_{it} we get

(20) $$s_{it} = c_t + \beta\phi_t x_i + \gamma\phi_t a_i + u_{it}$$

where $u_{it} = \phi_t(\pi\theta_{it} + \varepsilon_{it})$. Hence, when I regress one's post-secondary institution's average test score on one's ability and father's education I am in reality estimating the coefficients *β* and *γ* scaled by some time varying factor ϕ_t. The estimated coefficients of equation (20) are

(21) $$\begin{bmatrix} \hat{\beta}\phi_{t\,OLS} \\ \hat{\gamma}\phi_{t\,OLS} \end{bmatrix} = \begin{bmatrix} \beta\phi_t \\ \gamma\phi_t \end{bmatrix} + \begin{bmatrix} \sigma_x^2 & \sigma_{xa} \\ \sigma_{xa} & \sigma_a^2 \end{bmatrix}^{-1} \begin{bmatrix} \sigma_{x\theta} \\ \sigma_{a\theta} \end{bmatrix} \phi_t = \begin{bmatrix} \beta\phi_t + \kappa\phi_t(\sigma_a^2\sigma_{x\theta} - \sigma_{xa}\sigma_{a\theta}) \\ \gamma\phi_t + \kappa\phi_t(\sigma_x^2\sigma_{a\theta} - \sigma_{xa}\sigma_{x\theta}) \end{bmatrix}$$

where $\kappa = \dfrac{1}{\sigma_x^2\sigma_a^2 - \sigma_{xa}^2}$. As equation (21) indicates there are two types of biases. The first is the omitted variables bias. *β* and *γ* are biased because *x* and *a* are correlated with *θ* and the degrees of correlation change based on whether the regime is tracking or mixing. Another source of bias comes from the scale factor ϕ_t. Because of the existence of both an omitted variables bias and a scale bias, it is difficult to predict the direction of biases for each coefficient.

Suppose we disregard the scale bias and focus on the omitted variable bias. Assuming that the marginal and joint distribution of income and ability do not change, i.e., σ_x, σ_a, and σ_{xa} are all positive and constant, then κ is also constant under both regimes. The theory indicates that under tracking there is no correlation between *x* and *θ* but a strong correlation between *a* and *θ*. So under tracking the omitted variable bias makes *γ* biased upward. Since *x* and *a* are positively correlated, *β* is downward biased. That is, under tracking $\sigma_{x\theta}^T = 0$ and

(22) $$\begin{bmatrix} \hat{\beta}\phi_T{}^T{}_{OLS} \\ \hat{\gamma}\phi_T{}^T{}_{OLS} \end{bmatrix} = \begin{bmatrix} \beta\phi_T - \kappa\phi_T(\sigma_{xa}\sigma_{a\theta}^T) \\ \gamma\phi_T + \kappa\phi_T(\sigma_x^2\sigma_{a\theta}^T) \end{bmatrix}.$$

Likewise, Under mixing the correlation between *x* and *θ* is positive and the correlation between *a* and *θ* is positive but weaker than before. The consequence is that the omitted variable bias term increases for *β* while it decreases for *γ*. That is, under mixing $\sigma_{x\theta}^M > 0,\ \ \sigma_{a\theta}^T > \sigma_{a\theta}^M$ and

(23) $$\begin{bmatrix} \hat{\beta}\phi_M{}^M{}_{OLS} \\ \hat{\gamma}\phi_M{}^M{}_{OLS} \end{bmatrix} = \begin{bmatrix} \beta\phi_M + \kappa\phi_M(\sigma_a^2\sigma_{x\theta}^M - \sigma_{xa}\sigma_{a\theta}^M) \\ \gamma\phi_M + \kappa\phi_M(\sigma_x^2\sigma_{a\theta}^M - \sigma_{xa}\sigma_{x\theta}^M) \end{bmatrix}.$$

However, the additional layer of bias generated by the scale factor ϕ_t may either scale down or up each coefficient estimate. Ultimately, we would not be able to disentangle the two effects and conclude which direction the overall bias works. In other words, we cannot directly compare the coefficients $\hat{\beta}^M$ and $\hat{\beta}^T$ or $\hat{\gamma}^M$ and $\hat{\gamma}^T$ because we cannot estimate the time dependent multipliers ϕ_M and ϕ_T separately. For instance, if $\phi_M > \phi_T$ then $\hat{\beta}^M > \hat{\beta}^T$ unconditionally but the inequality for *γ* can go either way. However, since the scale coefficients affect *β* and *γ* in the same manner we can still compare ratios of the coefficients, and unconditionally we should find that $\dfrac{\hat{\beta}^M}{\hat{\gamma}^M} > \dfrac{\hat{\beta}^T}{\hat{\gamma}^T}$. This ratio of coefficients is what the theory predicts and is the main interest in the following empirical examination of intergenerational education mobility.

Appendix D. Inter-city migration and changing fixed effects

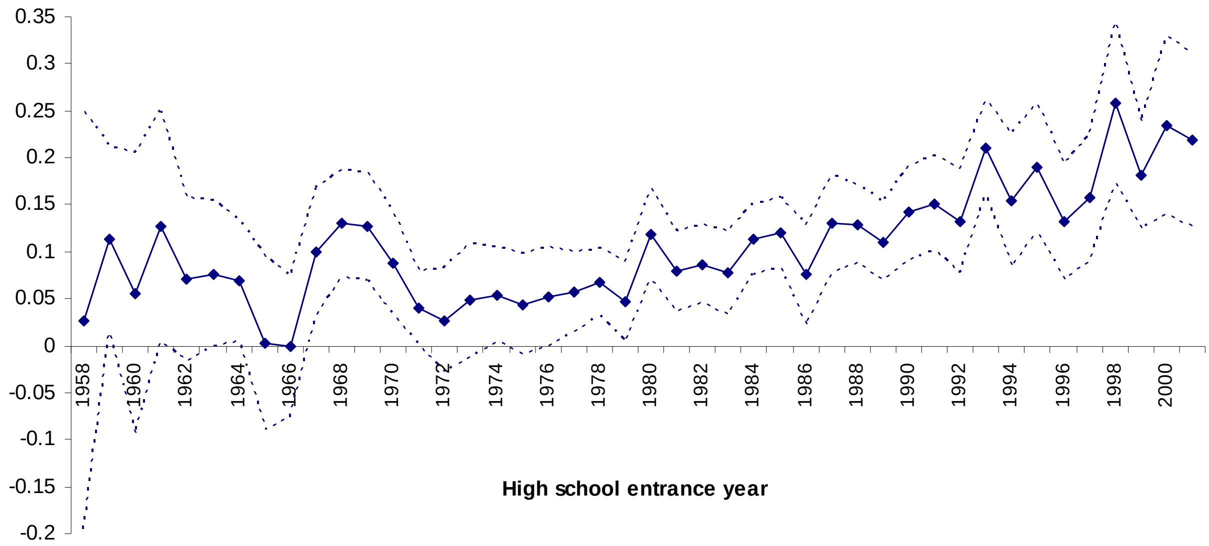
To better understand the relevance of changing fixed effects in estimating biased coefficients, I analyze a simplified situation where there are only two cities, city 1 and city 2, with city 1 being of the high profile city. I drop the age and year fixed effects for simplicity. The estimates of the coefficients are:

(28) $$\begin{bmatrix} \hat{\beta} \\ \hat{\gamma} \\ \hat{\mu}_1 \\ \hat{\mu}_2 \end{bmatrix} = \begin{bmatrix} \beta \\ \gamma \\ \mu_1 \\ \mu_2 \end{bmatrix} + \begin{bmatrix} \sum x_i x_i & \sum x_i a_i & \sum x_i I_1 & \sum x_i I_2 \\ \sum x_i a_i & \sum a_i a_i & \sum a_i I_1 & \sum a_i I_2 \\ \sum x_i I_1 & \sum a_i I_1 & N_1 & 0 \\ \sum x_i I_2 & \sum a_i I_2 & 0 & N_2 \end{bmatrix}^{-1} \begin{bmatrix} \sum x_i\theta_i \\ \sum a_i\theta_i \\ \sum \theta_i I_1 \\ \sum \theta_i I_2 \end{bmatrix}$$

and I am interested in $(\hat{\beta}^M - \hat{\beta}^T)$ and $(\hat{\gamma}^M - \hat{\gamma}^T)$. As before I assume that the marginal and joint distribution of income and ability does not change. If there is no inter-city migration then the *X'X* matrix remains unchanged under both regimes because households are simply reshuffling within city without changing the overall city average income and ability, i.e, $\sum x_i I_j$, $\sum a_i I_j$, and N_j all remain the same. Hence the change in bias only comes from the within city migration captured in the *X'U* matrix and there is no need to allow changing fixed effects when estimating the change in coefficients.

On the other hand, if in reality there is inter-city migration so that the average income and ability changes but I assume constant city fixed effects, then I would be incorrectly averaging out the effect that the change in city specific income and ability would have on the coefficients.

Figure 1. Coefficient of log father's years of education* high school year



Notes: Figure plots the coefficients β_k from the equation, $y_{ik} = \sum_k (fedu_{ik} \times D_{ik}) \times \beta_k + \mu_k + \varepsilon_{ik}$ where D_{ik} is a dummy variable equal to one if individual i entered high school in year k and μ_k are birth year fixed effects. $fedu_{ik}$ is log of father's years of education and y_{ik} is the log of achievement measured by the average entrance exam score based on one's college status. I use a nationally representative sample and pool all individuals who went to high school in the Korea Labor and Income Panel Survey (KLIPS) to generate the coefficients. The shift in high school admission policy from tracking to mixing started in 1974 beginning with Seoul and then gradually rolled out to other cities until 1980. Some of the cities that maintained a tracking admission policy during 1974-1980 later shifted to mixing after 2000. Cohorts born during the Korean War (1950-53) would have entered high school around 1966-70. Hence, the hump around 1967-1970 could likely reflect selective birth of higher educated households. 95% Confidence bands are in dotted lines.

