Does Human Capital Spillover Beyond Plant Boundaries?: Evidence from Korea

WooRam Park *

November 2012

Abstract

Human capital externalities exist if plants in a region with a high level of human capital can produce more—given similar inputs—than plants in a region with less human capital. It is difficult to assess this assertion because human capital levels are endogenous. To address this issue, this paper explores the July 30 Education Reform in South Korea, that was initiated in 1981 for political reasons. This reform discretely and mechanically increased the number of students that each college was allowed to admit based on the college's enrollment prior to the implementation of the policy. As a result, this reform exogenously increased cross-region differences in the supply of college graduates starting in the mid 1980s. Using annually collected plant level data, I explore the effect of changes in human capital levels induced by this reform on plant productivity. My results suggest that this effect is limited. I find a correlation between level of human capital and plant productivity which is similar to the correlation observed in the U.S. However, after addressing the endogeneity by using an instrumental variable, the effect of the overall level of human capital on productivity decreases and becomes statistically insignificant. Moreover, consistent with the results using the value added of plants as an outcome, the wages of workers are not higher in regions with higher levels of human capital.

JEL code: J30, L30, O40

Keywords: human capital externalities, plant productivity, exogenous variation, instrumental variable, college education, Korea

*Ph.D. Candidate, Department of Economics, Columbia University; 420 West 118th Street, New York, NY 10027; Contact at wp2135@columbia.edu. I am very grateful for support, guidance, and many helpful comments from Miguel Urquiola, Bentley MacLeod and Eric Verhoogen. I thank Giacomo De Giorgi, Jonas Hjort, Supreet Kaur, Suresh Naidu, Wojciech Kopczuk, Cristian Pop-Eleches and Till von Wachter for the suggestions and the participants of the Columbia University Development Workshop and Applied Microeconomics Colloquium for their comments. All errors are my own.

1 Introduction

Whether human capital externalities raise productivity is of interest to economists and policymakers, because this issue has substantial implications regarding government subsidies for education and economic growth (Lucas, 1988).[1] However, there is no consensus on the empirical magnitude of human capital externalities. In particular, there is no agreement on whether they are large enough to help explain cross-regional differences in growth rates, or to justify subsidies for education.[2] For instance, Rauch (1993) and Moretti (2004a) find positive and sizable human capital externalities on productivity, whereas Acemoglu and Angrist (2000) and Rudd (2000) find little evidence that these human capital externalities are significant in practice.[3]

In part, these conflicting results may reflect that the literature has struggled to deal with endogeneity. In particular, reliable evidence on magnitude of human capital externalities is scarce due to several empirical challenges. First, the level of human capital in a given region is usually endogenous since it is likely to be correlated with unobserved characteristics that affect productivity in the region. Additionally, changes in human capital levels are in most cases gradual; and thus, sufficient variation for reliable estimation is hard to find. This matter exacerbates endogeneity concerns when the literature uses within region variation, since even a small amount of endogeneity could bias results by a considerable amount.

In this paper, I provide new evidence on the magnitude of human capital externalities

[1] A substantial body of theoretical literature suggests possible mechanisms behind human capital externalities. Some papers propose spillovers through personal interactions. This idea at least goes back to Marshall (1890) and is more recently suggested by Arrow (1962) and Romer (1986). According to this view, geographical proximity between employees is an important condition for human capital spillovers, as the exchange of ideas through personal interaction is assumed to be decreasing with distance. This view on the mechanism of human capital spillovers, has been used to argue for the importance of cities as engines of economic growth (Jacobs, 1970). More recently, Niehaus (2012) argues that the increase in education levels will lead to knowledge spillovers by increasing workers' ability to learn skills from other workers. Alternatively, Acemoglu (1996) proposes that the increase in human capital could have a positive external effect on productivity without involving technology when there is a costly search between workers and firms. This type of human capital externality does not necessarily involve knowledge spillovers.

[2] Other types of social return to human capital documented by the literature include, effects on non-economic outcomes such as citizenship (Dee, 2004; Milligan, Moretti, & Oreopoulos, 2004), health of one's children (Currie & Moretti, 2003) and criminal activity (Lochner & Moretti, 2004).

[3] Several papers document the positive relation between productivity and average years of schooling using cross-country data (de la Fuente & Domenech, 2001). However, cross-country evidence is unlikely to reveal the magnitude of human capital externalities since average levels of human capital are likely correlated with characteristics that can affect productivity (e.g. social infrastructure) (Hall & Jones, 1999).

on plant productivity. That is, I examine whether a given plant's productivity is affected by the human capital level of other plants in nearby locations. In particular, I closely follow the methodology of Moretti (2004c) but use a new instrument for human capital levels, one based on arguably exogenous variation in the supply of college graduates across regions induced by the July 30 Education Reform in Korea.

Specifically, in 1980, the Korean government implemented an education reform that resulted in a large and discrete increase in the number of students entering college in 1981. Between 1979 and 1981, the number of freshmen increased from 98,000 to 190,000. This college-level increase in freshmen enrollment was mechanically determined based on the enrollment prior to the reform. In particular, for every college, freshmen enrollment in 1981 was generally *proportional* to freshmen enrollment in 1979. Thus, cross-college differences in enrollment increased discretely as the colleges which had a large freshmen enrollment in 1979 experienced larger absolute increases in enrollment as a result of the reform. Similarly, regions with larger initial freshmen enrollment-levels experienced larger increases in enrollment. This discrete increase in cross-region variation in college enrollment resulted in an increase in the cross-region variation in the supply of college graduates starting in the mid 1980s. Due to this fact, there were rapid increases in the proportion of workers with college degrees in each region starting in the mid 1980s, and there was a sizable variation in the change in this proportion across regions as well.

I construct an instrumental variable which extracts the portion of the region-specific change in the level of human capital predicted by this exogenous reform. That is, I rely on the fact that the reform had a larger impact on regions with higher initial enrollments. My broad approach is closely related to Moretti (2004a) who uses the presence of land grant colleges as an instrument for the human capital level in a given region. However, while Moretti (2004a) uses cross-sectional variation, I exploit the arguably exogenous *increase* in college enrollment to instrument for region-specific *increases* in the level of human capital. I provide further evidence on the exogeneity of this change in human capital by showing that it is positively correlated with the initial freshmen enrollment levels only after the impact of the reform. This evidence helps to address concerns, for example that the impact of the reform was correlated with pretrends, or that the reform simply amplified developing differences in the trends of human capital levels across regions.

I implement this idea using the 1982-1996 data from the Mining and Manufacturing Survey which is collected by Statistics Korea. The data provide detailed information on output, labor and capital, and other plant-specific characteristics such as ownership type, age, industry and location. In particular, by estimating a production function at the plant level, I examine whether region-specific increases in the share of college graduates had a positive effect on plant productivity after controlling for plant-specific inputs and characteristics.

Overall, I find little evidence of human capital spillovers beyond plant boundaries—after controlling for plants' own level of human capital, the proportion of workers who are college graduates in a given region does not have a meaningful effect on plant productivity. The magnitude of the simple correlation between the regional level of human capital and plant productivity is similar to that observed using plant level data in the U.S. In particular, pooled regressions suggest that a one percentage point increase in the proportion of workers who are college graduates is associated with a 0.7 percentage increase in productivity. However, after instrumenting for the human capital level, the effect decreases and becomes statistically insignificant. The results from the instrumental variable analysis show that there is a positive bias in the correlation between the level of human capital and productivity. As a robustness check, I also analyze workers' wages rather than plant productivity as an outcome variable. The results again do not support empirically meaningful externalities. Overall, the findings are consistent with the recent paper by Huber (2012) which questions the presence of human capital spillovers beyond establishment boundaries. In particular, by surveying workers in the R&D complex in England, he finds that the workers have limited interactions with the workers outside their establishment; internal resources are sufficient and preferable for innovation.

To my knowledge my paper is one of few papers to examine human capital externalities using plant level data while addressing the endogeneity of regional human capital. Previous literatures on human capital externalities mainly uses wages as an outcome variable (Acemoglu & Angrist, 2000; Moretti, 2004a; Iranzo & Peri, 2009).[4] The idea that if human

[4] Lange and Topel (2006) and Moretti (2004b) provide a good summary of the literature which uses wage data to document the social returns of education.

capital externalities exist, workers in a region with a higher level of human capital will be more productive and thus they will earn higher wages, forms the basis for the papers in question. However, Moretti (2004a) and Ciccone and Peri (2006) describe, interpreting the results is complicated due to the imperfect substitution between skilled and unskilled workers. This difficulty along with the downward sloping demand for the human capital could yield a positive relation between the level of human capital and average wages even without the existence of externalities. Thus, the literature often fails to give conclusive implications on the magnitude of externalities. Moretti (2004c) is one of few exceptions as using plant level output allows results to have a more direct bearing on the magnitude of human capital externalities. In constrast to my findings in Korea, Moretti (2004c) finds sizable human capital spillovers beyond plant boundaries in the manufacturing sector in the U.S.

My paper also contributes to the literature on the subject by examining possible externalities resulting from college education, which could be different from externalities from K-12 education (Iranzo & Peri, 2009). Finally, this paper adds to the literature by exploring the effect of human capital externalities in the context of a developing country. Particularly, despite the widespread belief regarding the importance of human capital externalities on plant productivity in the growth of South Korea—e.g. Lucas (1988, 1993)—little is known about their empirical magnitude in South Korea. My results suggest the human capital spillovers beyond establishment boundaries might not have been a crucial factor in Korea's growth during the 1980s and 1990s.

The rest of this paper is organized into the following sections. Section 2 introduces the institutional background and Section 3 describes the data set. Section 4 presents the identification strategy. Section 5 presents results and a series of robustness checks. In Section 6, the magnitude of the externalities is examined using the wages of workers as an outcome variable. The last section discusses the conclusion.

2 Institutional Background

Korea offers a unique institutional setting for this type of study in that the central government controls the supply of college graduates by setting the freshmen quota, or entrance quota, for *both* private and public colleges.[5] The freshmen quota was strictly enforced during 1970s through 1980s and colleges faced severe penalties for admitting freshmen beyond the assigned quota, such as the loss of government funding and a decrease in their quota for the following years.[6] Moreover, the government also controlled the number of colleges by granting permission for the establishment of new institutions. The number of colleges was more or less stable across regions during the period of interest.[7] In short, this setting was quite different from countries such as the U.S. in which college enrollment is not set in centralized manner. Because of way that the college enrollment was determined in Korea, the supply of college education was less likely to be responsive to a time-varying region-specific characteristics.

Until 1980, the government only allowed a gradual increase in the freshmen quota despite a large increase in the demand for college education in 1970s. As a result, the number of 'repeat applicants', who were forced to apply to colleges repeatedly for more than one year to receive higher education, accumulated as the freshmen quota was not sufficient to accommodate all of the students who wanted to enter college.

However, in 1981 the freshmen quota discretely jumped due to an unexpected education reform announced on July 30, 1980 (Choi, 1996).[8] The main purposes of the reform were: i) to increase the probability that students from disadvantaged backgrounds would receive college education, and ii) to reduce the number of 'repeat applicants'. The major component of the reform was a discrete increase in the freshmen quota to accommodate more students.[9] Figure 1 describes the reform's mandated increase in the freshmen quota and corresponding increase in the freshmen enrollment in 1981. It is clear that this large increase was a onetime event, as the freshmen quota was stable during the 1980s after the initial increase in 1981.

