ESSAYS ON ECONOMICS OF ENVIRONMENT, HEALTH, AND CULTURE

By

Jae Il Cho

Dissertation

Submitted to the Faculty of the

Graduate School of Vanderbilt University

in partial fulfillment of the requirements

for the degree of

DOCTOR OF PHILOSOPHY

in

Economics

May 10, 2024

Nashville, Tennessee

Approved:

Christopher Carpenter, Ph.D.

Kevin Huang, Ph.D.

Michelle Marcus, Ph.D.

Paige Skiba, Ph.D.

ACKNOWLEDGMENTS

I wish to express my sincere appreciation to my advisor, Professor Carpenter, and the committee members, Professor Huang, Professor Marcus, and Professor Skiba, for their invaluable insights and guidance throughout the semesters in graduate school. Additionally, I extend my gratitude to Professor Bethmann for constructive discussions regarding potential research ideas with me. The third chapter, “How Are They Doing? The Academic Performance and Mental Wellbeing of World Cup Babies” is published in the SSM-Population Health (Volume 25) March issue in 2024.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Executive Summary

The importance of an in-depth understanding of environmental, health, and cultural issues in economics is increasing. The first dissertation essay, "Feeling Blue and Seeing Red from Yellow Dust: The Effect of Air Pollution on Adolescent Mental Health" investigates the causal effect of air pollution on Korean adolescent mental health. Most of the existing research, however, concentrates on the effect air pollution has on adult physical health such as respiratory and (or) cardiovascular diseases. Less attention is paid to mental health outcomes such as symptoms of depression and aggression, in particular among adolescents. Identifying the causal effects of air pollution on health outcomes is a difficult task as air pollution is directly linked with economic activity which can impact aggregate health measures through changing income, for example. In this study, I follow prior work using wind speed and direction as instruments to measure the causal effects of air pollution on mental health statuses among South Korean adolescents. In Korea, west wind is known to carry "Yellow Dust" which contains harmful industrial pollutants. Regression results show strong adverse effects of PM10 and CO on adolescent mental health, making students more depressed and also more aggressive, especially in the younger cohort.

The second essay, "Gangnam Style and the Housing Market in the Eponymous District: How a Global Pop Culture Phenomenon Boosted Property Prices" investigates how the viral hit of "Gangnam Style" affected the housing market in Gangnam District. Almost overnight, the 2012 hit "Gangnam Style" made a district in South Korea's capital Seoul world famous. Using a difference-in-differences framework and data on all real estate purchases in Korea's megacity between 2009 and 2022, I examine how Gangnam's sudden popularity affected its housing market. The results show that after the release of the song prices increased significantly while the number of transactions significantly declined – a clear indication that the market was driven by supply-side factors, particularly rapidly adjusting expectations, a view that is supported by the concomitant and significant increase in the number of available hotel rooms in Gangnam.

The third essay, “How Are They Doing? The Academic Performance and Mental Wellbeing of World Cup Babies” is co-worked with Professor Dirk Bethmann at Korea University and published by *SSM-Population Health* in 2024. This paper investigates the quantity-quality trade-off of children using the 2002 Korea & Japan World Cup induced upward fertility blip as an experiment. In June 2002, South Korea cohosted the 17th FIFA World Cup. Unexpected wins carried the Korean National Football Team to the semi-finals and sparked an unprecedented euphoria among Koreans. Die-hard fans and occasional football viewers, young and old, women and men flocked the streets side by side, cheered for their team, and partied through the nights. In the subsequent spring of 2003, the country experienced a temporary and significant increase in its fertility rate. Using a difference-in-differences design, we exploit the quasi-experimental nature of this episode to investigate the Beckerian trade-off between the quantity and quality of children born to parents in South Korea. Our results support the notion of an adverse effect on child quality. Students born approximately ten months after the World Cup tend to perform significantly worse in school. Moreover, our results uncover a hitherto overlooked aspect: the same students exhibit significantly higher degrees of mental wellbeing.

CHAPTER 1

Feeling Blue and Seeing Red from Yellow Dust:

The Effect of Air Pollution on Adolescent Mental Health

1.1 Introduction

Recently, an increasing number of economists and social scientists alike have directed their attention towards the various pathways that influence mental health. It is noteworthy that mental health not only stands as a principal determinant and indicator of societal welfare (Dwyer et al., 2020; Horwitz et al., 2010; Wickham et al., 2020) but also exerts a pivotal role in shaping the broader economy, given its strong correlation with workforce productivity (de Oliveira et al., 2023; Goetzel et al., 2003; Kuroda & Yamamoto, 2018). Another significant motivation for economic research into mental health stems from concerns related to negative externalities and spill-over effects. Pollution (Cao et al., 2023) and work-related stress (Leiter & Durup, 1996; Casas & Benuto, 2022), byproducts of economic activities, adversely influence individual mental health and pose important economic issues regarding negative externalities (in the case of pollution) and spill-over effects to proximate workers and family members (in case of work-related stress). These topics underscore the necessity of mental health research in economics.

Despite the growing literature on mental health, less emphasis is made on adolescent and child (youth) mental health, notwithstanding the distinctive mental health attributes inherent to the younger demographic and their crucial role in human capital formation. Due to the developmental stage of the brain, children and adolescents exhibit unique patterns in mental health different from adults (Powers et al., 1989). More significantly, mental health challenges encountered during the nascent phases of life affect the mental health trajectories in adulthood

(Johnson et al., 2018; McLeod et al., 2016). Hence, adolescent and child mental health research has inherent importance by analyzing the distinctive vulnerabilities and possible consequences on human capital formation. Indeed, this paper contributes to the current discussion on youth mental health by showing the causal effect of pollution on adolescent mental health. I use Korean meteorological, air pollution, and mental health data as South Korea is experiencing serious air pollution with enough regional variations, giving the opportunity to analyze the causal effect of air pollution on adolescent mental health[1].