Figure 2. Illustration of single-crossing in income and ability : Indifference curves in *r-θ* space

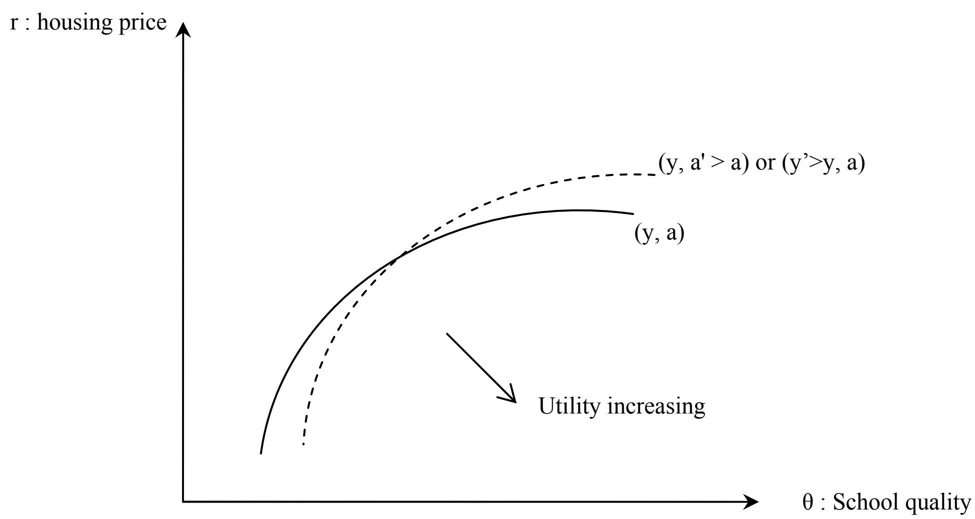
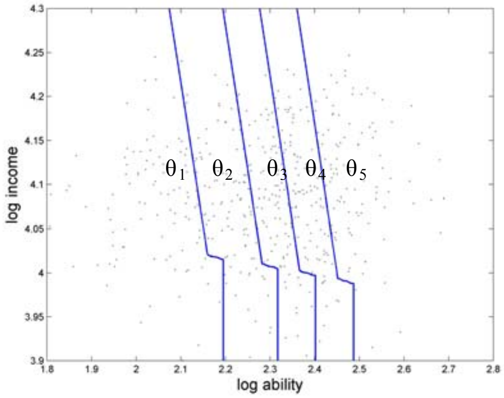
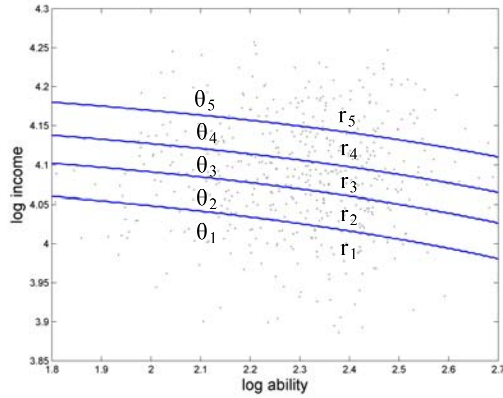


Figure 3. School allocation, tutoring choice, and achievement under tracking and mixing equilibrium (δ=0.5, α=0.2, β=0.6)

A. School allocation under tracking

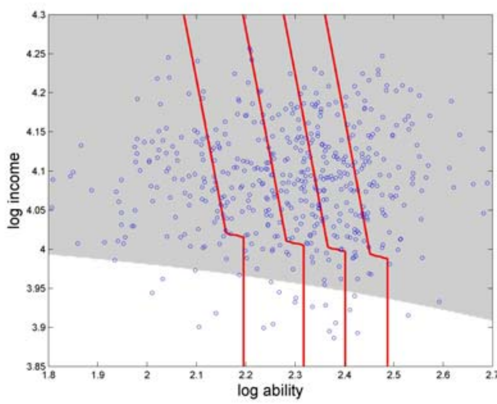


B. School allocation under mixing

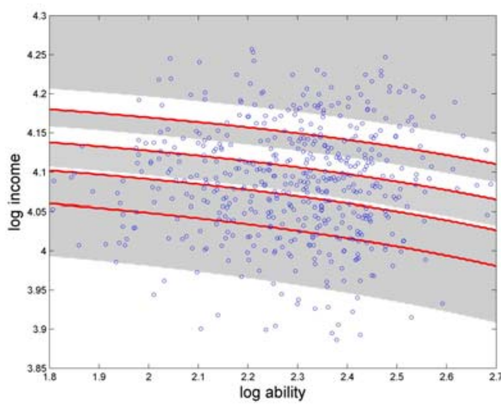


Notes: There is one school per neighborhood/district and each school is represented by school quality θ where $\theta_1 < \theta_2 < \theta_3 < \theta_4 < \theta_5$. Under tracking all neighborhoods pay the same price for housing. Rent premium emerges under mixing and $r_1 < r_2 < r_3 < r_4 < r_5$. Each dot represents a household.

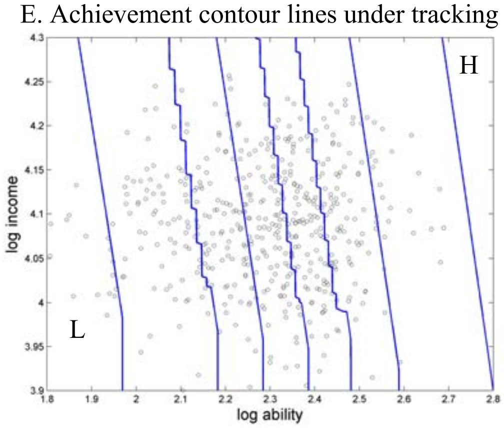
C. Tutoring choice under tracking
- Shaded region choose tutoring



D. Tutoring choice under mixing
- Shaded region choose tutoring



E. Achievement contour lines under tracking



F. Achievement contour lines under mixing

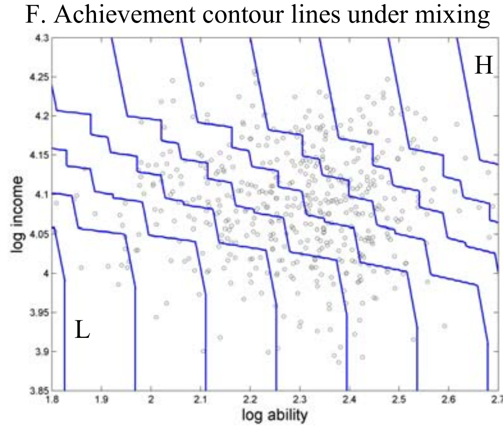
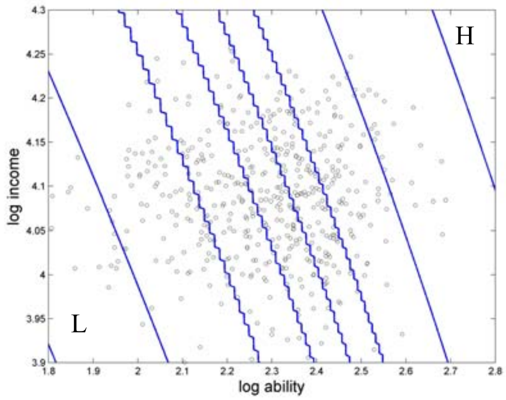


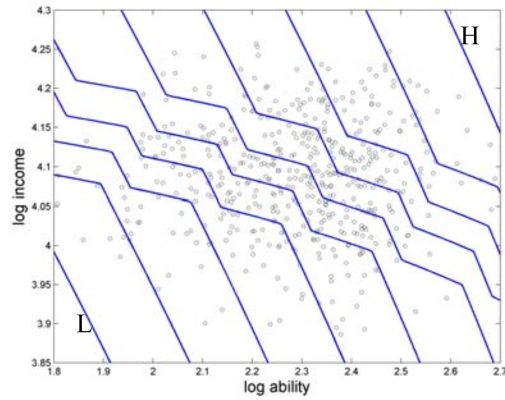
Figure 4. Achievement contour lines under alternative specifications