[5]The government decided not only the freshmen quota but also the admission guidelines for *both* private and public colleges.

[6]The government allowed only a small number of disadvantaged students to be accepted over the freshmen quota. Further, the government provided an incentive for colleges to keep the actual enrollment lower than the freshmen quota. For instance, the government increased the subsidy for colleges if the actual enrollment for a given college was lower than the freshmen quota set by the government.

[7]The government eventually relaxed the regulation for establishing new colleges in 1996.

[8]President Park Chung-hee who was in office for over 15 years was assassinated by his body guard on October 26, 1979. After the assassination, the military junta lead by General Chun Doo-hwan, gained control after a series of coups. Many people hoped a democratic government would be established after the assassination, and, as a result, this military junta was not popular. To gain popularity, the junta announced the education reform on July 30, 1980.

[9]Other components include prohibiting private tutoring, and abolishing college-specific entrance exams.

More importantly, the central government forced each and every college to increase the freshmen quota in essentially the same way. That is, the magnitude of increase was not endogenously adapted to each college to accommodate region-specific demand for higher education. In particular, the implementation of the increase in the freshmen quota was more or less mechanical—in general, the government set the freshmen quota in 1981 for each college proportionally to the freshmen quota of previous years. Thus, the colleges which happened to have a large freshmen quota in 1979 experienced larger absolute increases by 1981. Figure 2 plots the freshmen quota for each college in 1979 and 1981 along with a 45 degree line. This figure shows that the relation between the freshmen quota in 1981 and in 1979 is linear, suggesting that the freshmen quota in 1979 primarily determined the increase. The figure also shows that the absolute differences in enrollment increase; the gap between the 45 degree line and the freshmen quota in 1981 increase as the initial freshmen quota increases.

Since the reform was consistently applied to each college, the relationship between the freshmen quota in 1979 and in 1981 in each region is similar. Figure 3 describes the correlation between the freshmen quota in 1979 and in 1981 in each region along with the 45 degree line. By comparing the freshmen quota in 1981 with the 45 degree line, one can see that regions which happened to have a higher freshmen enrollment in 1979 experienced larger absolute increases by 1981. Thus, the reform exogenously increased the difference in the supply of college graduates in each region after the mid-1980s. Furthermore, this figure confirms that the freshmen quota in 1981 in each region was indeed mostly determined by the proportional increase in the freshmen quota in 1979—the relation between the quotas in 1979 and in 1981 is linear. The figure thus provides evidence against the claim that the increase in the freshmen quota in 1981 was endogenously determined by the government. In fact, Chun Doo-Hwan, the head of the military junta, did not favor his hometown regions—Gyeongbuk and Daegu—by increasing their quotas more than in other regions.

As a result of these changes in the supply of college graduates across regions, there is a substantial cross-regional variation in the change in the proportion of the labor force composed of college graduates. To see this variation, one supposes that there are two regions: A and B where each region has 1000 people entering the workforce each year. One then assumes that the initial freshmen enrollments—which could be endogenous—in region A and region B are 200 and 100, respectively. If there is no inter-regional mobility, in the long run the proportion of college graduates in the workforce in region A and region B would be 0.2 and 0.1, respectively.[10] Now, one supposes that the government imposed a reform that forces each college to increase its enrollment by 100%. Then, the college enrollment in region A and region B will increase to 400 and 200, respectively. As a result of the reform the proportion of college graduates in the workforce in each region will increase to 0.4 and 0.2. This example demonstrates that the region with a larger freshmen quota would have experienced a larger increase in share of workers with college degrees due to the reform.

The validity of the reform as a quasi-experiment, therefore, depends on whether the change in human capital levels across regions are indeed induced by the reform. I provide additional evidence on issue by showing that the trend in the share of college graduates in each region is positively correlated with the freshmen quota—which determined the size of the impact of the reform—only after the reform. Since the reform increased the existing difference in the freshmen quota across regions discretely, the regions with a relatively large freshmen quota are expected to experience a larger increase in the supply of college graduates, and thus a larger increase in share of workers who are college graduates.[11] If the change in the level of human capital was truly the result of the reform, the proportion of college graduates among workers in the regions which had a larger number of freshmen in 1979 should have increased more than in other regions *only* after the mid-1980s. Otherwise, the variation in the increase in the share of college graduates could simply reflect unobserved underlying trends and/or region-specific shocks.

Figure 4 illustrates the correlation between the size of the initial freshmen quota—measured by the freshmen quota in 1979 divided by the number of total employees in each region—and the trend in the proportion of college graduates prior to the reform which is measured by the change in proportion of white-collar workers between 1982 and 1986. There is little correlation between the pre-trend and the size of the freshmen quota in 1979—the correlation is 0.008.

[10] The analysis assumes 100% graduation rates.

[11] As a result of the mandatory military service in Korea which usually last 2-3 years, students who entered college in 1981 entered the job market between 1986 and 1988.

This correlation implies that the regions which had a larger freshmen quota in 1979 did not experience larger increases in human capital prior to the impact of the reform, that is prior to the exogenous increase in the supply of college graduates induced by the reform.

Figure 5 describes the correlation between the size of the initial freshmen quota and the change in the level of human capital after the mid 1980s—measured by the change in the proportion of the college educated labor force between 1988 and 1992. The figure confirms that higher initial freshmen quotas are correlated with a greater increase in the proportion of workers with college degrees after the reform—the correlation is 0.6118. Moreover, Figure 4 and 5 confirm a large cross-region variation in the change of the share of college graduates after the mid 1980s. The change in share of college graduates among workers prior to the reform in each region—described in y-axis of Figure 4—varies from -0.002 to 0.033 with the standard deviation 0.0096. The change in the share of college graduates after the impact, however, varies from 0.034 to 0.0963 with the standard deviation equal to 0.017.

Overall, these findings support the argument that the change in the proportion of college graduates across regions was driven by the education reform, and that the variation in the impact of the policy across regions was driven by the predetermined characteristics of each region.

3 Data

To examine human capital externalities on plant productivity, I use the Mining and Manufacturing Survey provided by Statistics Korea. Statistics Korea has been collecting data since 1968, but the micro data are only available starting in 1982. Also, since the manufacturing sector of Korea was heavily affected by the Asian financial crisis in 1997, I only use data prior to 1997. These data have been collected annually from mining and manufacturing plants which have five or more workers. The survey contains detailed information about individual plants such as industry classification, output, production cost, location, and tangible assets including capital. The data also contain information on the total number of employees and the number of white-collar (non-production) employees. However, like most plant level data, there is no information on the educational attainment level of workers. Thus, I proxy the change in the proportion of college graduates by the change in the proportion of white-collar workers.

To explore the validity of this proxy, I use the Basic Wage Structure Survey. These data have been collected by the Ministry of Employment and Labor of Korea and are designed to represents the wages of workers in establishments which hire more than 10 employees. The survey collects data from individual workers from a sample of establishments representing each sector. The data contain information on wages, education, occupation and industry.[12] Using these data, I show that the trend of workers with college degrees coincides with the trend of proportion of white-collar among workers. Also, I provide evidence that the increase in the proportion of college graduates actually induced the increase in proportion of white-collar workers.

Figure 6 displays the share of college graduates and white-collar among workers in the manufacturing sector using the Basic Wage Structure Survey and Mining and Manufacturing Survey. The time trend for white-collar workers tracks the trend of college graduates closely; the proportion of both trends shows little increase until the mid-1980s and then starts rising steeply after 1987. I provide an additional justification for using white-collar workers as a proxy for workers with college degrees by separately plotting the proportion of college graduates and white collar workers for a young cohort—workers aged between 25 and 29—and an older cohort aged between 40 and 44. Since the policy increased the supply of college graduates beginning in the mid-1980s, one expects the increase in the proportion of college graduates only among the young cohorts—only the cohort affected by the policy would experience the increase in the proportion of workers with a college education. Thus, if the proportion of white-collar workers is indeed a good proxy for the proportion of workers with college degrees, the proportion of white-collar workers should increase only among young cohorts starting in the mid-1980s. Figure 7 illustrates the share of college graduates and white-collar among workers separately for the younger and older cohorts of workers. The information in this figure confirms the prediction, as proportions of both college graduates and white-collar workers among the young cohorts increase rapidly after mid 1980 whereas the both proportions among the older cohorts shows little if any change during the 1980s until the early 1990s.

[12]Unfortunately, the data do not have location identifier for the establishment, thus I cannot use these data for the main analysis.

This information implies that, during this period, the increase in the proportion of white-collar workers was driven by the increased availability of college graduates.

Finally, Figure 8 provides the time trend for the proportion of workers who have a high school diploma as their highest degree. The figure shows a considerable increase in the share of workers with high school education prior to the mid-1980s. Thus, one can verify that the trend of white-collar workers does not coincide with the trend of the high school graduates. Overall, the evidence shows that the increase in workers with college degrees caused the increase in the proportion of non-production workers in this period. Thus, hereafter I use the changes in the proportion of white-collar worker as a proxy for the proportion of college graduates without further distinction.

In my main analysis, in order to compare my study with previous literature using plant level data, I focus on the manufacturing sector. Additionally, I omit years when the Mining and Manufacturing Survey was conducted as part of the Industrial Census—1983, 1988 and 1993—since variable definitions and the samples in those years are not consistent with other years.

Table 1 provides summary statistics. The first two columns contain the mean and the standard deviation during the period prior to the impact of the reform and columns (3) and (4) describe the mean and the standard deviation of the variables after the impact of the reform. All monetary values are in 1990 Korean Won. One can verify that both the value of the output and the value-added of individual plants increased rapidly during this period. The average output increased about 50 percent between the two periods from a base of 2.5 billion Won (1USD is about 1,000 Won). The average value added of each plant also increased by a large amount: about 100 percent.[13] Moreover, the average capital stock also increased rapidly during the period of interest. The average capital stock of each plant was around 723 million Won during the years 1982-1986, whereas it was about 1.4 billion Won during the years 1987-1996.

More importantly, the average proportion of white-collar workers increased by a considerable amount. In particular, the average proportion of white-collar workers within a plant increased by around 25 percent, or about five percentage points, after the reform went into ef-

[13]The value added of each plant is defined as the value of output less the cost of production which includes cost of materials and electricity.

fect. Moreover, the increase in the proportion white-collar workers did not differ significantly across differently sized plants. The proportion of white-collar workers was about 20 percent prior to the impact of the reform for all plant sizes, and about 25 percent after 1987, regardless of plant size. Consistent with the increase in the proportion of white-collar workers in individual plant, the average proportion of white-collar workers in a given region also increased by a similar amount after the mid-1980s. The average payment to workers increased during this period as the total payment to workers increased by a substantial amount despite the decrease in the number of employees. The average age of an individual plant and the building area used by each plant were both stable during the years 1982-1996. Overall, the summary statistics show that many plant level characteristics—with a few exceptions—significantly changed after the impact of the policy.

4 Research Design and Empirical Specification

In this section, I provide detailed information on how I use this annually collected data to examine human capital spillovers beyond plant boundaries.