There are inherent challenges in determining causal effects. Correlations between air pollution levels and various other factors associated with health, such as changes in economic conditions and weather, can be significant. Simple regression models, therefore, often encounter significant endogeneity issues due to omitted variable biases. As a result, several studies have adopted an instrumental variable approach to ascertain a causal relationship (Chen et al., 2018; Deryugina et al., 2019; Gu et al., 2020). To overcome the endogeneity issues associated with measuring the causal effect of air pollution on adolescent mental health, I use seasonal wind patterns, the frequency and strength of west winds during the spring season, as instrumental variables. More specifically, I construct two sets of instruments. Regarding the main set of instruments ***Gobi Wind***, I count the total number of hours during the spring season when the wind blows from Bayannur in the Gobi Desert and add information about the average wind speed during that time. Regarding the alternative set ***West Wind***, I also use the average wind speed

[1] According to Environmental Performance Index Report 2016, air quality in South Korea ranked 173 out of 180 countries (EPI, 2016; Kim et al., 2020). Pollutants such as particulate matter (PM10) and carbon monoxide (CO) are transported from the Chinese industrial areas and the West coast of South Korea, where major power plants are located, to the inner Korean peninsula. Despite the South Korean government's efforts[1] to decrease air pollution levels, the country is expected to endure the most significant economic costs (and thus welfare losses) of outdoor air pollution among OECD countries by 2060 due to internationally transported air pollutants (Kim et al., 2019; OECD, 2016).

information but here I count the number of hours when the wind blows straight from the west instead.

The spring season wind patterns are chosen for instruments because concentrations of air pollutants exhibit strong seasonal patterns in South Korea and regularly peak during springtime. Underlying this pattern are the predominating west winds during the Korean spring. For centuries springtime west winds have been known to carry the so-called "Yellow Dust" – fine soil particles from the Gobi Desert of China and Mongolia that turn the sky into a characteristic yellowish color.[2] With the industrialization of modern day China, "Yellow Dust" particles started to pick up air pollutants on their way to the east. As a consequence, the inner Korean peninsula is confronted with high PM10 and CO levels, especially during the spring season (Park & Shin, 2017; Park & Hwang, 2017; Yoo et al., 2020). For the instruments to be valid, it must be the case that seasonal wind patterns are highly correlated with air pollution levels (relevance), monotonically increase the air pollution levels (monotonicity), but are uncorrelated with factors such as economic conditions, demographics, and weather that may affect adolescent mental health (exclusion restriction). In the empirical analysis, I produce evidence that these assumptions hold.

Using the Korean Children and Youth Panel Survey (KCYPS), the two-stage least squares (henceforth 2SLS) regression results indicate that the PM10 and CO levels adversely affect a number of adolescent mental health indicators. Both seventh and tenth-grade students experience more depressive and more aggressive symptoms when exposed to higher levels of air pollution. My findings also show that seventh-grade students are more seriously affected by the

[2] Yellow Dust is the literal translation of the Korean expression Hwang Sa (황사). Other terms referring to the same phenomenon include Asian Dust, Yellow Sand, Yellow Wind and China Dust Storm. Yellow Dust clouds can typically be observed in China, Korea, Japan, and the Russian Far East. The historical records of the Yellow Dust phenomenon date back as far as AD 850 (Chun, 2004) indicating that the springtime wind patterns were consistent for more than a thousand years.

PM10 and CO levels than tenth graders. More specifically, a one standard deviation increase in PM10 levels worsens depressive (aggressive) symptoms approximately by 0.22 (0.14) standard deviations for seventh-grade students (0.12 respectively 0.14 standard deviations for tenth-grade students). Similarly, a one standard deviation increase in CO levels worsens the depressive (aggressive) symptoms approximately by 0.10 (0.06) standard deviations for seventh-grade students (0.07 and 0.12 standard deviations for tenth-grade students). The results therefore strongly support the notion that air pollution (PM10 and CO levels) causes serious mental health problems among South Korean adolescents and that the more vulnerable younger cohort is disproportionately affected.

This study contributes to the ongoing debate about the effect of air pollution on health outcomes in three distinct ways. First, this study broadens and extends the current discussion on the effect of air pollution on health outcomes by considering adolescent mental health. Even though there are many studies on the effect of air pollution on adult mental health (e.g. Bishop et al., 2017; Buoli et al., 2018; Chen et al., 2018; Gu et al., 2020; Bakolis et al., 2020), the link between air pollution and adolescent mental health is less examined. Some studies document a negative correlation between air pollution exposure and adolescent mental health (Buoli et al., 2018; Joo et al., 2022; Lin et al., 2019; Szyszkowicz et al., 2020). However, most of this research lacks the quasi-experimental design needed for a causal interpretation (King et al., 2022). Fortunately, the KCYPS includes a number of variables related to depressive and aggressive symptoms, that allow a comprehensive analysis of the air pollution effect on adolescent mental health.