Alternative specification 1: δ=1.0, α=0.4, β=0.4

A. Achievement contour lines under tracking

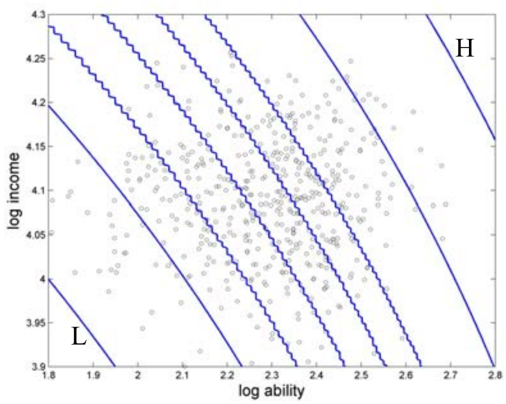


B. Achievement contour lines under mixing

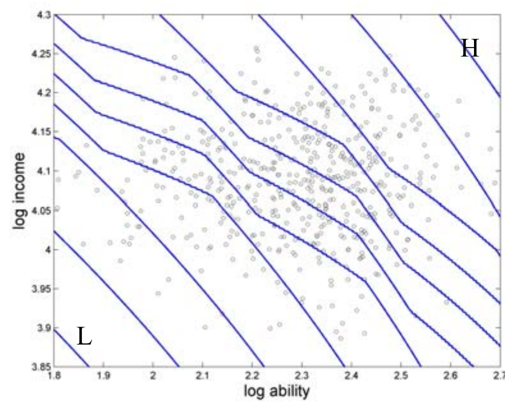


Alternative specification 2: δ=2.0, α=0.6, β=0.2

C. Achievement contour lines under tracking

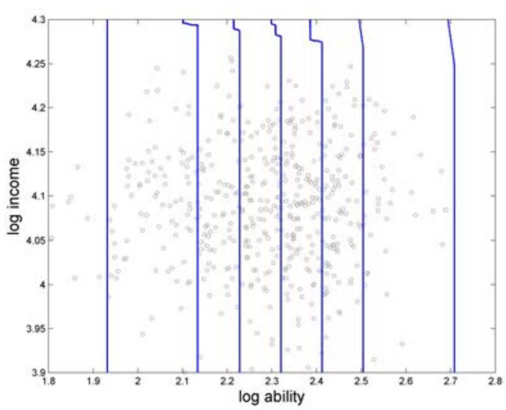


D. Achievement contour lines under mixing



Alternative specification 3: δ=0.2, α=0.6, β=0.2, p=3

E. Achievement contour lines under tracking



F. Achievement contour lines under mixing

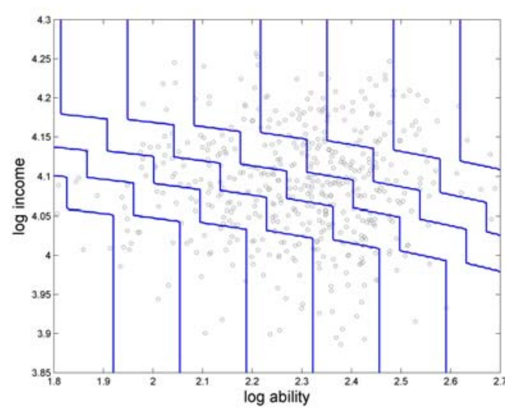
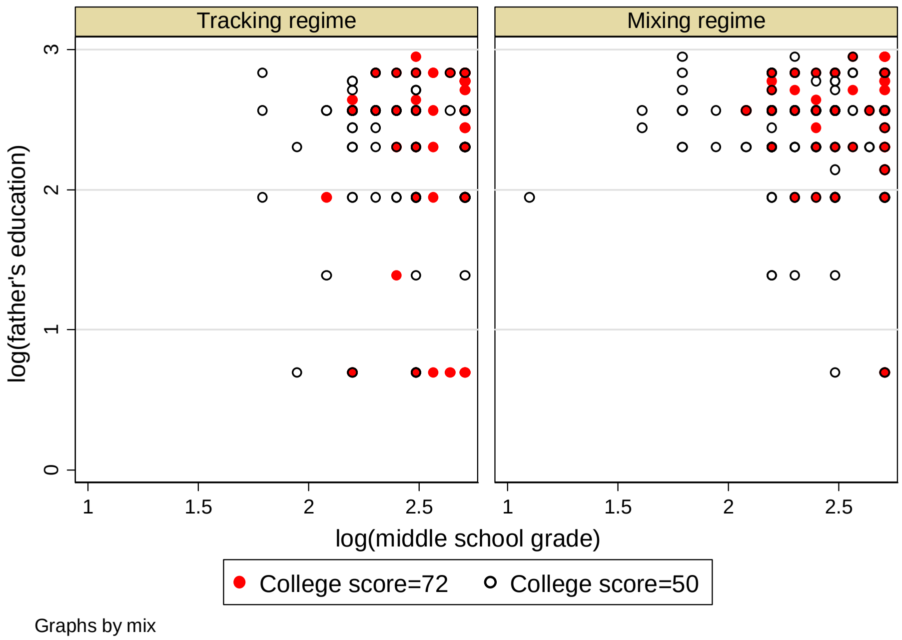
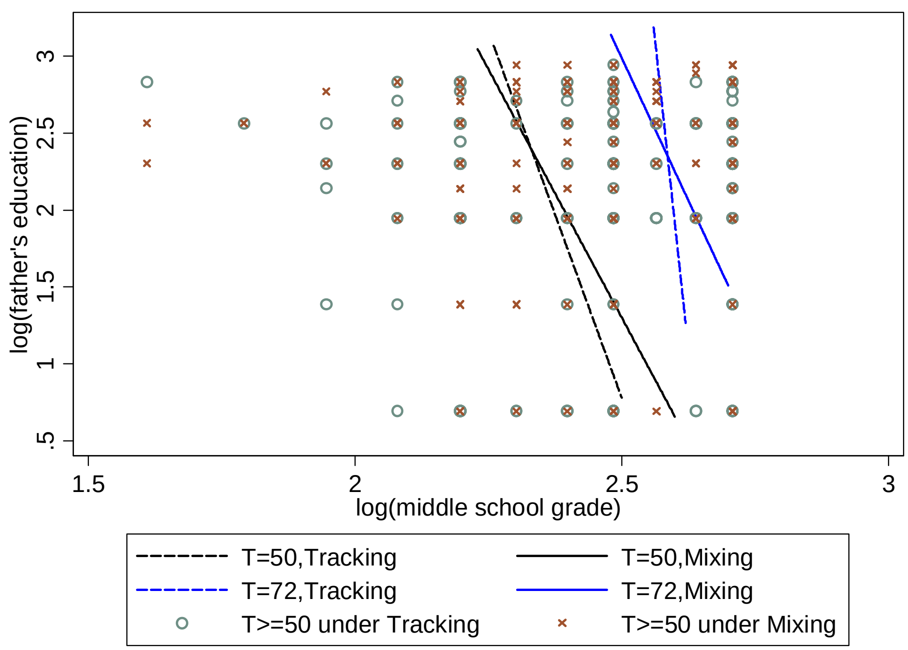


Figure 5. Scatter plot of of father's education and middle school grade by regime



Notes: Solid circle are for individuals with a college score of 70 and hollow circle are for those with a score of 50. Left panel is under tracking and right panel is under mixing.

Figure 6. Iso-achievement lines by regime:
Estimated based on OLS coefficients on log of own middle school grade and log of father's years of education



Notes: Dashed line are based on coefficients under the tracking regime and solid lines under the mixing regime. Coefficients were estimated using all observations with a college score (T) above 72 and 50, respectively.

Figure 7. Evolution of $\frac{\hat{\beta}}{\hat{\gamma}}$ by policy cohorts for all treatment and control cities

(Hypothetical policy year for control cities is 1980)

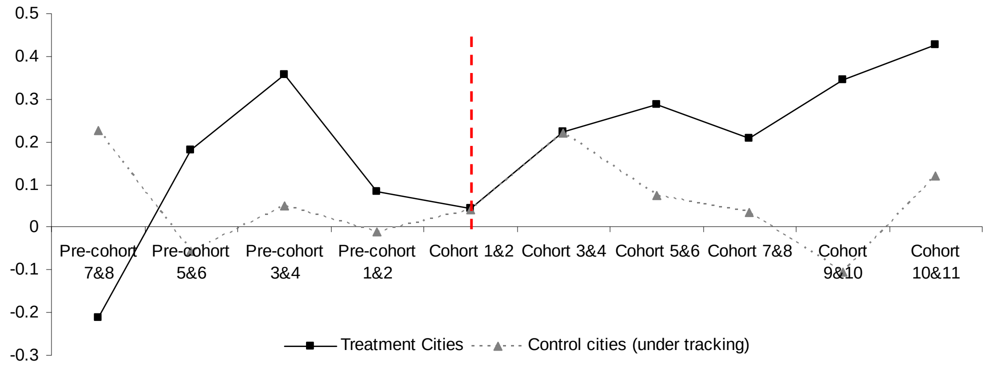
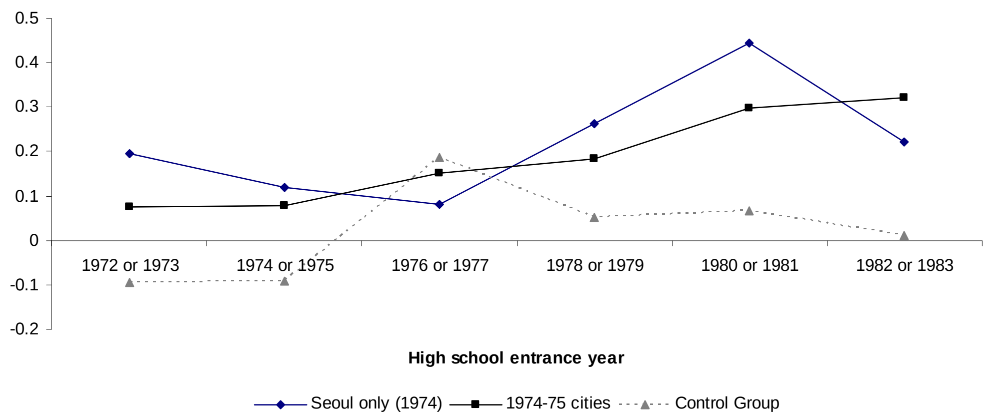


Figure 8. Evolution of $\frac{\hat{\beta}}{\hat{\gamma}}$ by high school entrance year by city groups

A. Cities that shift regimes in 1974&75



B. Cities that shift regimes in 1979&80

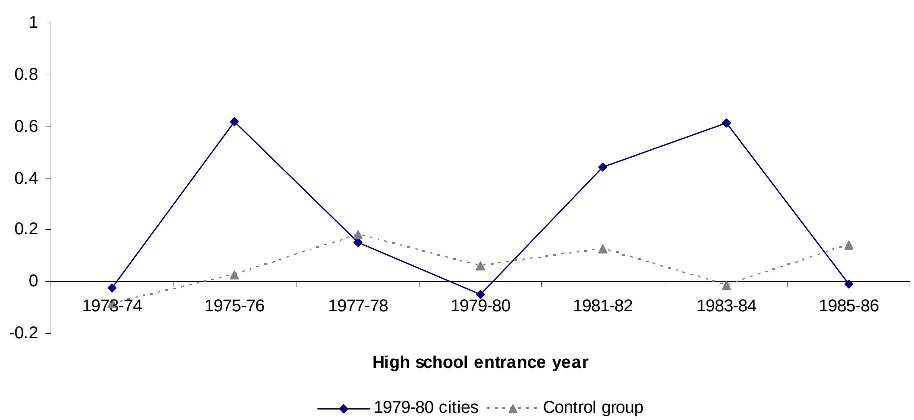
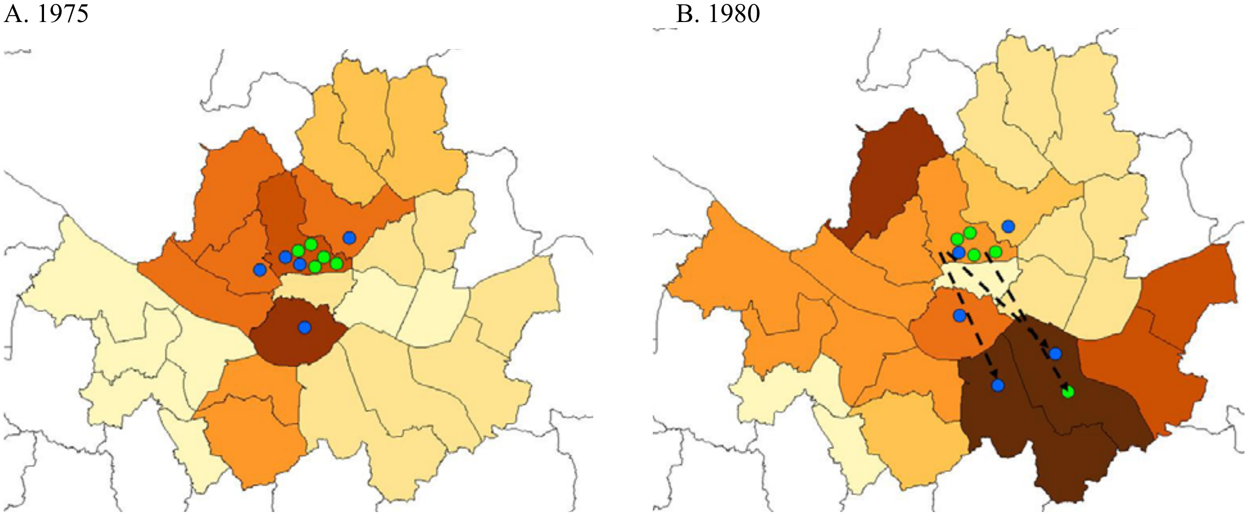


Figure 9. Movement of top-tier high schools from the Unified Central District to District 8 and the change in the ratio of college educated adults in Seoul

A. 1975 B. 1980



Note: Blue dots indicate public high schools and green dots private high schools. The three schools that moved during this period opened their new campuses in the 8th district in 1976, 1978, and 1980 with announcements of campus moves coming two to three years earlier. Darker shades imply a higher percentage of college educated adults in the administrative districts.