The existence of human capital externalities implies that plants located in regions with higher levels of human capital will be more productive. Thus, one can assess the magnitude of such externalities by examining the relation between the level of the human capital and plant productivity in each region. However, empirically estimating externalities is challenging because the change in the level of human capital is endogenous in most cases. That is, unobserved factors affecting regional plant productivity can also have a positive effect on the overall level of human capital. For instance, the establishment of a "million dollar plant" can have a positive effect on the productivity of existing plants and also attract workers with higher human capital (Greenstone, Hornbeck, & Moretti, 2010). In this case, a positive relation between the level of human capital in a given region and its average plant productivity could exist even in the absence of human capital externalities. In other words, a positive correlation between the level of human capital and average plant productivity does not necessarily imply the existence of human capital externalities.

In the remainder of this section, I describe the endogeneity issue in detail using an empirical strategy adopted from Moretti (2004c). I also explain how the empirical setting in this paper helps mitigate some associated concerns.

I assume a Cobb-Douglas production function;

$$Y_{ijrt} = A_{ijrt} {B_{ijrt}}^{\alpha_b} {W_{ijrt}}^{\alpha_w} {K_{ijrt}}^{\beta} \tag{1}$$

where Y_{ijrt} is output of the plant i, in industry j, in region r, at year t. B_{ijrt}, W_{ijrt} and K_{ijrt} denote inputs: blue-collar workers, white collar-workers, and capital, respectively. Total factor productivity is represented by A_{ijrt}. If plant productivity depends on the regional level of human capital, then, $\ln A_{ijrt}$ can be expressed as follows;

$$\ln A_{ijrt} = \gamma H_{rt} + \epsilon_j + \epsilon_r + \epsilon_t + \epsilon_{rt} + \epsilon_{ijrt} \tag{2}$$

where H_{rt} is the proportion of workers who are college graduates, which measures the level of human capital in a given region. The coefficient of H_{rt}, γ, indicates the effect of regional human capital on productivity. Thus, the magnitude of human capital externalities on total factor productivity will depend on the size and significance of γ.

After taking logs of the production function (1) and substituting for $\ln A_{ijct}$, (1) can be rewritten as;

$$y_{ijrt} = \gamma H_{rt} + \alpha_b b_{ijrt} + \alpha_w w_{ijrt} + \beta k_{ijrt} + \mathbf{X}'_{ijrt}\Phi + d_j + d_r + d_t + \epsilon_{rt} + \epsilon_{ijrt} \tag{3}$$

where y_{ijrt} is log of the value added of the plant. b_{ijrt} and w_{ijrt} are the log of labor input of white-collar and blue-collar workers, respectively. k_{ijrt} denotes the log capital stock. One advantage of using repeated cross section data is that one can control for year fixed effects and region fixed effects.[14] d_j, d_r and d_t are industry fixed effects, region fixed effects, and year fixed effects, respectively. In addition to fixed effects, I also control for the basic characteristics of plant i, $\mathbf{X}_{ijrt}$, such as age of plant, the type of ownership and area of the plant's building, which can affect the productivity of the plant.

[14]The survey does contain the plant id. Unfortunately, the id is not assigned consistently across the year, thus one cannot control for the plant fixed effects.

The main source of endogeneity arises if time-varying region-specific shocks, ϵ_{rt}, are positively correlated with the change in the share of college graduates, H_{rt}. Moreover, since the change in the level of human capital is gradual in most cases, the endogeneity issue can be exacerbated by controlling for region fixed effects. In other words, since there will be only a small amount of variation left in H_{rt} after controlling for region fixed effects, even a small correlation between H_{rt} and ϵ_{rt} could bias estimates by a large amount.

This paper addresses these challenges by using the policy change that induced a sharp increase in the supply of college graduates starting in the mid-1980s (see Section 2). This raises the probability that there will be a sufficient amount of within region variation in the level of human capital to estimate γ reliably. In other words, even after controlling for region fixed effects, the variation left in H_{rt} should be sufficient to estimate γ. Furthermore, there is a sizable amount of variation in the increase in the proportion of workers who are college graduates across regions. Importantly, since this cross-regional variation arguably induced by the exogenous reform, the variation in H_{rt} during this period is unlikely to be correlated with the region-specific time shock ϵ_{rt}. Moreover, since the magnitude of impact of the reform was largely determined by the predetermined characteristics of each region—freshmen quota in 1979—controlling for region fixed effects will help address the endogeneity in the change in level of human capital.

However, Figure 5 also provides some evidence of the endogenous sorting of human capital, as there are regions which experienced distinctly smaller increases in human capital compare to the size of their initial freshmen quota. These regions include Jeonbuk and Jeonnam, which rose up against the military junta, and as a result their development was impeded during the 1980s. Another region is Gangwon, where the economy heavily depended on coal mining; the demand for coal fell sharply and the region declined rapidly during the period. Since the deviation in the increase in the level of human capital from the share of the freshmen quota is non-random, the result from longitudinal specification which controls for the region fixed effects would still be biased.

Thus, to further address this concern, I instrument the change in the proportion of college graduates by exploiting the exogenous timing and the size of the positive supply shock in college graduates induced by the reform. My approach is to extract the exogenous portion of change in the level of human capital by only using the change in the level of human capital predicted by the exogenous education reform. In particular, I instrument the change in the level of human capital using the freshmen quota prior to the reform, which determined the size of the impact, interacted with a dummy variable indicating the post-period. Formally, the first stage is as follows;

$$H_{rt} = \text{POST} * \text{PropFresh79}_{rt} + \theta_b b_{ijrt} + \theta_w w_{ijrt} + \kappa k_{ijrt} + \mathbf{X}'_{ijrt}\Pi + d_j + d_r + d_t + \epsilon_{rt} + \epsilon_{ijrt} \quad (4)$$

where H_{rt} is proportion of the college graduates and the POST is a dummy variable that takes a value of one after 1986. The relative size of the freshmen quota in 1979, PropFresh79_{rt}, is defined as follows:

$$\text{PropFresh79}_{\text{rt}} = \frac{fq1979_r}{emp_{rt}} \quad (5)$$

where $fq1979_r$ indicates the freshmen quota in 1979 in each region and the emp_{rt} is the total number of employees in the region, r at time t.

The second stage of the IV regression uses the predicted value of the proportion of college graduates from the first stage , $\hat{H}_{rt}$, and estimates the following equation.

$$y_{ijrt} = \gamma \hat{H}_{rt} + \alpha_b b_{ijrt} + \alpha_w w_{ijrt} + \beta k_{ijrt} + \mathbf{X}'_{ijrt}\Phi + d_j + d_r + d_t + \epsilon_{rt} + \epsilon_{ijrt} \quad (6)$$

Again, the coefficient of interest is γ. Since $\hat{H}_{rt}$ is a function of variables that are not correlated with ϵ_{rt}, ideally γ will not suffer from endogeneity bias and thus will address the bias associated with OLS estimates.

5 Results

In this section I provide the estimation results. I begin by documenting results from a 'pooled regression' specification which omits the region fixed effects from equation (3). Thus, the results from this specification describe the correlation between the proportion of college graduates among workers in a given region and plant productivity.

Table 2 documents the pooled regression results for various specifications. All specifications control for capital stock, labor input by the type of worker and year fixed effects. Moreover, to control for time-varying region-specific productivity shocks, I also control for the log of capital stock per worker in each region. Labor inputs are measured by the number of employees, and the capital stock is measured as the monetary value of assets excluding the value of land. Columns (2)-(4) control for additional characteristics of plants, such as age, type of ownership, and industry at the 2 digit level. Column (3) also controls for year $\times$ industry dummies, and Column (4) allows the technology to differ across industries by allowing the coefficients of labor inputs and the capital stock to vary across industries defined at the 2 digit level. The results exhibit a positive correlation between the level of human capital and plant productivity; this coincides with cross-sectional results in the U.S. as documented by Rauch (1993) and Moretti (2004c). The coefficient is consistently sizable and statistically significant across specifications. In particular, a percentage point increase in the proportion of white-collar workers in a given region—which is used as a proxy for the share of college graduates—is associated with a 0.75 percent increase in plant productivity. Although the definition of the variables is not exactly the same as given by Moretti (2004c), the magnitude of the correlation between the share of workers with a college degree and plant productivity is similar.[15] Overall, the results from Table 2 show that a positive correlation between plant productivity and level of human capital does exist in Korea. In the remainder of this section, I show that the magnitude of the coefficient decreases as I further address the endogeneity by exploring the implementation of the reform.

Table 3 reports an longitudinal specification which controls for region-specific dummies as well as year fixed effects. Again, all specifications control for the labor input of each type of workers, capital stock, area of the building used by plants, and the log of capital stock per worker in each region. The additional controls and specifications used in each column are the same as in Table 2. Since the change in the proportion of college graduates in this period is mostly caused by the reform, the region fixed effects will address a substantial portion of the endogeneity in the change in proportion of college graduates.

[15] In Moretti (2004c), the cross sectional estimate of the correlation between the share of the college graduate workers and the productivity of plants varies from 0.807 to 0.847.

As expected, the magnitude of the main coefficient decreases, compared to the baseline pooled regression result. The size of the coefficients vary from 0.3 to 0.4 which is about half the magnitude of the simple correlation between the level of human capital and productivity. However, the point estimate of the main coefficient in columns (1) and (2) is still somewhat sizable and statistically different from zero. Since this specification cannot completely rule out endogeneity, I further address the issue by using my instrumental variable.

Table 4 conveys the result for the IV regression and the corresponding reduced form result. Columns (1) and (2) report the first stage, and Columns (3) and (4) report the second stage. Additionally, Columns (2) and (4) control for the individual characteristics of plants. The first stage of both specifications is strong as the coefficient of the interaction term is statistically significant at 1%. Moreover the F-statistics of the first stage regression are sufficiently larger than 10 for both specifications. The results from the second stage provide further evidence against the existence of human capital spillovers as the magnitude of the main coefficient is smaller than that of the specification in the previous section. The magnitude of the coefficient from the preferred specification is close to zero, 0.08, and statistically indistinguishable from zero. In addition to the 2SLS result, Columns (5) and (6) report the reduced form result. Consistent with the 2SLS result, the coefficient on the $\text{POST} * \text{PropFresh79}_{rt}$ is close to zero and statistically insignificant. Instrumenting the proportion of the college graduates further addresses endogeneity and further corrects the bias in the coefficient of interest. Moreover, the decrease in the magnitude of the coefficient compared to the longitudinal estimate seems plausible, as the longitudinal estimate is most likely to be positively biased. Thus, the instrumental variable analysis of this paper provides more compelling estimates.

In short, despite finding a sizable correlation, the effect of the human capital level on productivity decreases and become statistically insignificant as I apply my instrumental variable which properly corrects the positive bias. Thus, the results provide little support for the existence of human capital spillovers beyond plant boundaries.

In remainder of this section, I perform several robustness checks.

5.1 Robustness Check

As a robustness check, I omit intermediate years when the impact of the reform was not fully realized and thus focus on the long term effects of the policy. In particular, I exclude the years between 1986 and 1993 when the proportion of college graduates in the labor force was increasing steeply as young college graduates were replacing low skilled workers.