Second, I identify seasonal wind patterns in East Asia (wind direction and speed) as convincing instruments to measure the causal effect on the mental health among Korean adolescents. In fact, using the instrumental variable approach seems advisable because air pollution levels and the macroeconomic situation might be highly correlated. During an economic boom, firms intensify production and employment which increases air pollution (with the reverse being true during a recession). Simultaneously, the economic situation may affect people's decisions on medical expenditure and treatment as well as avoidance behavior. Simple regression models may therefore exhibit serious endogeneity issues through omitted variable biases. Consequently, several other studies also follow the instrumental variable approach to uncover the causal relationship: thermal inversions (Chen et al., 2018), wind speed and direction (Deryugina et al., 2019), and maximum wind speed (Gu et al., 2020). My choice of instruments (seasonal wind patterns proxied by wind direction and speed) leads to particularly robust implications as the frequency and the strength of West winds increase during the spring season thereby transporting diverse air pollutants from the industrial regions in China and the West coast of South Korea (relevance and monotonicity). Importantly, I provide evidence that the instruments are uncorrelated with factors that can affect adolescent depressive and aggressive symptoms such as gross regional domestic product, regional employment rate, regional population, and regional precipitation, suggesting that the exclusion restriction is not violated.

Third, in contrast to an already rich body of literature on the effect of air pollution on health outcomes in Western countries, this study focuses on the case of South Korea with its comparatively serious air pollution problem. Considering that air pollution levels may affect human health outcomes in a non-linear fashion (Arceo et al., 2016), the extrapolation of estimated health effects in Western countries with typically lower air pollution levels to the high-pollution countries of East Asia might lead to a significant underestimation of the true effects[3]. This study addresses this issue by producing robust estimates in an environment where severe

[3] The annual average PM10 level (milligrams per cubic meter) for 2016 in Seoul was 50.43 (Kwak et al., 2022), whereas Central LA exhibited an PM10 level 32.4 during the same period (South Coast AQMD, 2016). In 2016, no province within South Korea demonstrated PM10 levels lower than those observed in Central LA.

air pollution levels affect adolescent mental health.

1.2 Related Literature

Numerous causal relationships between pollution and human health have been established. For instance, groundwater pollution is known to cause an array of physical illnesses and mental health problems. Bacterial pollution of groundwater may cause a spectrum of illnesses such as hepatitis, cholera, dysentery, cryptosporidiosis, giardiasis, diarrhea, and typhoid (Cutler & Miller, 2005; Jalan & Ravallion, 2003; Roushdy et al., 2012; Wang & Yang, 2016; World Bank, 2006). Furthermore, chemically polluted groundwater may induce longer-term afflictions such as carcinogenic diseases that may precipitate cancer (Ebenstein, 2012; Lin et al., 2000; Lu et al., 2015; Morales-Suarez-Varela et al., 1995). The dissolution of heavy metals, notably arsenic, nickel, mercury, cadmium, and lead, into groundwater has also been documented to engender serious physical and mental disorders (Ayuso-Álvarez et al., 2019; Bhagure & Mirgane, 2011). Among many different types of pollution, measuring the effect of air pollution on health outcomes has recently garnered increasing attention within the field of economics. The literature primarily focuses on physical health and finds the negative consequences of air pollution on infant mortality and birth outcomes (see, for example, Chay & Greenstone, 2003; Currie & Neidell, 2005; Currie et al., 2009; Currie & Walker, 2011; Sanders & Stoecker, 2015) as well as on adult (including elderly) physical health (e.g. Chay et al., 2003; Chen et al., 2013; Deschênes et al., 2017; Deryugina et al., 2019; Hollingsworth & Rudik, 2021).

Despite a large literature focusing on the effect of air pollution on physical health, there is a limited number of studies examining the effects of air pollution on mental health. Bishop et al. (2017) examine the link between dementia and chronic air pollution exposure. Using the China Family Panel Studies and thermal inversion as instruments, Chen et al. (2018) show that a one standard deviation increase in the average PM2.5 level increases severe mental illness by 6.67 percentage points (or by 0.33 standard deviations).[4] Using wind speeds as instruments, Persico & Marcotte (2022) finds a unit increase in daily PM2.5 is associated with 0.49% increase in daily suicides. Other studies that explore the effect of air pollution on mental health include Buoli et al. (2018), Gu et al. (2020), and Bakolis et al. (2020). These studies generally show the negative link between air pollution and mental health outcomes using a holistic literature review (Buoli et al., 2018), maximum wind speeds as instruments (Gu et al., 2020) and multilevel generalized linear model regressions (Bakolis et al., 2020). While these findings mainly apply to adult mental health outcomes, little is known about the impact on adolescents. This is surprising as the developmental stage likely makes adolescents more vulnerable to exposure to air pollution (Slack & Webber, 2007; Paul et al., 2013; Lamb & Murphy, 2013) and therefore warrants separate investigation.

There are several pathways where air pollution affects adolescent mental health. One possible mechanism is through hindering the brain development of the young (through white surface matter in the left hemisphere of the brain, see for example Binter et al., 2022; Lopuszanska & Samardakiewicz, 2020; Peterson et al., 2015; Roberts et al., 2019). As a consequence, the health effects of air pollution may be more severe among adolescents compared to those of adults as adolescents exhibit different mental health patterns compared to those of adults (Slack & Webber, 2007; Paul et al., 2013; Lamb & Murphy, 2013). Moreover, the reduced white surface matter of the brain may also cause cognitive as well as behavioral problems including attention-deficit/hyperactivity disorder symptoms and conduct disorder problems (Peterson et al., 2015) such as being irresponsible, skipping school, stealing (or violating the rights of others), and physically harming other people or animals (Johns Hopkins Medicine,

[4] Chen et al. (2018) added six five-way categorical variables: depression, nervousness, restlessness, hopelessness, difficulty, and worthlessness. The resulting sum (the so-called K6 Score) ranges from zero to 24. Severe mental illness is measured by a binary variable that indicates whether the K6 Score is greater than twelve or not.