Figure 10. Boundaries of the Unified Central District and District 8

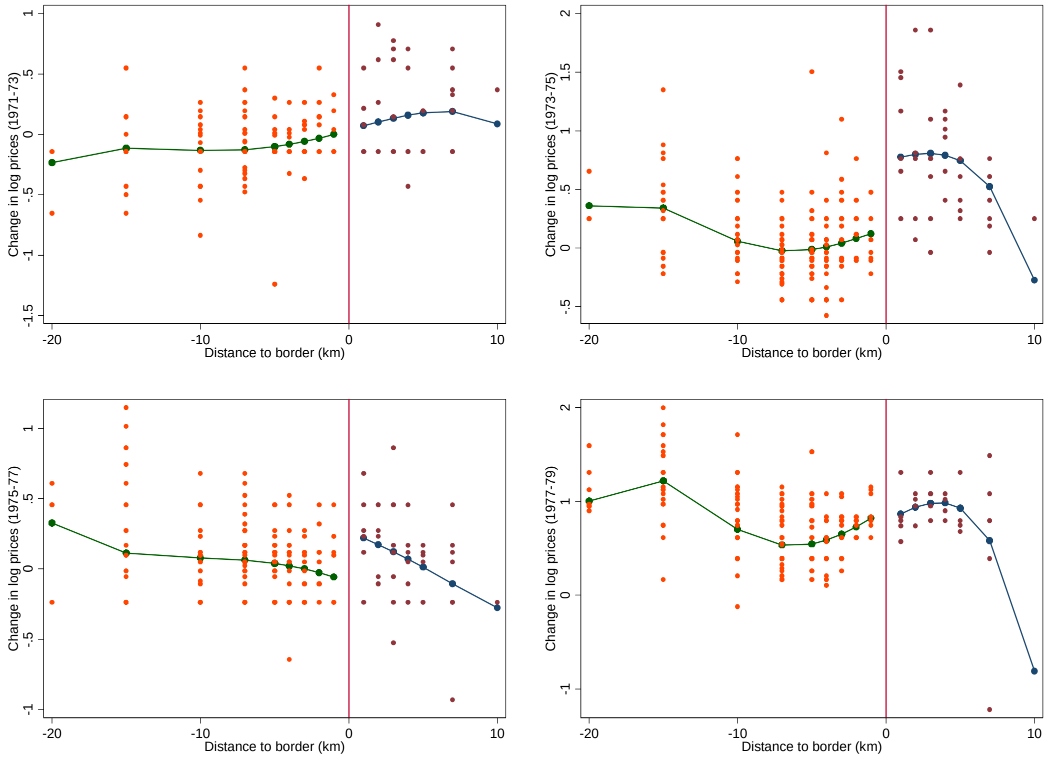
A. Unified Central District with 0.5 km bounds B. District 8 with 1km bounds

Note: Red dots indicate the geographic center of each neighborhood(dong). Distance from district borders are measured as the shortest distance from each dot to the both borders and then assigned based on the integer number in kilometers away from the boundary lines.

Figure 11. Change in medium quality housing land prices by distance from district borders

(Fourth order polynomials fitted. Years are 71-73 , 73-75, 79-77, 75-77 in clockwise order)

A. Change in housing land price by distance from District 8 border



B. Change in housing land price by distance from Unified Central District border

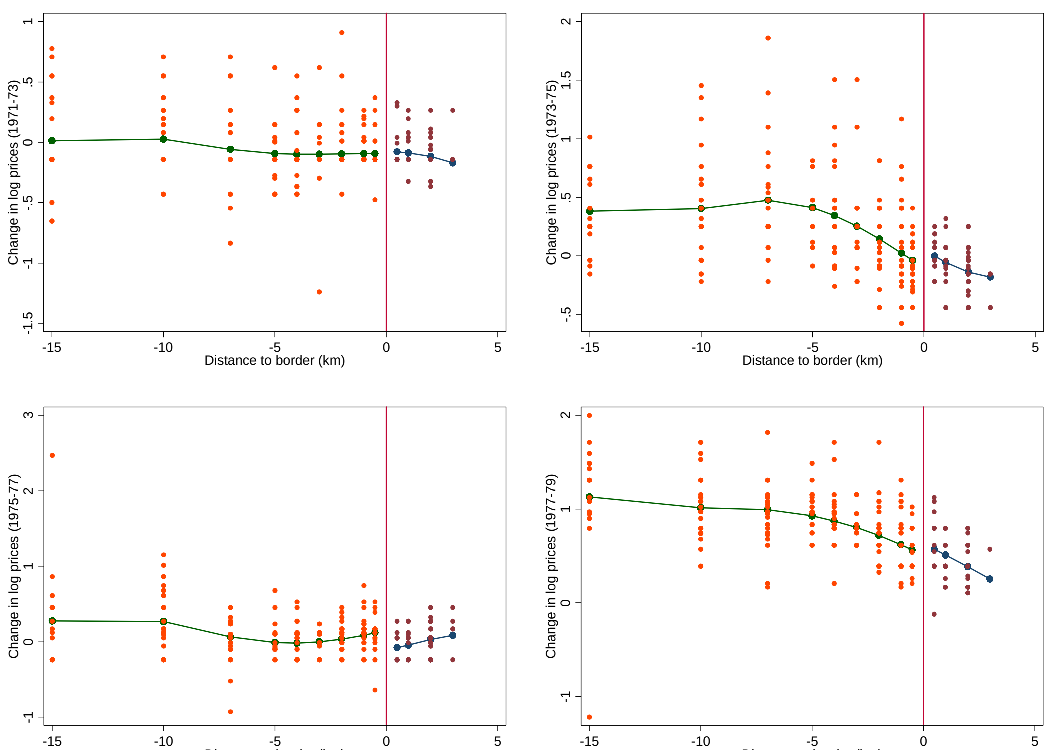
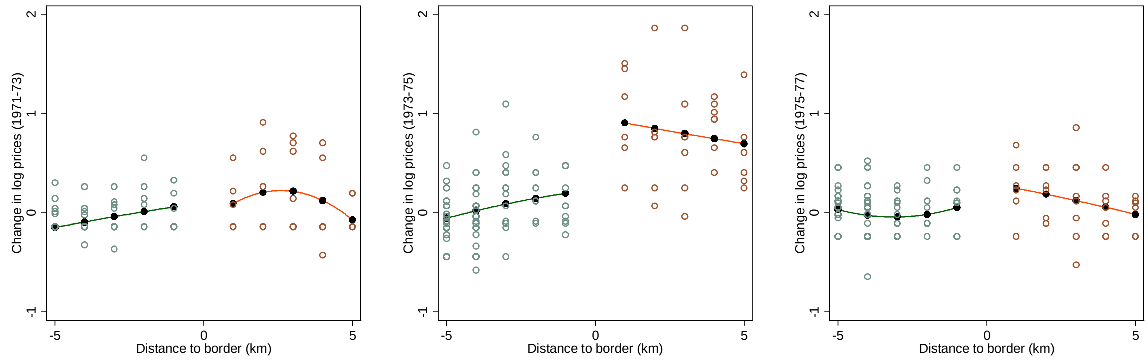
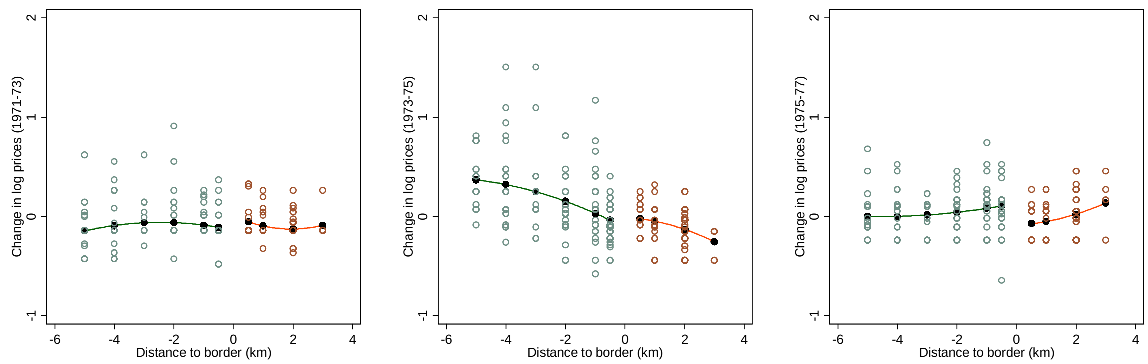


Figure 12. Border discontinuity estimates using narrower (5km) windows from district boundary

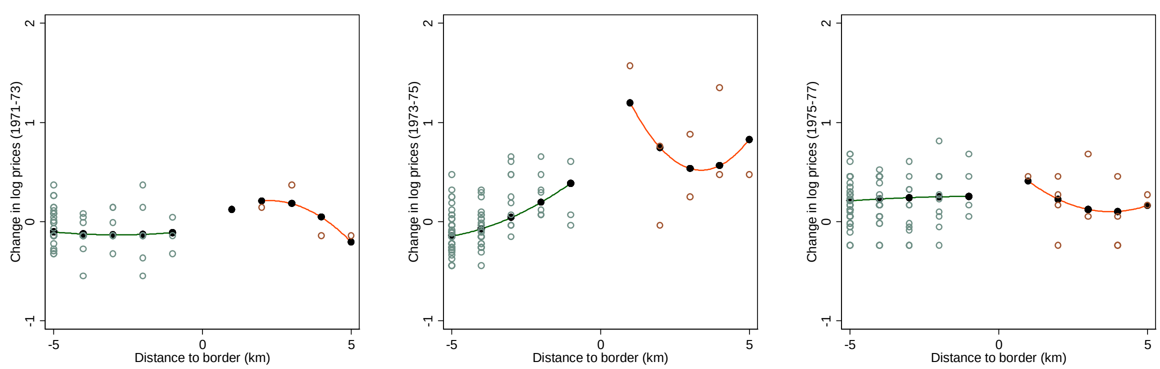
A. Change in housing land price by distance from District 8 border