Table 5 describes the result from pooled regression and the longitudinal estimate. As expected, the size of the coefficient becomes larger than that in the specification which included the intermediate years. However, as described in Table 6, the estimated effect of human capital on productivity decreases by a large amount and becomes statistically indistinguishable from zero after I use the IV analysis.

Lastly, I use a more general form for the production function—a translog specification—instead of the Cobb-Douglas specification and estimate (3) including the square of each log input and the interaction between each log input. Table 7 reports the results under this production function. The magnitude of the coefficient from this specification is slightly different from the one obtained under the Cobb-Douglas specification. Consistent with the analysis using Cobb-Douglas production function, the results from Table 7 shows the sizable and significant correlation between the level of human capital and plant productivity.

Table 8 reports the results of the instrumental variable analysis under translog production function. The coefficient of interest is somewhat sizable compared to the main specification, however, it is not statistically different from zero. Thus, this analysis shows that the main results are robust in regards to changes in the functional form of production function.

5.2 Internal Validity of the Reform

Although the variation of impact across regions induced by this reform was large and was exogenously imposed, one might worry about the internal validity of this policy. In particular, it is debatable whether the workers who became college graduates due to this reform are comparable in terms of human capital with college graduates who entered college prior to the education reform. In other words, if the reform simply increased the number of college graduates without adding any human capital, then one would not expect externalities from the increased proportion of college graduates.

To address this concern, I show that there were plant level responses to the increase in the share of college graduates. First, there was an increase in the proportion of white-collar workers, which I show closely mirrors the increase in the proportion of college graduates. If the workers who earned a college degree due to the increased freshmen quota had no additional human capital relative to high school graduates, there would be no reason for plants to hire them as white-collar workers. Moreover, the trend for the average capital per worker coincides with the increase in the proportion of college graduates. Figure 9 shows the average capital per worker in the manufacturing sector during the period of interest.[16] Its trend closely follows the trend of the share of college graduates. This suggests that plants increased their investment in response to the increase in the share of college graduates.

Finally, I examine the wages by cohorts to see whether the increase in college education led to an increase in the overall level of human capital. In particular, due to the compulsory school entrance law in Korea, the cohorts that were born after 1962 were more likely to enter college in 1981 or later. Thus, the cohorts born after 1962 would have a higher chance of earning a college education than the cohorts born prior to 1961. Park and Son (2012) exploit this idea by adopting a regression discontinuity design and by comparing the share of individuals with some college education and the average wage by cohorts born around 1962. The authors document a discrete jump in the share of college graduates and a corresponding jump in income and wages for the cohort born in 1962. Figures 10 and 11 describe the proportion of college graduates among workers and age-adjusted average monthly wage, respectively. In each case the x-axis denotes birth year centered at 1961. The figures verify the discrete increase in both the level of education and wages, where the latter is a proxy of human capital, for cohorts born after 1962. Overall, the evidence shows that college education after the policy had sizable value added in terms of human capital; and thus, the increased proportion of college graduates due to the policy is valid for analyzing human capital externalities.

[16]The monetary values are adjusted to the 1990 Korean Won.

6 Human Capital Externalities and Wages

In this section, I examine the human capital externalities by using the wages of workers as an outcome variable. I first present an equilibrium framework that links the human capital externalities and the wage/rent of a given region and test the predictions of the model using plant level data.

A substantial body of theoretical literature suggests possible mechanisms for human capital externalities on the wages of workers. Most relevant to this study, Moretti (2004a, 2004c) builds a model based on Roback (1982, 1988) that explains the effect of human capital externalities on wages/rent when workers and plants are perfectly mobile. Thus, the model is appropriate for documenting the effect of variation in the human capital level within a country where workers and plants can relocate across regions with little restriction.

In essence, the model claims that regions with a higher share of skilled workers will have higher rents if there are human capital externalities. Otherwise, all plants will be located in the regions with a higher level of human capital. To compensate workers for the high cost of living due to higher rents in the regions with a higher share of skilled workers, their wages will be higher than those of workers with similar characteristics in regions with relatively lower levels of human capital. Thus, in equilibrium, human capital externalities will lead rents and wages to be higher in regions with higher levels of human capital, making plants and workers indifferent between regions. To summarize, the model implies that human capital externalities on plant productivity can be examined indirectly by testing whether regional wages and/or rents depend positively on the level of human capital.[17]

I elaborate on the model which illustrates how an exogenous increase in the supply of college graduates in a given region can affect rent and wages in that region. For simplicity, I assume that plants use the same technology to produce a nationally traded product. I further assume that there are only two regions, A and B. Commuting costs between the two regions are assumed to be prohibitively high in order to rule out the possibility of workers living in

[17]Several paper such as Rauch (1993) and Acemoglu and Angrist (2000) use the wages and/or rents of each region to examine the externality of human capital on workers' productivity. Lange and Topel (2006) provides a summary on the literature on social returns to education.

region A and working in region B.[18] There are two types of goods: a nationally traded good, Y, and locally traded land, L. Thus there is a single price for the nationally traded good, whereas the price of land—rent—varies across regions. I normalize the price of the nationally traded good to be one. Plants produce, Y, using labor and capital. Moreover, I assume the technology to be Cobb-Douglas:

$$Y = AH^{\alpha}K^{\beta} \tag{7}$$

where H and K denotes labor and capital, respectively. A indicates productivity.

The existence of human capital externalities would imply that plants can produce more in regions with a higher level of human capital after controlling for inputs. Thus, if there is a human capital externality on productivity, then, A, will depend positively on, $\bar{S}$, the level human capital in a given region.[19] On the other hand, if there are no human capital externalities, then increases in $\bar{S}$ will not affect productivity. In addition, I assume that workers in a region will have externalities on the plants only within the same region. Thus the change in the level of human capital will not affect the wage and rent of other regions. Although physical constraints have limited social interactions to a lesser extent in recent years due to the rapid development in communication technology, such interactions were likely to be limited by physical proximity in Korea during the 1980s and 1990s. Also, this property—that physical proximity matters for the magnitude of externalities—is consistent with the theoretical literature illustrating possible mechanisms for human capital externalities.[20]

In equilibrium, each plant will earn a zero profit, in that its unit cost will be equal to the price of its product which is assumed to be one. Moreover, in equilibrium, the unit cost should be equal in both regions. Otherwise, plants will have an incentive to move the other location with lower unit cost. Therefore, for plants in either region, the unit cost $C(\cdot)$ should satisfy the

[18] Making this assumption ensures that the workers working in a region with high rents will be compensated by higher wages.

[19] Following (Moretti, 2004c), I empirically measure the level of human capital with the proportion of workers with a college degree in a given region.

[20] For instance, Marshall (1890) argues that the social interaction among workers is the main source of knowledge spillover. More recently, economists have argued that firms locate themselves in urban areas to acquire ideas from their neighbors. Thus the existence of local spillover has been used as an argument to explain why urban areas exists Glaeser (1999). Moreover, Jaffe, Trajtenberg, and Henderson (1993) shows that the knowledge spillover is geographically localized using patent citations

following condition:

$$C(W_A, r_A, \bar{S}_A) = C(W_B, r_B, \bar{S}_B) = 1 \tag{8}$$

where W_M denotes the wages of workers in region M, and r_M indicates the rent in region M. The workers face a two-part choice. First, they must decide where to live and, at the same time, where to work. Then, the workers supply labor and choose quantities of a composite good, Y, and a residential area to maximize their utility.[21] Since the price of the composite good is the same across regions, only the price of the residential area and the wages will affect the workers' choice regarding residency. In equilibrium, a worker should be indifferent between locations. Otherwise, they will have an incentive to move. Thus, in equilibrium, the indirect utility of the workers should satisfy the following condition:

$$V(W_A, r_A) = V(W_B, r_B) = I \tag{9}$$

where the $V(\cdot)$ denotes indirect utility. Figure 12, describes the initial equilibrium wage and rent and the change in the equilibrium wage and rent for this simple setting when there is an exogenous increase in the level of human capital in region B. For illustration, I assume that regions A and B are identical at the initial period. The left panel of Figure 12 describes the set of wages and rents that satisfies the equilibrium condition for workers and the set of workers' wages and rents that satisfies the equilibrium condition for the plants located in regions A and B in period 1. In particular, the upward sloping curve illustrates the set of wages and rents that will give utility equal to I, and the downward curve illustrates set of the wages and rents that make the unit cost of the plant in each region equal to one. Thus, the wage for workers and the rent for each region is characterized by the point where the equilibrium condition for both workers and plants are satisfied. Point 1 in each panel describes the set of wages and rents that satisfy this condition at the initial state. One may note that the equilibrium wage and rent is identical in each region—$W_A^* = W_B^*$ and $r_A^* = r_B^*$—at the initial point since the region A and region B are assumed be identical i.e. $\bar{S}_A = \bar{S}_B$.

Now, one might consider what would happen if the supply of college graduates increased

[21]Following Roback (1982), I ignore the standard labor-leisure decision of workers and assume the workers do not consume leisure.

in region B due to an exogenous policy change. For instance, if the proportion of workers who work in the region where their alma mater is located does not change radically, an exogenous increase in the college enrollment in region B would result in an exogenous increase in the average level of human capital in region B; thus, $\bar{S}_A < \bar{S}_B$.

If human capital externalities exist, an increase in the overall level of human capital will cause an increase in the plant productivity in region B. As a result, the isocost curve for a plant located in region B will shift to the right as the plant can produce more with less labor input. The right panel of Figure 1 describes the change in set of wages and rents that satisfies the equilibrium condition for the plants located in region A and B in period 2. Point 2, where the isocost of plant in region B and the indifference curve of workers intersects, indicates the new equilibrium wage and rent in region B under the existence of human capital externalities. Thus, the equilibrium wages and rent will be greater in region B—$W_A^{**} < W_B^{**}$ and $r_A^{**} < r_B^{**}$—after the increase in the human capital level in region B. In the new equilibrium, the workers in region B will receive higher wages but will also have to endure a higher rent, thus, workers will remain indifferent between the two regions. Similarly, the plants located in region B are more productive but also face higher labor costs and rent; thus the plants located in region A would not move to region B. The model shows that if there is a human capital externality on productivity, an exogenous increase in the level of human capital will have a positive effect on the wages of workers.

Using this model, I test whether the relation between the wages and the overall level of human capital coincides with my main result, and I assess the magnitude of human capital spillover on plant productivity. However, since the data do not have information on individual workers, I cannot directly compare the wages of otherwise similar individuals working in regions with different levels of human capital. Thus, I proxy the wages of workers by using the average payment to workers and examine its relation with overall levels of human capital.[22] If highly skilled workers are more likely to sort into regions with a higher share of college graduates, failure to control for the individual characteristics will yield estimates with a positive bias. Thus, the result presented in this section may provide an upper bound of the

[22]The data contain information about of the average number of the workers and total payroll for workers for each year.

overall effect of human capital in a given region on the wage of workers.

Formally, I regress the following equation:

$$p_{ijrt} = \tau H_{rt} + \alpha_w w_{ijrt} + \alpha_b b_{ijrt} + \mathbf{X}'_{ijrt}\Lambda + d_j + d_r + d_t + \epsilon_{ijrt} \tag{10}$$

where p_{ijrt} denotes the log of average wage of workers and $\mathbf{X}_{ijrt}$ indicates characteristics of the plant which can affect the average wage of workers such as the age of the plant and the type of ownership. The coefficient of H_{rt}, τ, captures the relation between the average wage of workers and level of human capital in a region, as measured by the share of workers with college degrees.