2023). Thus, several recent studies suggest a connection between exposure to air pollution and crimes (Bondy et al., 2020; Burkhardt et al., 2019; Herrnstadt et al., 2021).

There are other possible mechanisms that may explain the negative association between air pollution and adolescent mental health. Diesel-exhaust particles may trigger proinflammatory factors and reactive oxygen species, known to cause severe mental health problems (Block et al., 2004; Sui et al., 2018). Moreover, air pollution may increase the chance of systemic oxidative stress (Kelly, 2003; Risom et al., 2005), which directly raises the risk of depression (Ng et al., 2008; Yanik et al., 2004). Finally, air pollution may reduce sleep quality (Becker et al., 2017; Hayashino et al., 2010), which has a strong link with depression (O'Leary et al., 2017) as well as behavioral problems (Demichelis et al., 2023). Therefore, studying the mental health outcomes in adolescents is essential in understanding the unique vulnerabilities and potential consequences of air pollution exposure during this critical developmental stage, and this paper fills this gap by studying the effect of air pollution on adolescent mental health in South Korea.

1.3 Data and Methodology

One objective of this work will be to identify the mechanisms through which the gold and

The main interest of this paper is to measure the effect of air pollution on the mental health of young Koreans. Therefore, most variables in the regression analyses contain individual information about adolescents provided by the Korean Children and Youth Panel Survey (KCYPS) which is conducted by the National Youth Policy Institute (NYPI) and administered by the Prime Minister's Office. The NYPI chooses schools based on the size and population of the seventeen primary administrative divisions. Using proportional stratified sampling, the NYPI then randomly selects individual students. The dataset traces both the first- and the fourth-grade cohort from 2010 to 2016. In the main analyses, I use the 2016 wave of the dataset such that the first-grade cohort had become seventh-grade students and the fourth-grade cohort had become tenth-grade students.[5] Thus, all respondents were secondary school students by the time of the survey. This should benefit data quality as primary school students may have difficulties to report their mental health statuses accurately.[6]

This study apply a 2SLS regression model with *Gobi Wind* and *Wind Speed* as instruments (in one set of the robustness checks I use *West Wind* and *Wind Speed* as alternative instruments), which is summarized in equation (1.1):

$$\mathrm{Y} = \beta_0 + \beta_1 \widehat{\mathrm{Air\ Pollution}} + \beta_2 \mathrm{Covariates} + \varepsilon \tag{1.1}$$

$$\mathrm{Air\ Pollution} = \alpha_0 + \alpha_1 \mathrm{Gobi\ wind} + \alpha_2 \mathrm{Wind\ speed} + \alpha_3 \mathrm{Covariates} + \delta$$

Y denotes a (four-way) categorical outcome variable related to depressive or aggressive symptoms. Air Pollution is the regional average of 2016 spring PM10 or CO levels. Covariates is a vector of individual characteristics controlled in the regressions. The error terms ε and δ of the 2SLS regression model are assumed to have the usual properties. In two additional sets of robustness checks, I clustered the observations on the school level and on the living district level[7] and thereby (slightly) changed the properties of ε and δ. Through clustering observations, I hope to resolve possible regional correlations. Clustering at the living district level, in particular, may further add to the precision of estimates because parents but not students choose the place of residence. Besides the standard 2SLS regression model, I also apply the ordered probit model for the second stage. In contrast to the standard 2SLS regressions the probit model leads to estimates

[5] The variables of interest, depressive and aggressive symptoms, were queried in the 2012, 2015 and 2016 waves of the dataset. Note, however, that in 2012 both focus cohorts were still attending elementary school (as third and sixth graders). In 2015, this was still the case for the younger cohort (attending grade six) while the older cohort was in its third year at secondary school (attending grade nine). Only in the 2016 wave of the dataset were all respondents secondary school students.

[6] Family characteristics (i.e., household income, monthly allowance, mother's education, and gender) are reported by parents. Personal characteristics and mental health outcomes (i.e., studying time, gaming time, depressive symptoms, and aggressive symptoms) are reported by students. Lastly, NYPI filled in the administrative information (i.e., student ID, school ID, and living districts). Students had one-to-one interviews with professional interviewers and were guaranteed privacy.

[7] South Korea has a total of 228 second-tier administrative divisions. Among them are 77 cities (Si), 69 urban districts (Gu), and 82 districts (Gun) in rural areas which I use in the living district level analysis.

that are easier to interpret when dealing with categorical variables.

The individual covariates controlled in the main regressions are natural log of household income, natural log of monthly allowance, mother's education level, total studying time in minutes, total computer gaming time in minutes, and gender. Socio-economic variables like household income, students' monthly allowance, and mother's education level are known to be highly correlated with adolescent mental health statuses (McLeod & Shanahan, 1996; Strohschein, 2005; Dearing, 2008; Sareen et al., 2011). Time-use variables such as total computer gaming time and (or) study time may also affect adolescent mental health (Wenzel et al., 2009; Gunnell, 2018). According to Wenzel et al. (2009), excessive computer game playing may deteriorate mental health and even cause some mental problems. Moreover, academic pressure and study time may also have a direct negative link to adolescent mental health (Gunnell, 2018). Last but not least, it is well known that there are gender differences in mental health (Astbury, 2001; Rosenfield & Mouzon, 2012; Kiely et al., 2019). Thus, I use all of the above mentioned covariates in the regression model. The descriptive statistics of control variables are displayed in Table 1.1.