B. Change in housing land price by distance from Unified Central District border



C. Change in commercial land price by distance from District 8 border



D. Change in commercial land price by distance from Unified Central District border

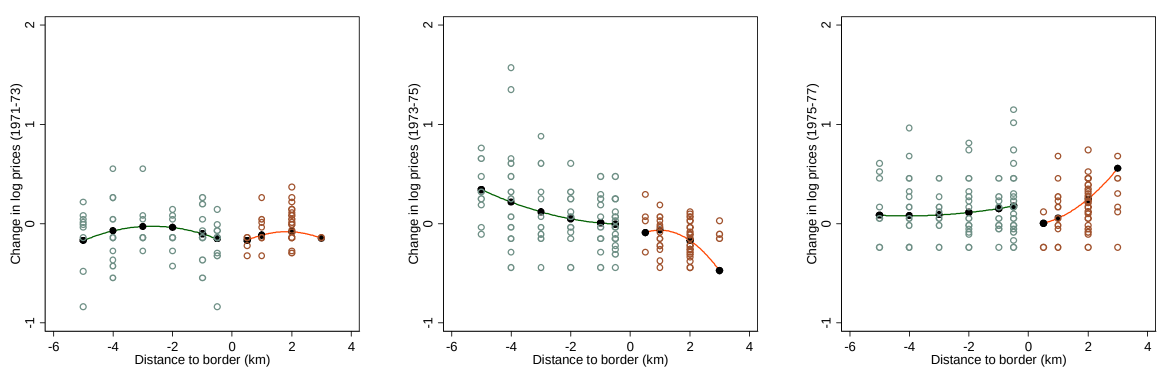


Table 1. Descriptive Statistics

Variable	Mean	Std. Dev.	Min	Max	Obs
Panel A: Individual level data used in the difference in difference analysis					
College score	47.439	10.030	41	80.93	2420
Father's year of education	7.545	4.197	1	18	2348
Middle school grade	10.530	2.715	3	15	2426
Received tutoring during high school	0.096	0.294	0	1	2438
Received tutoring in middle school	0.050	0.219	0	1	2437
General high school student	0.669	0.471	0	1	2456
Male	0.551	0.497	0	1	2456
Year of birth	1962.693	4.450	1952	1972	2456
Year of high school entrance	1978.698	4.410	1970	1985	2456
Age when entering high school	16.006	0.581	12	20	2456
Went to any college	0.398	0.490	0	1	2455
Went to top 3 college	0.014	0.119	0	1	2456
Went to top 10 college	0.036	0.186	0	1	2456
Went to top 20 college	0.057	0.233	0	1	2456
High school code	1085.063	641.030	1	2186	2312
City code	754.651	512.451	100	1602	2380
High school in group 1 (1974) cities	0.275	0.447	0	1	2456
High school in group 2 (1975) cities	0.130	0.337	0	1	2456
High school in group 3 (1979) cities	0.077	0.267	0	1	2456
High school in group 4 (1980) cities	0.083	0.275	0	1	2456
Panel B: Neighborhood(dong) level price and distance data used in the border discontinuity design					
Distance to 8th District boundary	-6.12931	5.029554	-20	10	464
Distance to Central District boundary	-2.503233	4.707249	-15	3	464
Housing land price in 1971	113.6	92.0	5.4	448.7	403
Housing land price in 1973	94.9	73.8	4.7	389.2	397
Housing land price in 1975	89.9	54.9	7.0	250.0	377
Housing land price in 1977	98.5	65.5	5.5	315.0	428
Housing land price in 1979	156.9	56.4	17.4	348.9	379
Commercial land price in 1971	504.4	511.3	26.9	2692.4	278
Commercial land price in 1973	441.6	460.4	31.1	2802.4	290
Commercial land price in 1975	396.3	377.6	30.0	2500.0	287
Commercial land price in 1977	422.9	467.7	39.4	3544.2	378
Commercial land price in 1979	679.4	635.9	145.4	4652	335

Notes : Land prices in Panel B are for medium quality housing and commercial land. Prices for high and low quality housing and commercial land are reported in the Appendix.

Sources: KLIPS (Round 1-11), Korea Land Appraisal Annals (1971-1979), Census (1975, 1980, 1985), Seoul Statistics Annal (1975, 78, 81, 85), Seoul Education Statistics Annal (1975, 78, 81, 85).

Table 2. Effect of the education regime shift on achievement production

Dependent variable: ln(college score)	(1)	(2)	(3)	(4)	(5)	(6)
City shifts to mixing (MIX)	-0.1906***	-0.2819***	-0.2802***	-0.2369***	-0.3146***	-0.1220*
	(0.0687)	(0.0732)	(0.0714)	(0.0695)	(0.0847)	(0.0728)
Ln(father's years of education)	0.0273***	0.0255***	0.0236***	0.0162**	0.0216**	0.0061
	(0.0099)	(0.0089)	(0.0086)	(0.0078)	(0.0088)	(0.0106)
Ln(father's education)*MIX	0.0421***	0.0464***	0.0476***	0.0397***	0.0353**	0.0249
	(0.0118)	(0.0106)	(0.0110)	(0.0096)	(0.0164)	(0.0160)
Ln(middle school grade)	0.2205***	0.2185***	0.2146***	0.1952***	0.2452***	0.0994***
	(0.0210)	(0.0206)	(0.0201)	(0.0203)	(0.0248)	(0.0239)
Ln(MS grade)*MIX	0.0430	0.0483*	0.0487*	0.0346	0.0670*	-0.0221
	(0.0291)	(0.0285)	(0.0278)	(0.0282)	(0.0363)	(0.0269)
MIX*GHS						-0.1534
						(0.0932)
Ln(father's education)*GHS						0.0181
						(0.0121)
Ln(father's education)*MIX*GHS						0.0082
						(0.0217)
Ln(MS grade)*GHS						0.1505***
						(0.0271)
Ln(MS grade)*MIX*GHS						0.0836**
						(0.0358)
General high school student(GHS)				0.1068***		-0.2969***
				(0.0066)		(0.0614)
Received tutoring when MS student			0.0616***	0.0463***	0.0477**	0.0415***
			(0.0136)	(0.0142)	(0.0226)	(0.0156)
Male	0.0681***	0.0673***	0.0686***	0.0704***	0.0826***	0.0671***
	(0.0075)	(0.0072)	(0.0077)	(0.0067)	(0.0071)	(0.0067)
City, Age, and HS year FEs	Y	Y	Y	Y	Y	Y
City FEs*MIX		Y	Y	Y	Y	Y
Observations	2,236	2,236	2,236	2,236	1,507	2,236
R-squared	0.3250	0.3358	0.3404	0.3963	0.4136	0.4220
Test-statistic:						
$t = \hat{\beta}_{mix}\hat{\gamma}_{track} - \hat{\gamma}_{mix}\hat{\beta}_{track}$	0.0081	0.0089	0.0091	0.0072	0.0072	0.0068
	(0.0032)	(0.0030)	(0.0030)	(0.0024)	(0.0047)	(0.0043)
p-value of one sided hypothesis test $H_0: \frac{\hat{\beta}_{track}}{\hat{\gamma}_{track}} = \frac{\hat{\beta}_{mix}}{\hat{\gamma}_{mix}}, \quad H_1: \frac{\hat{\beta}_{track}}{\hat{\gamma}_{track}} < \frac{\hat{\beta}_{mix}}{\hat{\gamma}_{mix}}$	0.0056***\	0.0018***	0.0015***	0.0017***	0.0623*	0.0571*

Note: Analysis is based on the sample of individuals who entered high school between 1970 and 1985. ***, **, and * indicates significance at the 1%, 5%, and 10% level, respectively. City clustered Huber-White standard errors are in parantheses.

Table 3. Achievment by cities with non-treatment cities as controls

	All high school students			General high school students		
	Seoul (1974)	1974&75 Cities	1979&80 Cities	Seoul (1974)	1974&75 Cities	1979&80 Cities
Dependent variable: ln(college score)	(1)	(2)	(3)	(4)	(5)	(6)
City shifts to mixing (MIX)	-0.2894***	-0.2557***	-0.0807	-0.3568***	-0.3180***	-0.3911*
	(0.0689)	(0.0595)	(0.0951)	(0.0671)	(0.0681)	(0.2155)
Ln(father's education)	0.0149	0.0125	0.0125*	0.0135	0.0139	0.0157*
	(0.0096)	(0.0097)	(0.0066)	(0.0111)	(0.0108)	(0.0091)
Ln(father's education)*MIX	0.0500***	0.0517***	0.0383**	0.0675***	0.0617***	0.0252
	(0.0102)	(0.0107)	(0.0147)	(0.0107)	(0.0105)	(0.0311)
Ln(middle school grade)	0.1945***	0.1954***	0.1859***	0.2454***	0.2466***	0.2303***
	(0.0244)	(0.0239)	(0.0229)	(0.0298)	(0.0292)	(0.0289)
Ln(ms grade)*MIX	0.0471*	0.0312	0.0143	0.0546*	0.0446	0.0610
	(0.0260)	(0.0259)	(0.0468)	(0.0291)	(0.0293)	(0.0919)
Male	0.0771***	0.0737***	0.0715***	0.0892***	0.0861***	0.0876***
	(0.0067)	(0.0073)	(0.0086)	(0.0076)	(0.0077)	(0.0116)
Received tutoring when MS student	0.0618***	0.0614***	0.0305	0.0703***	0.0695***	0.0284
	(0.0153)	(0.0133)	(0.0272)	(0.0196)	(0.0171)	(0.0321)
General high school student(GHS)	0.0909***	0.0952***	0.0921***			
	(0.0136)	(0.0114)	(0.0098)			
City, Age, HS year FEs and City FEs*MIX	Y	Y	Y	Y	Y	Y
Observations	1,397	1,562	1,308	995	1,067	894
R-squared	0.3901	0.3915	0.4044	0.4194	0.4163	0.4637
Test-statistic: $t = \hat{\beta}_{mix}\hat{\gamma}_{track} - \hat{\gamma}_{mix}\hat{\beta}_{track}$	0.009	0.0097	0.0069	0.0158	0.0146	0.0049
	(0.0028)	(0.0028)	(0.0031)	(0.0042)	(0.0040)	(0.0078)
p-value of one sided test: $H_0 : \frac{\hat{\beta}_{track}}{\hat{\gamma}_{track}} = \frac{\hat{\beta}_{mix}}{\hat{\gamma}_{mix}}, \quad H_1 : \frac{\hat{\beta}_{track}}{\hat{\gamma}_{track}} < \frac{\hat{\beta}_{mix}}{\hat{\gamma}_{mix}}$	0.0008***	0.0004***	0.0141**	0.0002***	0.0002***	0.2676

Note: Analysis is based on the sample of individuals who entered high school between 1970 and 1985. ***, **, and * indicates significance at the 1%, 5%, and 10% level, respectively. City clustered Huber-White standard errors are in parantheses.