As before, I start by documenting a simple correlation between the wages of workers and the level of human capital in regions by regressing equation (10) without region fixed effects. Column (1) and (2) of Table 9 presents the simple correlations. The size of the coefficient ranges between 0.6 and 0.7 across different specifications implying that a one percent point increase in the level of human capital is associated with a 0.6 percent increase in wages. Consistent with the previous result using plant value added, the simple correlation is sizable and statistically significant. However, as described in Column (3) and (4) of Table 9, the effect of the human capital decreases and becomes statically insignificant at the 5% level after controlling for region fixed effects. Table 10 summarizes the results of instrumental variable analysis which further addresses the endogeneity of human capital in a given region. After controlling for the endogeneity, the effect of human capital on the wage of workers further decreases by a large amount and becomes statistically indistinguishable from zero. Overall, the results using wages as an outcome variable are generally consistent with the results using plant productivity as an outcome variable. Thus, the results further confirm that the increased proportion of skilled workers in other plants does not affect the productivity of workers in a given plant.

7 Conclusion and Discussion

The magnitude of the return to human capital has important policy implications, for example, on the appropriate level of government subsidies to education. Thus, there is a large body of literature documenting the private benefit from human capital accumulation on the various aspects of life—including wage, health and happiness (Card, 1999; Oreopoulos & Salvanes, 2009). However, social returns to human capital, despite their important implications on growth and public finance, have been documented to a lesser degree.

To address this partial gap in the literature, this paper the examines magnitude of human capital externalities on productivity using plant level data. In particular, it tests whether plants located in regions with higher levels of human capital can produce more with a similar amount of input. To address this issue, I explore the education reform in Korea, which exogenously changed the level of human capital across regions by increasing existing differences in the supply of college graduates starting in mid-1980s.

Using this exogenous change, I empirically estimate the relation between the regional level of human capital and plant productivity. In particular, I explicitly control for the endogeneity in the change in the level of human capital by using an instrumental variable that extracts the portion of the change in human capital induced by this reform. Overall, I find little evidence of human capital spillovers beyond plant boundaries. That is, the productivity of a given plant in a given region is not affected by the level of human capital outside that plant after controlling for the plant's own human capital and other characteristics. Consistent with this result, the wages of workers are not affected by the human capital levels of workers in other plants. Thus, this paper suggests that the radius of human capital spillovers generally do not exceed the plant level.[23]

The results of this paper coincide with the recent literature that questions the plausibility of human capital spillover beyond plant boundaries. In particular, the main channel of the externalities proposed in the literature, is the personal interaction between workers employed at different establishments. Thus, the plausibility of the results in this paper will depend on whether the workers in the manufacturing sector interact regularly with workers in other plants and, learn from them. However, personal interaction between workers beyond plant boundaries are less likely to happen in Korea due to time constraints. The workers in the manufacturing sector in Korea worked about 52 hours per week during years 1982-1996, thus it is

[23]Previous literature on human capital or knowledge spillover implicitly assumed that the maximum range of human capital spillover is bounded at the country level.

unlikely that they had enough time to interact with workers in other plants. More importantly, it is not clear why skilled workers would teach their skills to less skilled workers working in other establishments. That is, even if the workers interacted with the workers in other plants, human capital spillovers beyond plant boundaries are unlikely to occur, since skilled workers have little incentive to pass along their skills without being compensated. Given the lack of convincing literature on a mechanism for human capital outside of plant affecting productivity, the results of this paper are not surprising.[24]

[24]It is important to note that this paper does not provides evidence that contradicts all types of human capital spillovers. For instance, human capital spillovers inside the plant or peer effects where workers are more likely to interact could still exists (Mas & Moretti, 2009).

References

Acemoglu, D. (1996). A microfoundation for social increasing returns in human capital accumulation. *The Quarterly Journal of Economics*, *111*(3), 779-804.

Acemoglu, D., & Angrist, J. (2000). How large are human-capital externalities? evidence from compulsory-schooling laws. In *Nber macroeconomics annual 2000, volume 15* (p. 9-74). National Bureau of Economic Research, Inc.

Arrow, K. J. (1962). The economic implications of learning by doing. *The Review of Economic Studies*, *29*(3), 155-173.

Card, D. (1999). The causal effect of education on earnings. In O. Ashenfelter & D. Card (Eds.), *Handbook of labor economics* (Vol. 3, p. 1801-1863). Elsevier.

Choi, K.-S. (1996). The impact of shifts in supply of college graduates: Repercussion of educational reform in korea. *Economics of Education Review*, *15*(1), 1-9.

Ciccone, A., & Peri, G. (2006). Identifying human-capital externalities: Theory with applications. *Review of Economic Studies*, *73*(2), 381-412.

Currie, J., & Moretti, E. (2003). Mother's education and the intergenerational transmission of human capital: Evidence from college openings. *The Quarterly Journal of Economics*, *118*(4), 1495-1532.

Dee, T. S. (2004). Are there civic returns to education? *Journal of Public Economics*, *88*(9-10), 1697-1720.

de la Fuente, A., & Domenech, R. (2001). Schooling data, technological diffusion, and the neoclassical model. *American Economic Review*, *91*(2), 323-327.

Glaeser, E. L. (1999). Learning in cities. *Journal of Urban Economics*, *46*(2), 254-277.

Greenstone, M., Hornbeck, R., & Moretti, E. (2010). Identifying agglomeration spillovers: Evidence from winners and losers of large plant openings. *Journal of Political Economy*, *118*(3), 536-598.

Hall, R. E., & Jones, C. I. (1999). Why do some countries produce so much more output per worker than others? *The Quarterly Journal of Economics*, *114*(1), 83-116.

Huber, F. (2012). Do clusters really matter for innovation practices in information technology? questioning the significance of technological knowledge spillovers. *Journal of Economic*

Geography, *12*(1), 107-126.

Iranzo, S., & Peri, G. (2009). Schooling externalities, technology, and productivity: Theory and evidence from u.s. states. *The Review of Economics and Statistics*, *91*(2), 420-431.

Jacobs, J. (1970). *The economy of cities*. New York: Vintage.

Jaffe, A. B., Trajtenberg, M., & Henderson, R. (1993). Geographic localization of knowledge spillovers as evidenced by patent citations. *The Quarterly Journal of Economics*, *108*(3), 577-98.

Lange, F., & Topel, R. (2006). The social value of education and human capital. In E. Hanushek & F. Welch (Eds.), (Vol. 1, p. 459-509). Elsevier.

Lochner, L., & Moretti, E. (2004). The effect of education on crime: Evidence from prison inmates, arrests, and self-reports. *American Economic Review*, *94*(1), 155-189.

Lucas, R. (1988). On the mechanics of economic development. *Journal of Monetary Economics*, *22*(1), 3-42.

Lucas, R. (1993). Making a miracle. *Econometrica*, *61*(2), 251-72.

Marshall, A. (1890). *The principles of economics*. New York: Macmillan.

Mas, A., & Moretti, E. (2009). Peers at work. *American Economic Review*, *99*(1), 112-45.

Milligan, K., Moretti, E., & Oreopoulos, P. (2004). Does education improve citizenship? evidence from the united states and the united kingdom. *Journal of Public Economics*, *88*(9-10), 1667-1695.

Moretti, E. (2004a). Estimating the social return to higher education: evidence from longitudinal and repeated cross-sectional data. *Journal of Econometrics*, *121*(1-2), 175-212.

Moretti, E. (2004b). Human capital externalities in cities. In J. V. Henderson & J. F. Thisse (Eds.), *Handbook of regional and urban economics* (Vol. 4, p. 2243-2291). Elsevier.

Moretti, E. (2004c). Workers' education, spillovers, and productivity: Evidence from plant-level production functions. *American Economic Review*, *94*(3), 656-690.

Niehaus, P. (2012). Education and knowledge spillovers. *UC San Diego, Mimeo*.

Oreopoulos, P., & Salvanes, K. G. (2009). *How large are returns to schooling? hint: Money isn't everything* (NBER Working Papers No. 15339). National Bureau of Economic Research, Inc.

Park, W., & Son, H. (2012). Impact of college education on labor market and non-pecuniary outcomes. *Columbia University, Mimeo.*

Rauch, J. E. (1993). Productivity gains from geographic concentration of human capital: Evidence from the cities. *Journal of Urban Economics*, *34*(3), 380-400.

Roback, J. (1982). Wages, rents, and the quality of life. *Journal of Political Economy*, *90*(6), 1257-78.

Roback, J. (1988). Wages, rents, and amenities: Differences among workers and regions. *Economic Inquiry*, *26*(1), 23-41.

Romer, P. M. (1986). Increasing returns and long-run growth. *Journal of Political Economy*, *94*(5), 1002-37.

Rudd, J. (2000). Empirical evidence on human capital spillovers. *Federal Reserve Board, Mimeo*.

Table 1: Summary Statistics

	1982-1986		1987-1996	
	Mean (1)	Std. Dev. (2)	Mean (3)	Std. Dev. (4)
total output (*1,000,000)	2519.0	30465.4	3844.2	52455.7
value added(*1,000,000)	874.8	9139.7	1596.8	23646.4
share of white-collar in region	0.198	0.038	0.270	0.049
share of white-collar in plant	0.206	0.139	0.251	0.152
employee <25	0.216	0.131	0.250	0.143
employee >25 and <50	0.198	0.144	0.250	0.164
employee>50	0.200	0.148	0.267	0.170
number of white-collar	14.4	88.9	12.6	94.0
number of blue-collar	59.3	292.6	35.8	208.7
capital (*1,000,000)	723.1	11989.0	1423.2	25697.5
average payment(*1,000,000)	2.685	1.159	7.162	3.593
area of building (m^2)	2084.8	17170.5	2275.2	25142.2
age of plants	8.005	7.152	8.155	8.246
number of obs	121573		447807	

Note : Monetary values are in 1990 Korean Won. 1 US dollar is approximately 1,000 Won

Table 2: Pooled Regression

	(1)	(2)	(3)	(4)
share of white-collar in region	0.724***	0.749***	0.769***	0.728***
	(0.099)	(0.081)	(0.078)	(0.083)
Additional Controls		y	y	y
Industry * Year effects			y	
Technology vary by Industry				y
adj. R-sq	0.809	0.815	0.816	0.819
N	569380	569380	569380	569380

All specification includes log number of white-collar and blue-collar workers, log of capital stock and year fixed effects. Specification (2),(3) and (4) additionally controls for individual plant specific characteristics such as industry and age of plants. The standard errors are clustered at region-year level.
*** statistical significance at the 99% level
** statistical significance at 95% level
* statistical significance at 90% level

Table 3: Longitudinal Specification

	(1)	(2)	(3)	(4)
share of white-collar in region	0.456**	0.441**	0.331*	0.436**
	(0.221)	(0.210)	(0.179)	(0.208)
Additional Controls		y	y	y
Industry * Year effects			y	
Technology vary by Industry				y
adj. R-sq	0.809	0.816	0.816	0.819
N	569380	569380	569380	569380