Table 1.1 : Descriptive Statistics for Control Variables

Variable	7th Grade		10th Grade		Definition
	Obs	Mean (sd)	Obs	Mean (sd)	
Income	1,918	5282.1 (2525.6)	1,787	4943.7 (2272.9)	Household income level in 10,000KRW
Allow	1,912	4.1443 (4.7864)	1,807	6.5269 (4.5714)	Monthly allowance in 10,000KRW
Study	2,001	158.12 (88.734)	1,964	115.12 (94.172)	Total studying time (tutoring time + assignment time) in minutes
Game	2,001	63.741 (70.261)	1,964	47.579 (64.133)	Total computer gaming time in minutes
Momeduc	1,876	3.0853 (1.0178)	1,734	3.1078 (1.0550)	Mother’s level of education 1 = Middle school or lower 2 = High school 3 = Community college 4 = University 5 = Graduate school or beyond
Gender	2,001	1.4808 (0.4998)	1,964	1.4756 (0.4995)	Gender 1 = Male 2 = Female
Spring PM10	2,001	59.382 (6.3318)	1,964	59.901 (6.0812)	Spring 2016 PM10 level
Spring CO	2,001	0.4711 (0.0613)	1,964	0.4706 (0.0606)	Spring 2016 CO level

Note: Standard deviations in parentheses.

Compared to seventh-grade students, tenth graders generally receive more allowance from their parents, conduct generally more self-study (approximately eight minutes more assignment time), and receive less private tutoring (approximately fifty minutes) which may reflect high schools' mandatory self-study sessions after the regular class schedule. Moreover, tenth-grade students spend 15 minutes less playing computer games than seventh-grade students.

Outcome variables measure different depressive and aggressive symptoms. KCYPS provides detailed information on adolescent mental health. I use eleven variables measuring depressive symptoms and five variables related to aggressive symptoms (i.e., I consider a total of sixteen outcome variables). Using sixteen different outcome variables, allows me to paint a comprehensive picture of how adolescent mental health statuses are affected by the PM10 and CO levels in South Korea. The summary statistics of the sixteen different outcome variables are displayed in Tables 1.2 and 1.3.

Table 1.2: Descriptive Statistics for Outcome Variables (Depressive Symptoms)

Variable	**7th Grade**		**10th Grade**		**Definition**
	Obs	Mean (sd)	Obs	Mean (sd)	
Unproductive	2,001	3.2829 (0.7556)	1,964	3.2363 (0.7252)	I am not productive and don't have energy 1 = strong yes 2 = yes 3 = no 4 = strong no
Depressed	2,001	3.4088 (0.7258)	1,964	3.3269 (0.7138)	I am depressed and sad (same categorical definition with Dep1)
Anxious	2,001	3.0095 (0.9069)	1,964	2.7994 (0.9193)	I am anxious (same categorical definition with Dep1)
Suicide	2,001	3.6417 (0.5942)	1,964	3.5784 (0.6127)	I talk about committing suicide (same categorical definition with Dep1)
Cry	2,001	3.1654 (0.8533)	1,964	3.2006 (0.8178)	I often cry (same categorical definition with Dep1)
Remorse	2,001	3.1564 (0.8167)	1,964	3.0601 (0.8122)	Wrong things caused by me (same categorical definition with Dep1)
Lonely	2,001	3.3388 (0.7957)	1,964	3.1589 (0.8220)	I am lonely (same categorical definition with Dep1)
Unmotivated	2,001	3.4548 (0.6790)	1,964	3.3768 (0.6662)	I am not motivated (same categorical definition with Dep1)
Pessimistic	2,001	3.3848 (0.7408)	1,964	3.1823 (0.8134)	I am not optimistic (same categorical definition with Dep1)
Tough	2,001	3.3963 (0.7123)	1,964	3.2454 (0.7586)	Everything is tough (same categorical definition with Dep1)
Insomnia	2,001	3.2569 (0.8331)	1,964	3.2938 (0.7882)	Cannot easily fall asleep or wake up (same categorical definition with Dep1)
DEPINEX	2,001	3.4208 (0.6905)	1,964	3.3198 (0.6783)	Index variable for depressive symptoms (same categorical definition with Dep1)

Note: Standard deviations in parentheses. Please refer to Appendix Table 1.11 to review the survey questions.

Note that except for Cry and Lonely all outcome variables listed in Table 1.2 contain information required by the widely-used Hamilton Depression Rating Scale. In fact, the list includes all information that is not physiological, somatic, or provided by a trained interviewer. Compared to seventh-grade students, tenth-grade students in general are marginally worse off in depressive measures. Tenth-graders felt more anxious, show more suicidal impulses, assess themselves less productive, and are less optimistic. As can be seen in the last row of Table 1.2, I also generated an index variable (DEPINDEX) as follows. For each student, I added the numerical values of the reported categories of all eleven depressive symptoms (Unproductive to Insomnia) and categorized the resulting sum.[8] Through generating the index variable, I hope to obtain a very precise measure as any inaccuracies that may occur in the individual measures of depressive symptoms should (partially) balance out during aggregation.