Table 4. City Group Specific Interaction Term Analysis with Group 5 Cities as Hypothetical Controls

	Seoul		1974&75 Cities			1979&80 Cities	
Dependent Variable:	Treated=1	Treated=0	Treated=1	Treated=0		Treated=1	Treated=0
ln(college score)	(1)	(2)	(3)	(4)		(5)	(6)
Coefficient of Log(father's education)*1(treated)*1(high school entrance year =)							
1970 or 1971	0.1714***	-0.0200	0.0694	-0.0189			
	(0.0233)	(0.0224)	(0.0526)	(0.0227)	1971 or 1972	-0.0685	0.0148
1972 or 1973	0.0675***	-0.0077	0.0172	0.0024		(0.0570)	(0.0198)
	(0.0245)	(0.0242)	(0.0361)	(0.0233)	1973 or 1974	0.0127	-0.0161
1974 or 1975	0.0385	-0.0127	0.0312	-0.0140		(0.0376)	(0.0252)
	(0.0236)	(0.0241)	(0.0246)	(0.0237)	1975 or 1976	0.0784**	0.0067
1976 or 1977	-0.0047	0.0257	0.0145	0.0206		(0.0341)	(0.0241)
	(0.0202)	(0.0207)	(0.0268)	(0.0206)	1977 or 1978	-0.0141	0.0356*
1978 or 1979	0.0537***	0.0093	0.0417**	0.0046		(0.0301)	(0.0188)
	(0.0122)	(0.0131)	(0.0180)	(0.0131)	1979 or 1980	-0.0216	0.0124
1980 or 1981	0.0957***	0.0144	0.0657**	0.0021		(0.0327)	(0.0181)
	(0.0202)	(0.0215)	(0.0332)	(0.0203)	1981 or 1982	0.0475	0.0318
1982 or 1983	0.0671***	0.0032	0.0927***	0.0003		(0.0360)	(0.0238)
	(0.0223)	(0.0225)	(0.0317)	(0.0222)	1983 or 1984	0.0635	-0.0027
1984 or 1985	0.0880***	0.0146	0.0535*	0.0226		(0.0411)	(0.0244)
	(0.0274)	(0.0277)	(0.0317)	(0.0266)			
Coefficient of Log(middle school grade)*1(treated)*1(high school entrance year =)							
1970 or 1971	0.1034**	0.1980***	0.0494	0.1788***			
	(0.0448)	(0.0551)	(0.0424)	(0.0539)	1971 or 1972	0.0035	0.1285***
1972 or 1973	0.2250***	0.0819**	0.1272***	0.1303***		(0.0808)	(0.0444)
	(0.0377)	(0.0334)	(0.0373)	(0.0345)	1973 or 1974	-0.0534	0.1927***
1974 or 1975	0.0778**	0.1392**	0.0960***	0.1278*		(0.0674)	(0.0452)
	(0.0337)	(0.0634)	(0.0338)	(0.0648)	1975 or 1976	-0.1017	0.2392***
1976 or 1977	0.1215***	0.1376***	0.1094**	0.1207***		(0.0735)	(0.0622)
	(0.0312)	(0.0354)	(0.0428)	(0.0345)	1977 or 1978	-0.0529	0.1938***
1978 or 1979	0.0566**	0.1837***	0.0734**	0.1802***		(0.0661)	(0.0434)
	(0.0283)	(0.0421)	(0.0332)	(0.0350)	1979 or 1980	-0.0155	0.2015***
1980 or 1981	0.0298	0.2176***	0.0610	0.1659***		(0.0635)	(0.0493)
	(0.0326)	(0.0521)	(0.0430)	(0.0485)	1981 or 1982	-0.0709	0.2510***
1982 or 1983	0.0547	0.2608***	0.0293	0.2607***		(0.0687)	(0.0495)
	(0.0356)	(0.0483)	(0.0454)	(0.0389)	1983 or 1984	-0.0989	0.1978***
1984 or 1985	0.0235	0.1551***	0.0600	0.1831***		(0.0692)	(0.0584)
	(0.0425)	(0.0469)	(0.0444)	(0.0482)			
R-squaured	0.42		0.41			0.42	
Observations	1473		1975			1375	

Notes: Analysis is based on the sample of individuals who entered high school between 1970 and 1985. Dummy variables for gender, middle school tutoring, and general high school are included as before. Also city, age, high school year fixed effects, and city fixed effects interacted with the policy variable are included. ***, **, and * indicates significance at the 1%, 5%, and 10% respectively. City clustered Huber-White standard errors are in parantheses.

Table 5. Effect of the education regime shift on tutoring choice and attending a general high school - Linear Probability Model

Dependent variable:	Received private tutoring when HS student				Attended a GHS	
	(1)	(2)	(3)	(4)	(5)	(6)
City shifts to mixing (MIX)	-0.0667	-0.0463	-0.0374	-0.0390	-0.1997*	-0.1945*
	(0.0574)	(0.0511)	(0.0734)	(0.0699)	(0.1018)	(0.1009)
Father's years of education	0.0084*	0.0056	0.0112**	0.0083*	0.0135***	0.0128***
	(0.0043)	(0.0037)	(0.0052)	(0.0044)	(0.0032)	(0.0032)
Father's education*MIX	-0.0005	0.0002	-0.0032	-0.0029	0.0090*	0.0092*
	(0.0041)	(0.0037)	(0.0045)	(0.0044)	(0.0054)	(0.0055)
Middle school grade	0.0112***	0.0077**	0.0117***	0.0068*	0.0192***	0.0183***
	(0.0039)	(0.0036)	(0.0044)	(0.0034)	(0.0060)	(0.0058)
Middle school grade*MIX	-0.0002	0.0009	-0.0012	0.0024	0.0150**	0.0153**
	(0.0052)	(0.0051)	(0.0061)	(0.0058)	(0.0066)	(0.0067)
Male	-0.0481***	-0.0377***	-0.0266	-0.0171	-0.0182	-0.0155
	(0.0136)	(0.0101)	(0.0181)	(0.0153)	(0.0270)	(0.0273)
Received tutoring when MS student		0.5110***		0.5752***		0.1314**
		(0.0425)		(0.0466)		(0.0580)
General high school student(GHS)		0.0011				
		(0.0171)				
City, Age, HS year FEs and City FEs*MIX	Y	Y	Y	Y	Y	Y
Observations	2,267	2,267	1,528	1,528	2,267	2,267
R-squared	0.1312	0.2628	0.1579	0.3392	0.1938	0.1973
Test-statistic:						
$t = \hat{\beta}_{mix}\hat{\gamma}_{track} - \hat{\gamma}_{mix}\hat{\beta}_{track}$	-0.000004	-0.000004	-0.00002	-0.00004	-0.00003	-0.00003
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0002)	(0.0002)
p-value of one sided test: $H_0: \frac{\hat{\beta}_{track}}{\hat{\gamma}_{track}} = \frac{\hat{\beta}_{mix}}{\hat{\gamma}_{mix}}, \quad H_1: \frac{\hat{\beta}_{track}}{\hat{\gamma}_{track}} \neq \frac{\hat{\beta}_{mix}}{\hat{\gamma}_{mix}}$	0.958	0.943	0.792	0.579	0.877	0.883

Note: Analysis is based on the sample of individuals who entered high school between 1970 and 1985. ***, **, and * indicates significance at the 1%, 5%, and 10% level, respectively. City clustered Huber-White standard errors are in parantheses.

Table 6. Tutoring choice by cities with non-treatment cities as controls - Linear Probability Model

	All high school students			General high school students		
	Seoul	1974&75	1979&80	Seoul	1974&75	1979&80
Dependent variable:	(1974)	Cities	Cities	(1974)	Cities	Cities
Received tutoring when HS student	(1)	(2)	(3)	(4)	(5)	(6)
City shifts to mixing (MIX)	-0.1027**	-0.0628	-0.3329***	-0.1298*	-0.1254*	-0.0119
	(0.0427)	(0.0532)	(0.0704)	(0.0744)	(0.0721)	(0.1186)
Father's years of education	0.0058	0.0059	0.0030	0.0097*	0.0100*	0.0049*
	(0.0049)	(0.0047)	(0.0023)	(0.0058)	(0.0055)	(0.0026)
Father's education*MIX	0.0005	-0.0014	0.0068	-0.0041	-0.0045	0.0062
	(0.0050)	(0.0041)	(0.0057)	(0.0057)	(0.0053)	(0.0062)
Middle school grade	0.0050	0.0057	0.0079***	0.0030	0.0039	0.0044
	(0.0037)	(0.0038)	(0.0029)	(0.0030)	(0.0032)	(0.0036)
Middle school grade*MIX	0.0060*	0.0038	0.0017	0.0105***	0.0104***	0.0009
	(0.0033)	(0.0047)	(0.0073)	(0.0029)	(0.0029)	(0.0109)
Male	-0.0240	-0.0272*	-0.0196	0.0038	-0.0000	0.0090
	(0.0170)	(0.0144)	(0.0141)	(0.0235)	(0.0211)	(0.0184)
Received tutoring when MS student	0.5906***	0.5760***	0.5627***	0.6457***	0.6333***	0.5909***
	(0.0318)	(0.0276)	(0.0624)	(0.0280)	(0.0286)	(0.0650)
General high school student(GHS)	0.0138	0.0144	-0.0223			
	(0.0169)	(0.0144)	(0.0167)			
City, Age, HS year FEs and City FEs*MIX	Y	Y	Y	Y	Y	Y
Observations	1,418	1,583	1,324	1,009	1,081	902
R-squared	0.3366	0.3161	0.3462	0.3974	0.3854	0.4363
Test-statistic:						
$t = \hat{\beta}_{mix}\hat{\gamma}_{track} - \hat{\gamma}_{mix}\hat{\beta}_{track}$	-0.00003	-0.00003	0.00005	-0.0001	-0.0001	0.00002
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
p-value of two sided test:						
$H_0: \frac{\hat{\beta}_{track}}{\hat{\gamma}_{track}} = \frac{\hat{\beta}_{mix}}{\hat{\gamma}_{mix}}, \quad H_1: \frac{\hat{\beta}_{track}}{\hat{\gamma}_{track}} \neq \frac{\hat{\beta}_{mix}}{\hat{\gamma}_{mix}}$	0.647	0.588	0.437	0.152	0.124	0.766

Note: Analysis is based on the sample of individuals who entered high school between 1970 and 1985. ***, **, and * indicates significance at the 1%, 5%, and 10% level, respectively. City clustered Huber-White standard errors are in parantheses.

Table 7. Robustness tests

Dependent variable: ln(college score)	Attend a top 3 college (1)	Attend a top 10 college (2)	Attend a top 20 colleges (3)	Reverse coding (4)	All schools (5)	Schools with at least 5 obs (6)
City shifts to mixing (MIX)	-0.2745***	-0.4856***	-0.4641***	0.0355	-0.1442	-0.0671
	(0.0996)	(0.1140)	(0.0000)	(0.0773)	(0.1117)	(0.1098)
Ln(father's education)	0.0028	0.0071	0.0034	-0.0170	0.0306	0.0337**
	(0.0029)	(0.0064)	(0.0069)	(0.0115)	(0.0186)	(0.0166)
Ln(father's education)*MIX	0.0117*	0.0275***	0.0252*	-0.0378**	0.0157	0.0087
	(0.0065)	(0.0064)	(0.0134)	(0.0162)	(0.0188)	(0.0151)
Ln(middle school grade)	0.0299**	0.0595***	0.1227***	-0.1999***	0.2143***	0.2413***
	(0.0131)	(0.0207)	(0.0270)	(0.0233)	(0.0397)	(0.0386)
Ln(ms grade)*MIX	0.0465**	0.0857***	0.0772**	0.0184	0.0243	0.0075
	(0.0235)	(0.0268)	(0.0371)	(0.0272)	(0.0477)	(0.0487)
Male	0.0104	0.0185*	0.0290***	-0.0806***	0.0870***	0.1170***
	(0.0074)	(0.0102)	(0.0108)	(0.0095)	(0.0242)	(0.0260)
Received tutoring when MS student	-0.0125*	0.0167	0.0055	-0.0574***	0.0504	0.0454
	(0.0073)	(0.0164)	(0.0351)	(0.0218)	(0.0349)	(0.0363)
General high school student(GHS)	0.0090***	0.0338***	0.0590***	-0.0984***	0.0719**	0.0897***
	(0.0033)	(0.0077)	(0.0118)	(0.0087)	(0.0280)	(0.0324)
City, Age, HS year FEs and City FEs*MIX	Y	Y	Y	Y	Y	Y
High school fixed effects					Y	Y
Observations	2,267	2,267	2,267	2,236	2,168	1,224
R-squared	0.0758	0.1045	0.1331	0.2740	0.6972	0.5994
Test-statistic:						
$t = \hat{\beta}_{mix}\hat{\gamma}_{track} - \hat{\gamma}_{mix}\hat{\beta}_{track}$	0.0002	0.0010	0.0028	0.0079**	0.0026	0.0019
	(0.0003)	(0.0007)	(0.0020)	(0.0033)	(0.0052)	(0.0047)
p-value of one sided test: $H_0: \frac{\hat{\beta}_{track}}{\hat{\gamma}_{track}} = \frac{\hat{\beta}_{mix}}{\hat{\gamma}_{mix}}, \quad H_1: \frac{\hat{\beta}_{track}}{\hat{\gamma}_{track}} < \frac{\hat{\beta}_{mix}}{\hat{\gamma}_{mix}}$	0.2297	0.0857*	0.0769*	0.0085***	0.3065	0.3472

Note: Analysis is based on the sample of individuals who entered high school between 1970 and 1985. ***, **, and * indicates significance at the 1%, 5%, and 10% level, respectively. City clustered Huber-White standard errors are in parantheses.