All specification includes log number of white-collar and blue-collar workers, log of capital stock, year fixed effects and region fixed effects. Specification (2),(3) and (4) additionally controls for individual plant specific characteristics such as industry and age of plants. The standard errors are clustered at region-year level.
*** statistical significance at the 99% level
** statistical significance at 95% level
* statistical significance at 90% level

Table 4: IV analysis

	first stage		IV		reduced form	
	(1)	(2)	(3)	(4)	(5)	(6)
after1987*freshmen quota79	0.264***	0.263***			0.062	0.023
	(0.040)	(0.040)			(0.117)	(0.112)
share of white-collar in region			0.235	0.087		
			(0.426)	(0.419)		
Additional Controls		y		y		y
First stage F	43.47	43.48				
adj. R-sq	0.978	0.978	0.809	0.816	0.809	0.816
N	569380	569380	569380	569380	569380	569380

All specification includes log number of white-collar and blue-collar workers, log of capital stock, year fixed effects and region fixed effects. Specification (2), (4) and (6) additionally controls for individual plant specific characteristics such as industry and age of plants. The standard errors are clustered at region-year level.
*** statistical significance at the 99% level
** statistical significance at 95% level
* statistical significance at 90% level

Table 5: Excluding intermediate years – pooled regression and longitudinal estimate

	Pooled Regression		Longitudinal	
	(1)	(2)	(3)	(4)
share of white-collar in region	0.751***	0.748***	0.675**	0.567**
	(0.137)	(0.1132)	(0.280)	(0.254)
Year Fixed Effects	y	y	y	y
Region Fixed Effects			y	y
Additional Controls		y		y
adj. R-sq	0.816	0.822	0.817	0.823
N	280021	280021	280021	280021

All specification includes log number of white-collar and blue-collar workers, log of capital stock. Specification (2) and (4) additionally controls for individual plant specific characteristics such as industry and age of plants. The standard errors are clustered at region-year level.
*** statistical significance at the 99% level
** statistical significance at 95% level
* statistical significance at 90% level

Table 6: Excluding intermediate years – IV analysis

	first stage		IV		reduced form	
	(1)	(2)	(3)	(4)	(5)	(6)
after1987*freshmen quota79	0.289***	0.263***			0.077	0.055
	(0.044)	(0.040)			(0.146)	(0.126)
share of white-collar in region			0.268	0.192		
			(0.485)	(0.428)		
Additional Controls		y		y		y
First stage F	41.71	41.66				
adj. R-sq	0.979	0.980	0.817	0.822	0.817	0.822
N	280021	280021	280021	280021	280021	280021

All specification includes log number of white-collar and blue-collar workers, log of capital stock, year fixed effects and region fixed effects. Specification (2), (4) and (6) additionally controls for individual plant specific characteristics such as industry and age of plants. The standard errors are clustered at region-year level.
*** statistical significance at the 99% level
** statistical significance at 95% level
* statistical significance at 90% level

Table 7: Translog production function – pooled regression and longitudinal estimate

	Pooled Regression		Longitudinal	
	(1)	(2)	(3)	(4)
share of white-collar in region	0.635***	0.692***	0.367*	0.397**
	(0.101)	(0.081)	(0.209)	(0.194)
Year Fixed Effects	y	y	y	y
Region Fixed Effects			y	y
Additional Controls		y		y
adj. R-sq	0.816	0.822	0.815	0.821
N	569380	569380	569380	569380

All specification includes log number of white-collar and blue-collar workers, log of capital stock, square term of each log input and interactions between the log input. Specification (2) and (4) additionally controls for individual plant specific characteristics such as industry and age of plants. The standard errors are clustered at region-year level.

*** statistical significance at the 99% level

** statistical significance at 95% level

* statistical significance at 90% level

Table 8: Translog production function – IV analysis

	first stage		IV		reduced form	
	(1)	(2)	(3)	(4)	(5)	(6)
after1987*freshmen quota79	0.263***	0.262***			0.079	0.068
	(0.040)	(0.040)			(0.104)	(.099)
share of white-collar in region			0.302	0.259		
			(0.383)	(0.419)		
Additional Controls		y		y		y
First stage F	43.12	43.12				
adj. R-sq	0.978	0.978	0.815	0.821	0.815	0.821
N	569380	569380	569380	569380	569380	569380

In addition to year and region fixed effects, all specification includes log number of white-collar and blue-collar workers, log of capital stock, square term of each log input and interactions between the log input. Specification (2), (4) and (6) additionally controls for individual plant specific characteristics such as industry and age of plants. The standard errors are clustered at region-year level.
*** statistical significance at the 99% level
** statistical significance at 95% level
* statistical significance at 90% level

Table 9: Human Capital Spillover and Wages – pooled regression and longitudinal estimate

	Pooled regression		Longitudinal	
	(1)	(2)	(3)	(4)
share of white-collar in region	0.564***	0.646***	0.256	0.301*
	(0.059)	(0.057)	(0.157)	(0.157)
Year Fixed Effects	y	y	y	y
Region Fixed Effects			y	y
Additional Controls		y		y
adj. R-sq	0.621	0.640	0.625	0.642
N	571214	571214	571214	571214

All specification includes log number of white-collar and blue-collar workers, log of capital stock. Specification (2) and (4) additionally controls for individual plant specific characteristics such as industry and age of plants. The standard errors are clustered at region-year level.

*** statistical significance at the 99% level

** statistical significance at 95% level

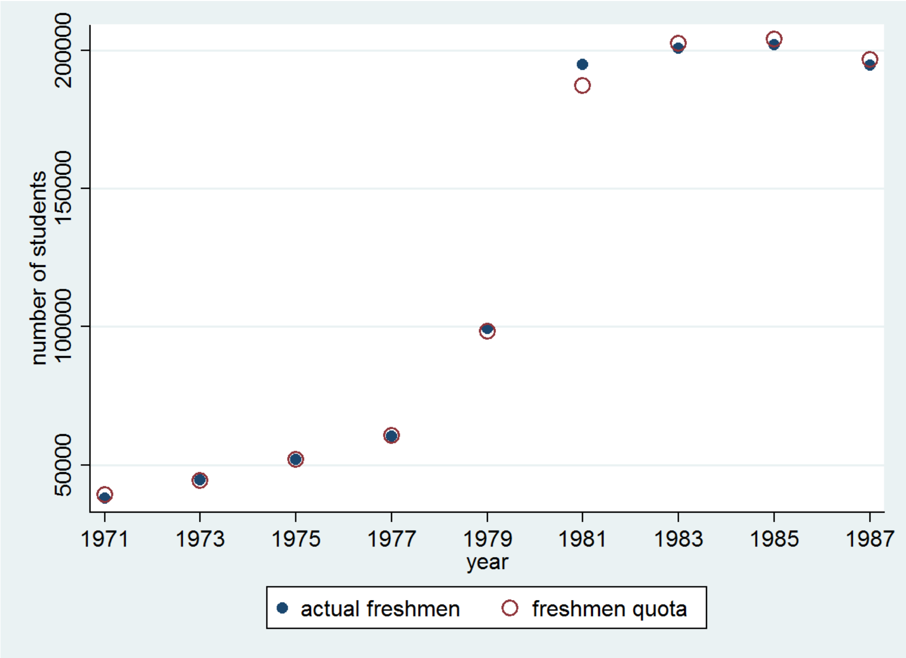
* statistical significance at 90% level

Table 10: Human Capital Spillover and Wages — IV analysis

	first stage		IV		reduced form	
	(1)	(2)	(3)	(4)	(5)	(6)
after1987*freshmen quota79	0.264***	0.263***			0.006	0.019
	(0.040)	(0.040)			(0.076)	(0.076)
share of white-collar in region			0.022	0.072		
			(0.286)	(0.282)		
Additional Controls		y		y		y
First stage F	43.26	43.26				
adj. R-sq	0.978	0.978	0.625	0.642	0.625	0.642
N	571214	571214	571214	571214	571214	571214

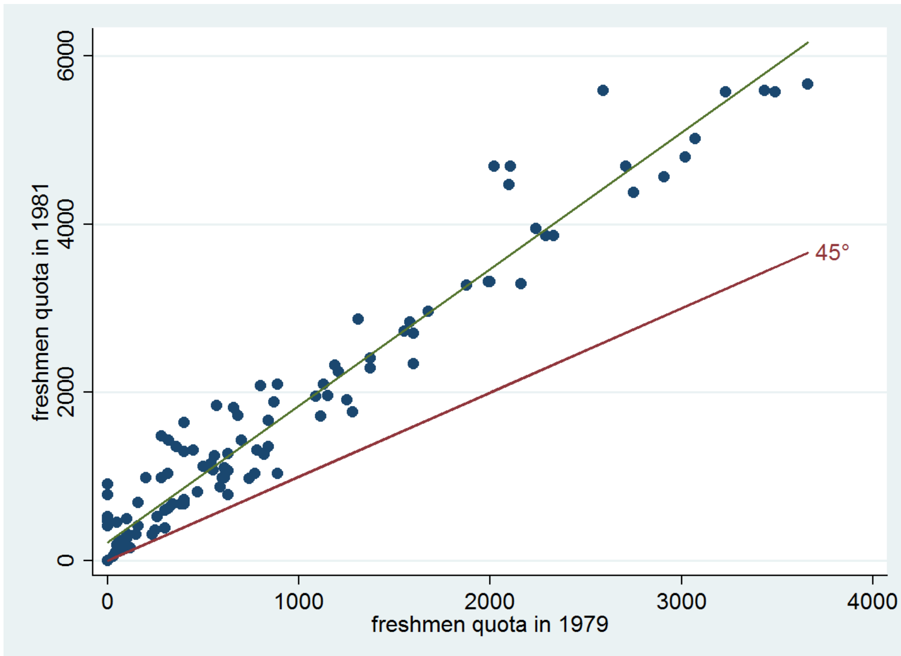
All specification includes log number of white-collar and blue-collar workers, log of capital stock, year fixed effects and region fixed effects. Specification (2), (4) and (6) additionally controls for individual plant specific characteristics such as industry and age of plants. The standard errors are clustered at region-year level.
*** statistical significance at the 99% level
** statistical significance at 95% level
* statistical significance at 90% level

Figure 1: Trend of Freshmen Quota and Actual Freshmen Enrollment



The figure plots the trend of the freshmen quota and actual freshmen enrollment for odd years during 1971 and 1987. The freshmen quota increased gradually until 1981 when it discretely increased due to July 30 Education Reform.