Table 1.3: Descriptive Statistics for Outcome Variables (Aggressive Symptoms)

Variable	**7th**		**10th**		**Definition**
	Obs	Mean (sd)	Obs	Mean (sd)	
Irritable	2,001	2.9990 (0.7931)	1,964	3.0056 (0.7584)	I become irritated even on small things 1 = strong yes 2 = yes 3 = no 4 = strong no
Disturb	2,001	3.0590 (0.7685)	1,964	3.0922 (0.7294)	I disturb or annoy friends (same categorical definition with Agr1)
Aggress	2,001	3.1579 (0.7563)	1,964	3.1054 (0.7497)	I become aggressive if I cannot do things in my way (same categorical definition with Agr1)
Fight	2,001	3.1534 (0.7720)	1,964	3.2149 (0.7135)	I often fight on trivial matters (same categorical definition with Agr1)
Angry	2,001	3.2249 (0.8102)	1,964	3.2439 (0.7554)	I am angry all the time (same categorical definition with Agr1)
AGRINDEX	2,001	3.2334 (0.7273)	1,964	3.2408 (0.6975)	Index variable for aggressive symptoms (same categorical definition with Agr1)

Note: Standard deviation in the parenthesis. Please refer to Appendix Table 1.11 to review the survey questions.

In contrast to the depressive symptoms, tenth-grade students generally are better off as they exhibit marginally less aggressive symptoms than seventh-grade students. Tenth-grade students are less likely to annoy their friends, fight on trivial matters, or be angry all the time. As with the depressive symptoms, I also aggregated aggressive symptoms measures (Irritable to Angry) to obtain a four-way index variable (AGRINDEX).[9]

I use air pollution data (PM10 and CO levels) from the Air Korea website. Air Korea

[8] After adding eleven depressive symptoms (Unproductive to Insomnia), DEPINDEX is set to 1 if the sum obtained is within the range [11,20]; DEPINDEX = 2 if the sum is within [21,28]; DEPINDEX = 3 if the sum is within [29,36]; DEPINDEX = 4 if the sum is within [37,44].

[9] After adding five aggressive symptoms (Irritable to Angry), AGRINDEX is set to 1 if the sum obtained is within [5,8]; AGRINDEX = 2 if the sum is within [9,12]; AGRINDEX = 3 if the sum is within [13,16]; AGRINDEX= 4 if the sum is within [17,20].

collects PM10 and CO levels at local air quality measurement stations on an hourly basis.[10] For the analysis, I calculate the regional averages of PM10 and CO levels in each of the seventeen first-tier administrative divisions of South Korea (eight metropolitan cities and nine provinces) during the spring of 2016.[11] Table 1.4 shows the descriptive statistics of PM10 and CO levels by region.

Table 1.4: Average 2016 Spring PM10 and CO by Region

West Regions	**PM10**	**CO**	**East Regions**	**PM10**	**CO**
Seoul	66.242	0.5658	Daegu	55.447	0.3887
Gyeonggi	68.515	0.5184	Ulsan	55.464	0.5077
Incheon	63.834	0.5358	North Gyeongsang	50.251	0.5077
Daejeon	57.426	0.4169	South Gyeongsang	54.440	0.4572
North Chungcheong	62.292	0.4534	Gangwon	56.949	0.4325
South Chungcheong	59.428	0.4228	Busan	55.428	0.4313
Sejong	55.443	0.6107			
North Jeolla	65.528	0.4473			
South Jeolla	48.297	0.4684	**(Province Level) Island**	**PM10**	**CO**
Gwangju	56.743	0.4580	Jeju	53.442	0.2960

Note: Units are milligrams (mg) per cubic meter for PM10 and parts per million (ppm) for CO. I first calculated the mean of each air quality measurement station from March 1st of 2016 to May 31st of 2016. Then, I matched air quality measurement stations to the seventeen first-tier administrative divisions of South Korea to calculate the regional average of Spring 2016 Spring PM10 and CO levels. Please refer to Appendix Figure 1.3 for a map of South Korea and regional pollution levels.

As can be seen regions in the West of South Korea (i.e., Seoul, Gyeonggi, Incheon, and North Jeolla) tend to have higher PM10 and CO levels than regions in the East (i.e., North Gyeongsang, South Gyeongsang, Gangwon, and Busan). This pattern is of course to be expected considering the origins of the pollutants in the heavily industrialized coastal areas of China and in the West coast of Korea.

Data related to the direction and speed of wind is collected from the Korea Meteorolo-

[10] Researchers can request historical air pollution data from the Air Korea archives dating back to 2001 (starting with 16 air quality measurement stations). In the spring of 2016, a total of 320 stations in all seventeen regions of South Korea were in use. In the analysis, I use hourly averages of PM10 and CO levels reported by the stations in a region. Weblink: https://www.airkorea.or.kr/web/ [last accessed on August 23, 2022].