Table 8. Border discontinuity estimates using the full sample

Dependent variables	District 8				Unified Central District			
	Difference in log prices from				Difference in log prices from			
	1971 to 73	1973 to 75	1975 to 77	1977 to 79	1971 to 73	1973 to 75	1975 to 77	1977 to 79
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Housing land								
	Medium quality housing land							
Inside district	0.007	0.576***	0.292*	-0.098	0.021	0.151*	-0.265**	0.124
	(0.083)	(0.151)	(0.154)	(0.194)	(0.066)	(0.070)	(0.101)	(0.214)
Observations	387	361	369	373	387	361	369	373
R-squared	0.149	0.413	0.033	0.307	0.040	0.337	0.122	0.403
	High quality housing land							
Inside district	-0.082	0.485***	0.283*	-0.034	-0.010	0.003	-0.217***	0.207
	(0.070)	(0.139)	(0.131)	(0.204)	(0.082)	(0.119)	(0.039)	(0.122)
Observations	420	420	427	416	420	420	427	416
R-squared	0.067	0.413	0.063	0.305	0.035	0.279	0.102	0.441
	Low quality housing land							
Inside district	0.109	0.588*	0.287	0.160	0.166	-0.043	-0.193	0.044
	(0.216)	(0.300)	(0.319)	(0.177)	(0.157)	(0.158)	(0.122)	(0.155)
Observations	343	311	319	281	343	311	319	281
R-squared	0.117	0.084	0.119	0.225	0.039	0.155	0.080	0.266
Panel B: Commercial land								
	Medium quality commercial land							
Inside district	0.555***	0.031	-0.031	0.444**	0.025	0.138	-0.321**	0.159
	(0.154)	(0.262)	(0.279)	(0.184)	(0.051)	(0.117)	(0.115)	(0.157)
Observations	269	261	281	329	269	261	281	329
R-squared	0.016	0.213	0.061	0.267	0.004	0.290	0.086	0.236
	High quality commercial land							
Inside district	0.252**	0.582***	0.029	0.569***	0.001	0.030	0.046	0.087
	(0.093)	(0.076)	(0.116)	(0.114)	(0.047)	(0.123)	(0.096)	(0.147)
Observations	309	335	364	389	309	335	364	389
R-squared	0.044	0.334	0.044	0.201	0.005	0.321	0.028	0.195
	Low quality commercial land							
Inside district	-0.692	0.235	0.656**	-0.091	0.076	-0.115	-0.201	0.025
	(0.429)	(0.148)	(0.259)	(0.479)	(0.066)	(0.089)	(0.117)	(0.205)
Observations	240	226	259	269	240	226	259	269
R-squared	0.046	0.195	0.029	0.210	0.041	0.286	0.048	0.324

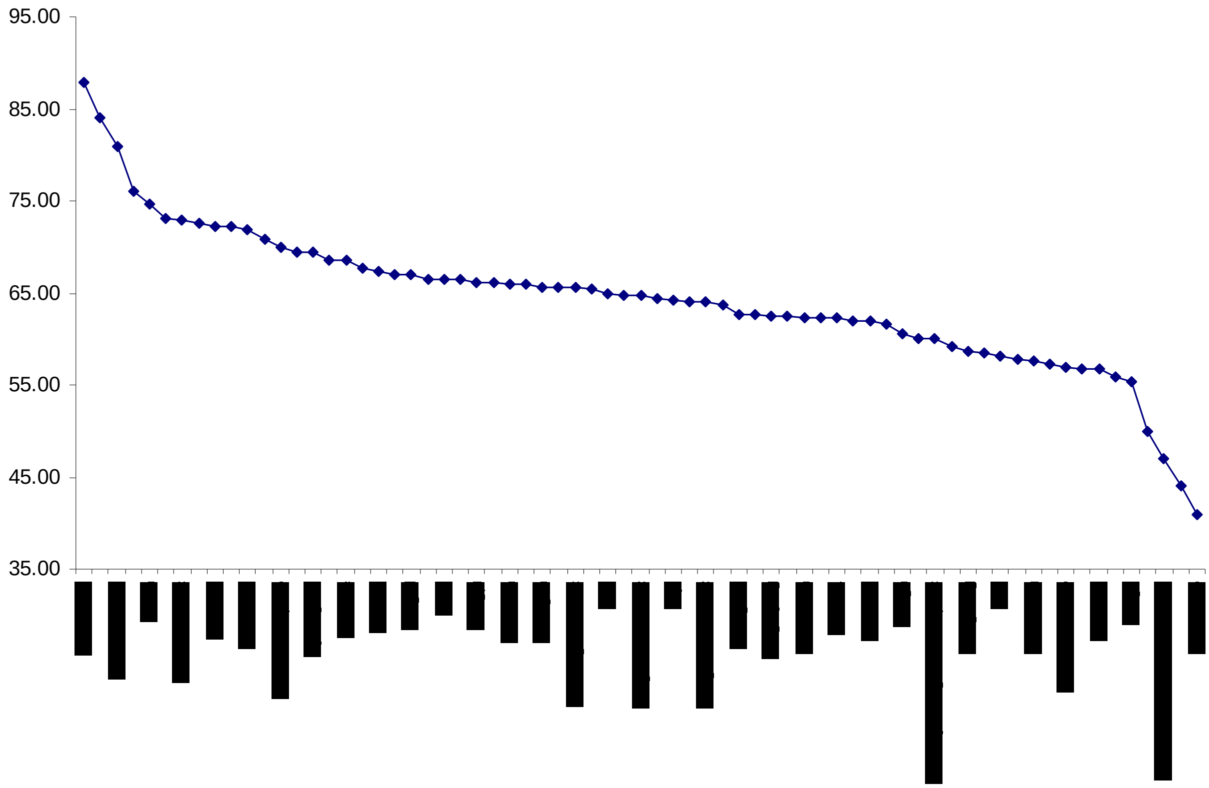
Notes: Running variable is integer distance from the district boundary. I use fixed polynomials on both sides of the boundary. 4th order polynomials are used for housing land prices and 3rd order polynomials are used for commercial land prices except for change in 1971-73 prices, in which quadratic is used because of lower observation counts. Administrative district (gu) clustered Huber-White standard errors are in parantheses. ***, **, and * indicates significance at the 1%, 5%, and 10% level, respectively.

Table 9. Border discontinuity estimates when sample is trimmed within 5km of district boundaries

Dependent variables:	District 8				Unified Central District			
	Difference in log prices from				Difference in log prices from			
	1971 to 73	1973 to 75	1975 to 77	1977 to 79	1971 to 73	1973 to 75	1975 to 77	1977 to 79
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Housing land								
	Medium quality housing land							
Inside district	-0.222	0.727**	0.140	-0.483*	0.146**	0.106	-0.221	0.352
	(0.135)	(0.311)	(0.109)	(0.221)	(0.057)	(0.075)	(0.129)	(0.222)
Observations	162	152	157	141	307	285	288	279
R-squared	0.217	0.471	0.089	0.317	0.022	0.289	0.071	0.384
	High quality housing land							
Inside district	-0.337***	0.572	0.012	-0.227	0.114	0.007	-0.203**	0.296*
	(0.056)	(0.347)	(0.070)	(0.142)	(0.078)	(0.153)	(0.079)	(0.157)
Observations	172	170	176	166	326	325	332	321
R-squared	0.111	0.538	0.073	0.301	0.050	0.213	0.070	0.368
	Low quality housing land							
Inside district	0.266	0.659*	0.307	-0.361	0.375*	-0.075	-0.414***	0.092
	(0.328)	(0.340)	(0.269)	(0.330)	(0.176)	(0.254)	(0.110)	(0.180)
Observations	132	115	121	121	276	234	227	197
R-squared	0.173	0.204	0.081	0.230	0.058	0.180	0.025	0.266
Panel B: Commercial land								
	Medium quality commercial land							
Inside district	0.005	1.263	0.418	0.192	-0.026	-0.159	-0.225	0.403
	(1.145)	(0.992)	(0.363)	(0.223)	(0.089)	(0.230)	(0.229)	(0.230)
Observations	126	131	141	168	245	236	245	285
R-squared	0.093	0.347	0.010	0.422	0.026	0.239	0.097	0.180
	High quality commercial land							
Inside district	0.328**	1.193*	0.215**	0.548**	0.024	0.110	-0.022	0.227
	(0.103)	(0.605)	(0.088)	(0.228)	(0.063)	(0.122)	(0.129)	(0.202)
Observations	145	155	167	180	278	295	317	336
R-squared	0.172	0.516	0.101	0.421	0.010	0.234	0.028	0.122
	Low quality commercial land							
Inside district	-1.133	1.186	0.339	0.155	-0.014	-0.038	-0.338*	0.229
	(1.307)	(0.697)	(0.300)	(0.262)	(0.061)	(0.070)	(0.170)	(0.315)
Observations	123	115	127	143	218	204	229	236
R-squared	0.048	0.457	0.035	0.226	0.025	0.245	0.059	0.248

Notes: Running variable is integer distance from the district boundary. I fit quadratic polynomials and functional form to change on both sides of the boundary. Administrative district (gu) clustered Huber-White standard errors are in parantheses. ***, **, and * indicates significance at the 1%, 5%, and 10% level, respectively.