Figure 2: Correlation between Freshmen Quota in 1979 and 1981 by Colleges



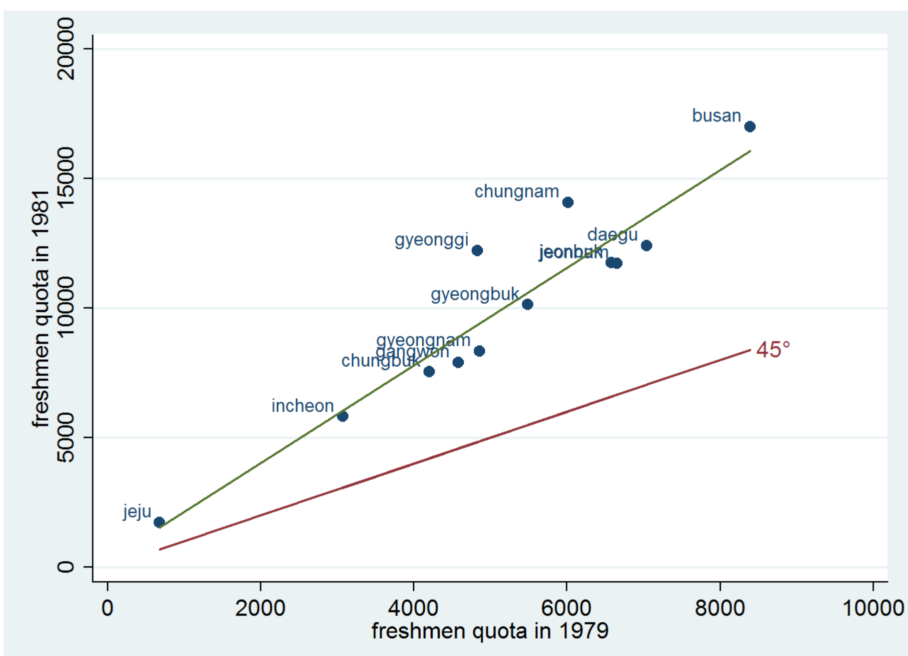


Figure 3: Correlation between Freshmen Quota in 1979 and 1981 by Region

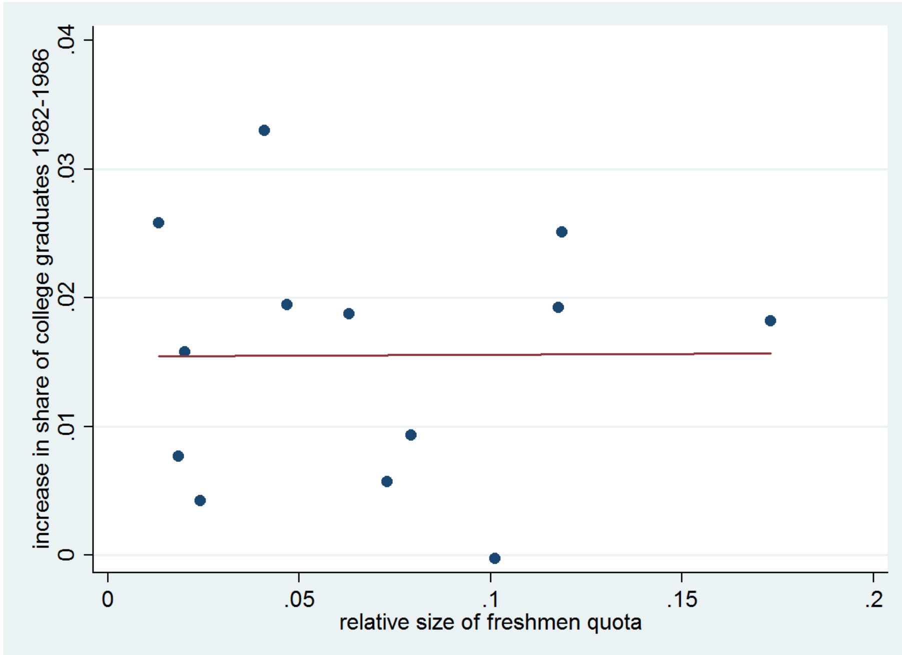


Figure 4: Correlation between Pre-trend of Change in Human Capital Level and Size of Freshmen Quota by Region

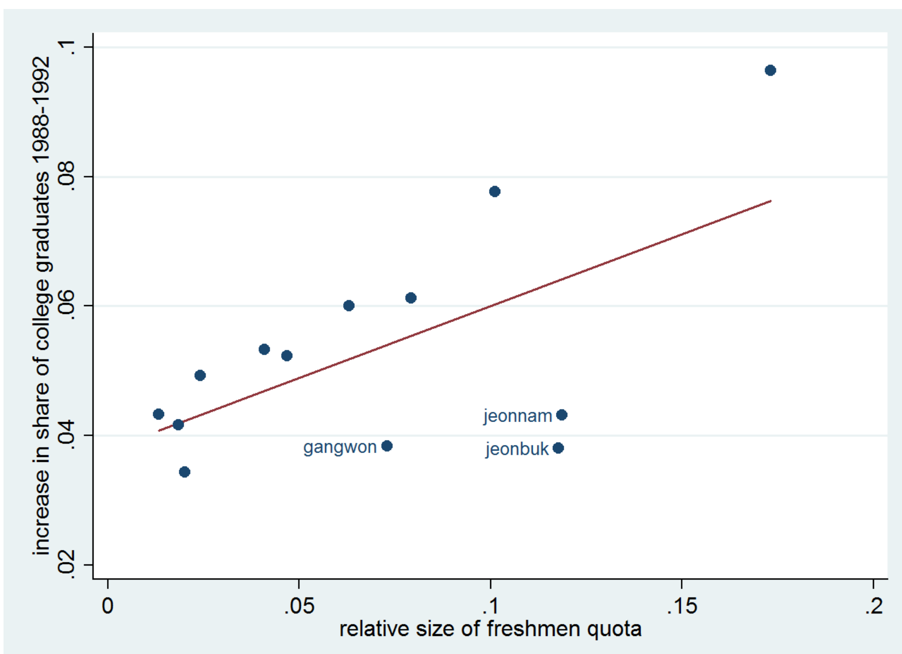


Figure 5: Correlation between Post-trend of Change in Human Capital Level and Size of Initial Freshmen Quota by Region

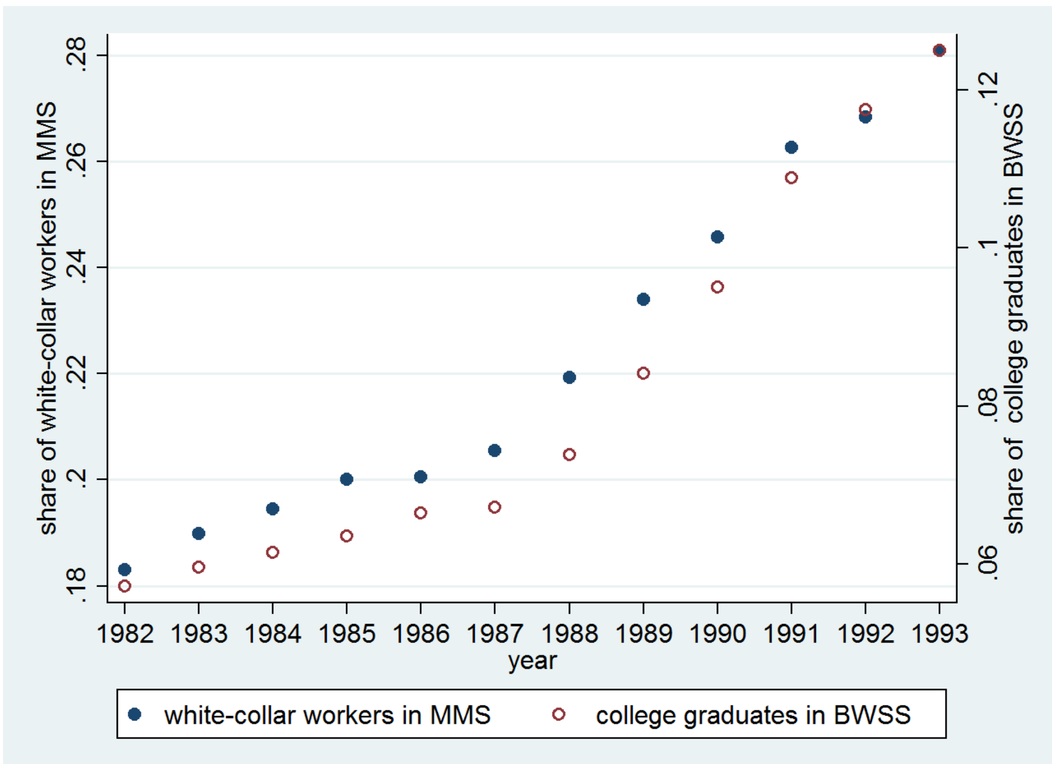


Figure 6: Proportion of College Graduate and White-Collar among Workers by Year

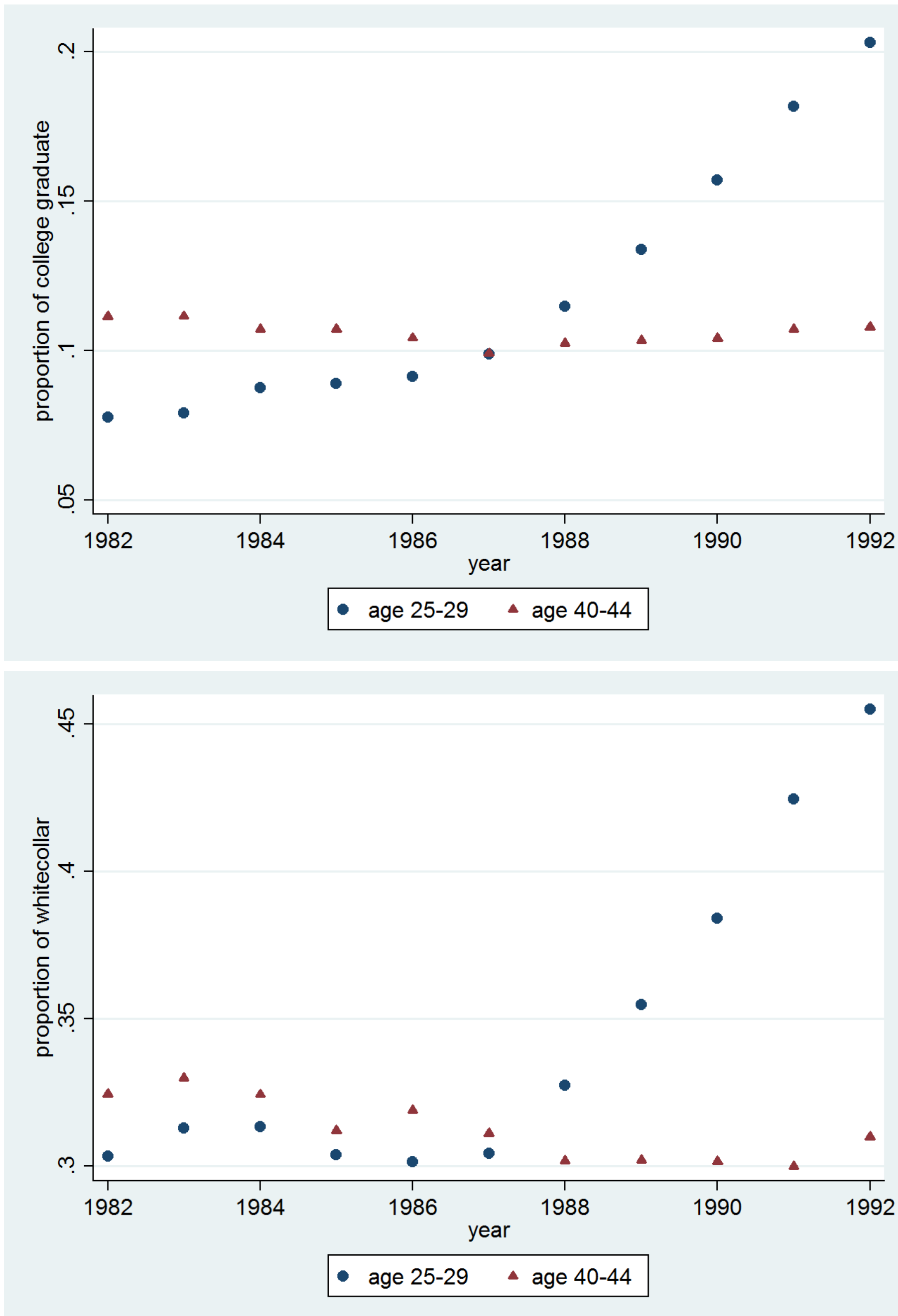


Figure 7: Trend of Proportion of College Graduates and White-Collar among Workers by Cohort