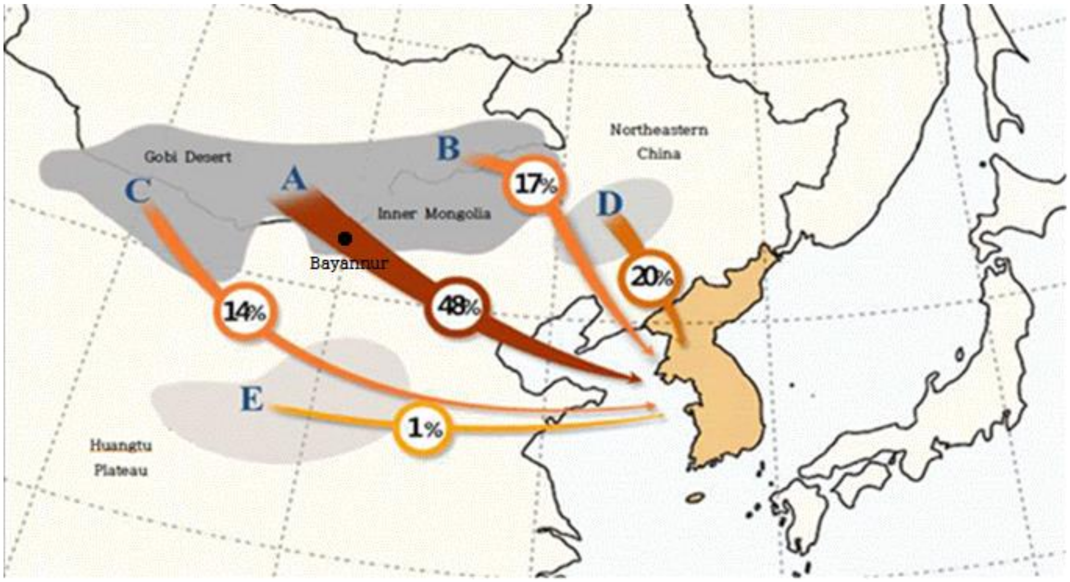
[11] Henceforth, spring 2016 addresses the period from March 1st to May 31st, 2016.

gical Administration (KMA).[12] The KMA provides hourly information on wind direction and speed at 95 meteorological stations. Information related to the direction of wind follows the standard wind rose: wind blowing exactly from the North obtains 0 degrees; East: 90 degrees; South: 180 degrees; and West: 270 degrees. Using this information, I construct the instrument *Gobi Wind*[13] by counting for every meteorological station the total number of hours during the spring season of 2016 when the wind blew from the city of ***Bayannur*** (see Figure 1.1) in the heart of the Gobi Desert plus-minus 90 degrees and then calculated regional averages based on the stations' locations. I chose ***Bayannur*** as the city is located in the middle of the southern border of the Gobi Desert, where the wind picks up the pollutants the most on the way to the Korean Peninsula.

[12] Researchers can request historical meteorological information from the archives dating back to 1904 (by selecting a year, month, and day). There is a total of 95 meteorological stations back in the spring of 2016. I primarily use hourly wind direction and speed reported by the 95 meteorological stations. Historical data can be requested at https://data.kma.go.kr/ [last accessed on August 23, 2022].

[13] Because I try to capture the pollutants carried by winds from the Gobi Desert, regions are assigned different wind directions for *Gobi Wind*. Seoul, Gyeonggi, South Chungcheong, North Chungcheong, Incheon, Sejong, and Daejeon are exposed to *Gobi Wind* when wind directions range between 195 and 15 degrees. Similarly, South Gyeongsang, North Gyeongsang, Busan, Gwangju, Ulsan, Daegu, Gangwon, South Jeolla, and North Jeolla are exposed to *Gobi Wind* when wind directions range between 200 and 20 degrees. For Jeju the range is 205 to 25 degrees.

Figure 1.1: Spring Wind Direction and Yellow Dust from China



Notes: The map shows the provenance of Yellow Dust particles dispersed in South Korea. It is based on a total of 143 Yellow Dust events between 2002 and 2017 tracked by the Korea Meteorological Administration. Source: Ministry of Public Administration and Security (2019).

From the perspective of the seventeen South Korean administrative divisions, ***Bayannur*** is in a slightly different direction. Thus, exposure to *Gobi Wind* slightly differs across regions even if the whole of Korea experiences winds from the exact same direction. To check for the robustness of the findings, I also construct the alternative instrument *West Wind* (wind blowing from 180 degrees to 355 degrees) which affects regions in the same way if winds blow from the exact same direction.[14] Finally, I define the instrument *Wind Speed* as an average of the measured wind speeds in a region during the spring of 2016. Table 1.5 shows the regional averages of *Gobi Wind*, *West Wind*, and *Wind Speed*.

[14] I excluded exact North (0 degrees) from the *West Wind* instrument as winds from Siberia are typically much less polluted than winds from China.

Table 1.5: Average Spring Wind Count and Speed

Location	Gobi Wind (hours)	West Wind (hours)	Wind Speed (m/sec)
Seoul	1546.0	1522.0	2.4216
Gyeonggi	1519.2	1369.8	1.9239
Incheon	1522.5	1606.5	3.8176
Daejeon	1548.0	1454.0	1.6308
North Chungcheong	1549.3	1259.0	1.9218
South Chungcheong	1451.1	1192.3	1.7803
Sejong	1548.0	1454.0	1.6308
North Jeolla	1505.1	1213.4	2.3098
South Jeolla	1459.4	1185.4	2.7448
Gwangju	1618.0	1322.0	1.7591
Daegu	1153.0	1025.0	2.3497
Ulsan	1440.0	1301.0	2.2009
North Gyeongsang	1500.7	1327.1	2.2773
South Gyeongsang	1487.0	1080.6	1.6579
Gangwon	1390.9	1219.1	2.0119
Busan	1281.0	1181.0	3.2785
Jeju	1320.3	1383.7	2.4917

Note: Because there was no meteorological station at Sejong in 2016, I instead use the regional average of nearby Daejeon. The distance between Sejong city and Daejeon city is about 16.6km. The unit for *Wind Speed* is meters per second. For *Gobi Wind* (*West Wind*), I counted the total number of hours with winds from *Bayannur* (the West) in each first-tier administrative division of South Korea from March 1st of 2016 to May 31st of 2016. The hours total is then divided by the number of meteorological stations in each division. Please refer to Appendix Figure 1.3 for a map of South Korea and regional wind patterns.