Appendix Figure 1. Distribution of Average Test Scores of Admitted Students by College



Appendix Table 1. Shift from tracking to mixing by city and year

Group	Policy year	City	City type	Obs.
Group 1	1974	Seoul, Busan	Capital and 2nd largest city	676
Group 2	1975	Daegu, Inchon, Gwangju	Cities with population over 1,000,000 in 1975	320
Group 3	1979	Daejeon, Suwon, Masan, Jeonju, Jeju, Chongju, Chuncheon	Province capitals	190
Group 4	1980	Jinju, Changwon, Andong, Mokpo, Gunsan, Iksan, Wonju, Chonan	Other major regional cities	203
Group5 (Others)	No policy change	All other regions		1067

Note: Among the Group 5 cities some shift to mixing after 2000. However, this period the policy analysis is confiscated by the plethora of Elite Special Purpose Schools. Hence I focuse on the policy change that occur during the mid to late 1970s.

Appendix Table 2. University scores used in the analysis

College	Score	College	Score	College	Score
POSTECH	87.90	Junnam	66.40	Samyuk	62.22
KAIST	84.00	Gongju	66.12	Other 4yr college in Seoul	62.00
Seoul National	80.93	Bukyung	66.08	Hansung	61.96
Yonsei	76.05	Kookmin	65.95	Gangwon	61.68
Korea	74.66	Sookmyung Women's	65.91	Donga	60.55
Sogang	73.17	Gwangun	65.66	Gyeongsung	60.06
Ewha Women's	72.85	Chungnam	65.60	Other 4yr college in Metropolis	60.00
Busan	72.63	Duksung Women's	65.56	Ulsan	59.13
Hanyang	72.28	Myungji	65.42	Wongwang	58.59
Korea Foreing	72.17	Inha	64.87	Sangjji	58.46
Seoul City	71.92	Seoul Women's	64.70	Silla	58.13
Sunkyunkwan	70.88	Sungshin Women's	64.67	Gyemyung	57.80
Korea Aerospace	69.91	Junbuk	64.36	Cheonggju	57.59
Joongang	69.49	Jeju	64.24	Gwandong	57.28
Gyeongbuk	69.38	Sejong	64.00	Other 4yr college	57.00
Kyunghee	68.63	Dongduk Women's	63.99	Chosun	56.82
Catholic	68.59	Chongshin	63.74	Mokwon	56.79
Dongkuk	67.75	Chungbuk	62.73	Baeje	55.80
Dankuk	67.29	Gyungsang	62.68	Daegu	55.41
Konkuk	67.03	Sangmyung	62.56	2yr Tech college in Seoul	50.00
Sungsil	66.97	Seogyung	62.54	2yr Tech college in Metropolis	47.00
Hongik	66.51	Youngnam	62.35	Other 2yr Tech college	44.00
Ajoo	66.47	Gyeonggi	62.34	No college	41.00

Appendix Table 3. The Location Change of the Top 10 High Schools in Seoul

High School	Location in 1974	Present Location	Year of Move
Top 5 Public High Schools			
Gyeonggi High School	Jongro	Gangnam	1976
Gyeongbok High School	Jongro	Jongro	
Seoul High School	Seodaemun	Seocho	1980
Yongsan High School	Yongsan	Yongsan	
Kyeongdong High School	Seongbuk	Seongbuk	
Top 5 Private High Schools			
Joongang High School	Jongro	Jongro	
Yangjung High School	Jongro	Yangcheon	1988
Baejae High School	Jongro	Gangdong	1984
Hweemun High School	Jongro	Gangnam	1978
Bosung High School	Jongro	Songpa	1989

Appendix Table 4. Effect of the education regime shift on tutoring choice and attending a general high school - Logit Model

Dependent variable:	Received private tutoring when HS student				Attended a GHS	
	(1)	(2)	(3)	(4)	(5)	(6)
City shifts to mixing (MIX)	0.9905	1.0884	1.2964	1.2423	-0.8230*	-0.8337*
	(0.9206)	(0.9578)	(1.0567)	(1.1655)	(0.4748)	(0.4812)
Father's years of education	0.1363**	0.1291**	0.1607**	0.1764***	0.0758***	0.0714***
	(0.0598)	(0.0625)	(0.0625)	(0.0678)	(0.0176)	(0.0179)
Father's education*MIX	-0.0631	-0.0699	-0.0931	-0.1165	0.0462*	0.0497*
	(0.0581)	(0.0640)	(0.0594)	(0.0721)	(0.0266)	(0.0277)
Middle school grade	0.1893**	0.1659**	0.1680***	0.1433**	0.1101***	0.1046***
	(0.0755)	(0.0740)	(0.0640)	(0.0573)	(0.0342)	(0.0334)
Middle school grade*MIX	-0.0863	-0.0758	-0.0738	-0.0386	0.0695**	0.0724**
	(0.0827)	(0.0860)	(0.0773)	(0.0778)	(0.0336)	(0.0345)
Male	-0.5999***	-0.5686***	-0.2857	-0.2398	-0.1012	-0.0884
	(0.1111)	(0.1184)	(0.1822)	(0.2008)	(0.1424)	(0.1457)
Received tutoring when MS student		3.0500***		3.7658***		0.9751**
		(0.2270)		(0.3213)		(0.4763)
General high school student(GHS)		0.0987				
		(0.2822)				
City, Age, HS year FEs and City FEs*MIX	Y	Y	Y	Y	Y	Y
Observations	1,626	1,626	993	993	2,092	2,092

Note: Analysis is based on the sample of individuals who entered high school between 1970 and 1985. ***, **, and * indicates significance at the 1%, 5%, and 10% level, respectively. City clustered Huber-White standard errors are in parantheses.

Appendix Table 5. Interaction term analysis of achievement with 1980 as the hypothetical policy year for control group

Coefficient of	Log(father's education)* 1(treated cities)* 1(policy cohort N)	Log(ms grade)* 1(treated cities)* 1(policy cohort N)	Log(father's education)* 1(policy cohort N)	Log(ms grade)* 1(policy cohort N)
	(1)	(2)	(3)	(4)
Dependent variable: ln(college score)				
Policy Cohort N= -9, -10	0.1093*	-0.0653	-0.0557	0.2460***
	(0.0650)	(0.0794)	(0.0373)	(0.0441)
Policy Cohort N= -7, -8	-0.0873	0.0879	0.0355	0.1577***
	(0.0537)	(0.0740)	(0.0322)	(0.0391)
Policy Cohort N= -5, -6	0.0442	-0.0185	-0.0114	0.1996***
	(0.0384)	(0.0707)	(0.0310)	(0.0365)
Policy Cohort N= -3, -4	0.0494	-0.0456	0.0109	0.2146***
	(0.0328)	(0.0665)	(0.0247)	(0.0346)
Policy Cohort N= -1, -2	0.0188	-0.0061	-0.0023	0.2042***
	(0.0286)	(0.0652)	(0.0190)	(0.0329)
Policy Cohort N= 1, 2	0.0033	0.0618	0.0077	0.1958***
	(0.0259)	(0.0413)	(0.0215)	(0.0345)
Policy Cohort N= 3, 4	0.0140	0.0597	0.0411*	0.1856***
	(0.0304)	(0.0450)	(0.0209)	(0.0313)
Policy Cohort N= 5, 6	0.0453**	0.0193	0.0138	0.1856***
	(0.0220)	(0.0418)	(0.0184)	(0.0346)
Policy Cohort N= 7, 8	0.0426	0.0299	0.0076	0.2113***
	(0.0410)	(0.0519)	(0.0281)	(0.0372)
Policy Cohort N= 9, 10	0.0984***	-0.0154	-0.0241	0.2304***
	(0.0322)	(0.0466)	(0.0280)	(0.0385)
Policy Cohort N= 11, 12	0.0590	0.0028	0.0223	0.1873***
	(0.0391)	(0.0473)	(0.0338)	(0.0425)
R-squaured	0.41			
Observations	2236			

Notes: Negative policy cohorts imply cohorts before the regime shift. Analysis is based on the sample of individuals who entered high school between 1970 and 1985. Dummy variables for gender, middle school tutoring, and general high school are included as before. Also city, age, high school year fixed effects, and city fixed effects interacted with the policy variable are included. ***, **, and * indicates significance at the 1%, 5%, and 10% respectively. City clustered Huber-White standard errors are in parentheses.

Appendix Table 6. Border discontinuity estimates using the narrower windows - 3km windows for District 8 and 2km windows for Unified Central District

Dependent variabl	District 8				Unified Central District			
	Difference in log prices from				Difference in log prices from			
	1971 to 73	1973 to 75	1975 to 77	1977 to 79	1971 to 73	1973 to 75	1975 to 77	1977 to 79
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Housing land								
	Medium quality housing land							
Inside district	-0.098	0.920**	0.188	-0.171	0.098	0.105	-0.224*	0.149
	(0.130)	(0.188)	(0.100)	(0.112)	(0.056)	(0.096)	(0.121)	(0.219)
Observations	58	59	60	59	233	217	218	205
R-squared	0.131	0.385	0.139	0.295	0.046	0.120	0.085	0.192
	High quality housing land							
Inside district	-0.217**	0.711*	0.089	-0.121	0.070	-0.036	-0.201**	0.250*
	(0.061)	(0.266)	(0.060)	(0.111)	(0.068)	(0.133)	(0.069)	(0.130)
Observations	59	61	61	61	246	243	248	235
R-squared	0.061	0.460	0.083	0.329	0.091	0.124	0.079	0.184
	Low quality housing land							
Inside district	0.418	0.808*	0.329	-0.223	0.268*	-0.052	-0.183	0.031
	(0.323)	(0.316)	(0.284)	(0.125)	(0.142)	(0.174)	(0.106)	(0.146)
Observations	51	54	58	58	199	162	157	138
R-squared	0.072	0.157	0.046	0.269	0.083	0.128	0.029	0.199
Panel B: Commercial land								
	Medium quality commercial land							
Inside district	-0.119	1.344**	-0.247	0.351*	0.063	0.015	-0.282*	0.255
	(0.223)	(0.182)	(0.113)	(0.153)	(0.085)	(0.113)	(0.133)	(0.188)
Observations	23	28	34	45	192	182	186	217
R-squared	0.258	0.342	0.102	0.485	0.042	0.160	0.071	0.115
	High quality commercial land							
Inside district	0.179	1.182***	-0.237*	0.500	0.019	0.024	0.038	0.160
	(0.394)	(0.024)	(0.092)	(0.251)	(0.044)	(0.097)	(0.097)	(0.170)
Observations	33	39	46	51	212	222	235	250
R-squared	0.259	0.347	0.110	0.490	0.009	0.124	0.034	0.065
	Low quality commercial land							
Inside district	-1.028*	1.191**	0.021	0.103	0.031	-0.150*	-0.233*	0.148
	(0.367)	(0.140)	(0.130)	(0.147)	(0.041)	(0.080)	(0.119)	(0.248)
Observations	26	24	31	40	166	148	160	171
R-squared	0.009	0.504	0.217	0.208	0.005	0.149	0.069	0.184

Notes: Running variable is integer distance from the district boundary. I fit linear functions and allow slopes to change on both sides of the boundary. Administrative district (gu) clustered Huber-White standard errors are in parantheses. ***, **, and * indicates significance at the 1%, 5%, and 10% level, respectively.