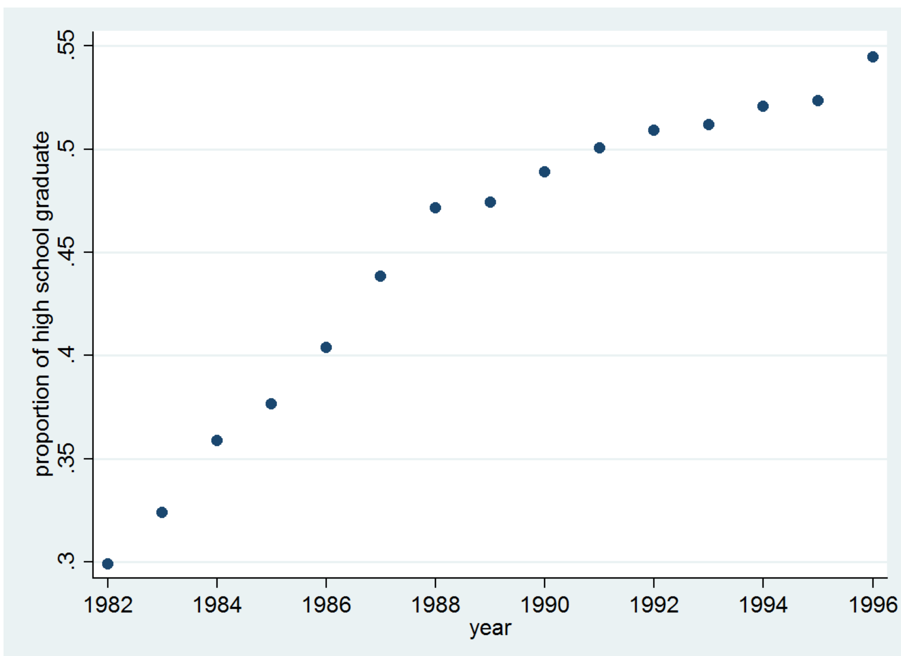


Figure 8: Proportion of High School Graduates among Workers by Year

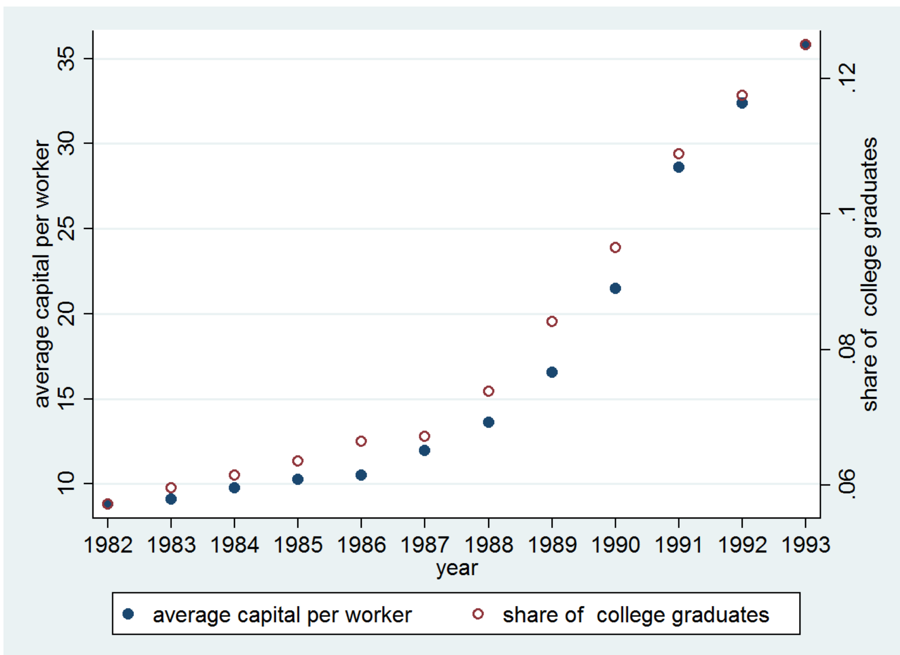


Figure 9: Trend of Average Capital Per Worker and Share of College Graduates among Workers

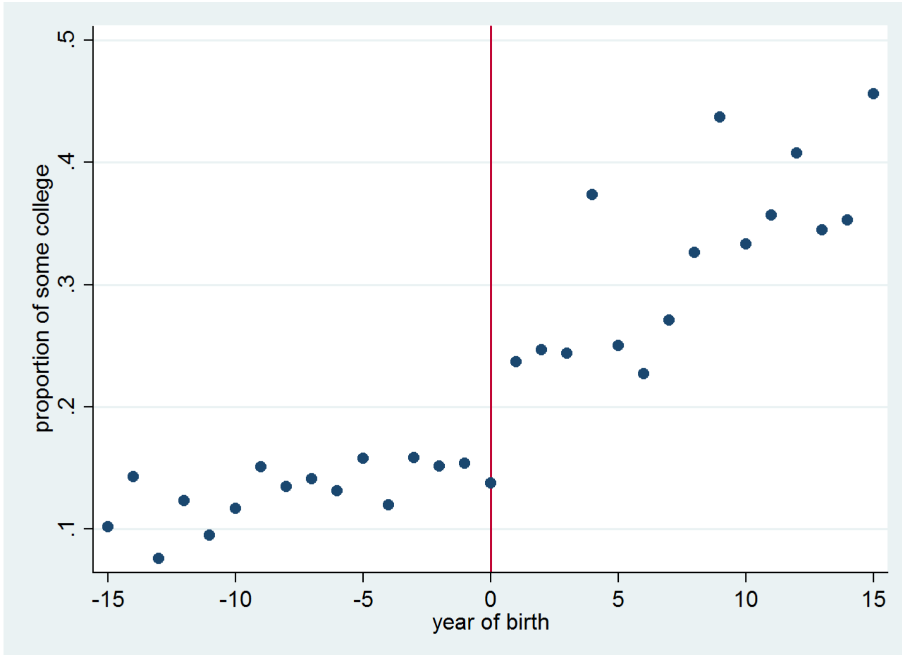


Figure 10: Proportion of individuals with Some College Education by Birth Cohort (Source: Park and Son (2012))

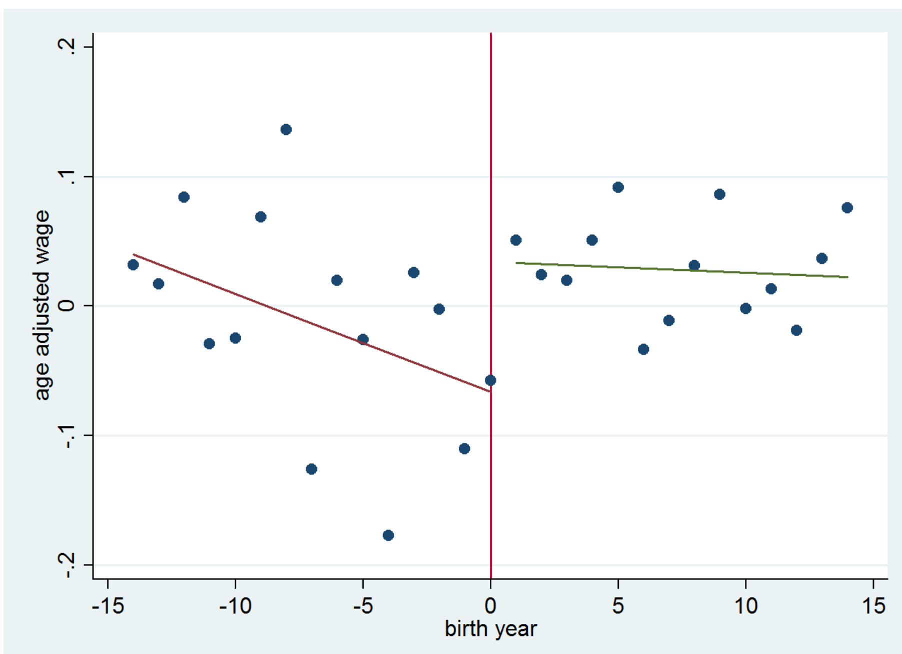


Figure 11: Age Adjusted Average Monthly Wage of Workers by Birth Cohort (Source: Park and Son (2012))

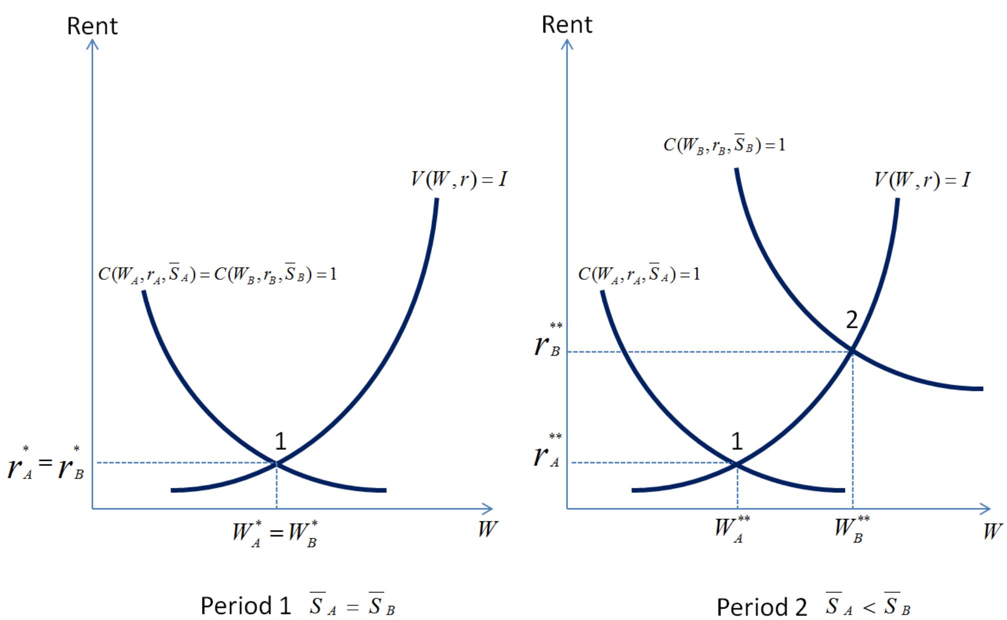


Figure 12: Equilibrium Wages and Rent in Two Region