1.4 Results

Before I present the main results, I examine the chosen instruments in greater detail. In a preliminary step, I regress Air Pollution on *Wind Speed* and *Gobi Wind* as well as on *Wind Speed* and *West Wind* during the spring of 2016 (cf. Tables 1.6 and 1.7). To be precise I use weekly averages of the air pollution and wind pattern measures in each of the seventeen regions such that I consider a total of 221 observations.[15] The results of these regressions are presented in the first and fourth columns. I also run extended specifications of the regression. In columns two and

[15] For all three measures, the 92 days from March 1st to May 31st result in 13 weekly observations, with the last observations reflecting eight-day averages.

five I add controls for location fixed effects (FE). Finally, columns three and six add controls for a linear-quadratic (LQ) weekly trend.

Table 1.6: Air Pollution and Wind Patterns I

	PM10			CO		
	(1)	(2)	(3)	(4)	(5)	(6)
Gobi Wind	0.1620***	0.1546***	0.2303***	0.0012***	0.0006***	0.0004**
	(0.0466)	(0.0477)	(0.0443)	(0.0003)	(0.0002)	(0.0002)
	[0.001]	[0.001]	[0.000]	[0.000]	[0.010]	[0.028]
Wind Speed	-2.2951*	-8.9396***	-13.705***	-0.0277***	-0.0873***	-0.0859***
	(1.3503)	(2.1354)	(2.0298)	(0.0090)	(0.0096)	(0.0085)
	[0.091]	[0.000]	[0.000]	[0.002]	[0.000]	[0.000]
Constant	44.870***	58.455***	43.243***	0.3916***	0.5474***	0.6233***
	(6.4505)	(7.7625)	(7.5709)	(0.0429)	(0.0349)	(0.0318)
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Location FE	no	yes	yes	no	yes	yes
LQ Time Trend	no	no	yes	no	no	yes
Observations	221	221	221	221	221	221
F-Statistics	8.43	14.76	35.12	14.67	46.00	52.53

Note: Standard errors in parentheses, p-values in brackets. *,**,*** denote statistical significance at the 10%, 5% and 1% level. F-tests check the joint null hypothesis that Gobi Wind = Wind Speed = 0.

Table 1.6 shows that *Wind Speed* and *Gobi Wind* are highly correlated with PM10 and CO levels during the spring of 2016. One additional hour of weekly *Gobi Wind* increases the PM10 level approximately by 0.16mg per cubic meter and the CO level raises by approximately 0.0012ppm. Moreover, with a slower *Wind Speed*, transported air pollutants from China (and the Korean west coast) stayed (or were trapped) in inner South Korea. There are no substantial changes when replacing *Gobi Wind* by *West Wind* (see Table 1.7).

Table 1.7: Air Pollution and Wind Patterns II

	PM10			CO		
	(1)	(2)	(3)	(4)	(5)	(6)
West Wind	0.1434***	0.1394***	0.2103***	0.0009***	0.0002	0.0003
	(0.0445)	(0.0489)	(0.0454)	(0.0003)	(0.0002)	(0.0002)
	[0.001]	[0.005]	[0.000]	[0.005]	[0.494]	[0.192]
Wind Speed	-3.7304***	-10.424***	-15.733***	-0.0373***	-0.0899***	-0.0881***
	(1.3656)	(2.1807)	(2.1180)	(0.0092)	(0.0099)	(0.0089)
	[0.007]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Constant	51.938***	64.888***	53.992***	0.4665***	0.5973***	0.6526***
	(5.0046)	(6.8411)	(6.5657)	(0.0337)	(0.0311)	(0.0275)
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Location FE	no	yes	yes	no	yes	yes
LQ Time Trend	no	no	yes	no	no	yes
Observations	221	221	221	221	221	221
F-Statistics	7.57	13.46	31.82	10.55	41.62	50.18

Note: Standard errors in parentheses, p-values in brackets. *,**,*** denote statistical significance at the 10%, 5% and 1% level. F-tests check the joint null hypothesis that West Wind = Wind Speed = 0.

Next, I examine whether air pollution (PM10 and CO) levels are endogenous to adolescent mental health outcomes using the Durbin and Wu-Hausman tests (see Appendix Table 1.12). The obtained Durbin Scores and the Wu-Hausman test statistics generally indicate a strong endogeneity problem and hence favor the instrumental variable specification. Moreover, the endogeneity issue is more pronounced in regression analyses using observations from seventh-grade students than in those related to tenth graders.[16] To further check the validity of the instruments, I inspect the F-statistics of the various first stage regressions. First stage F-statistics bigger than ten generally indicate that the instruments are robust for 2SLS regressions (Bound et al., 1995; Staiger & Stock, 1997; Wooldridge, 2011). The following Table 1.8 shows the first stage F-statistics for a total of 24 different regression specifications.[17]

[16] Corresponding OLS regression results for Durbin and Wu-Husman tests using index variables (DEPINDEX and AGRINDEX) are displayed in Appendix Table 1.13.

[17] Note that we consider three different clustering levels (individual, school, and living district), two alternative pairs of instruments (first *Gobi Wind* and *Wind Speed* then *West Wind* and *Wind Speed*), two air pollutants (PM10 and CO) as well as two cohorts of young adolescents (seventh graders and tenth graders).

Table 1.8: First Stage F-Statistics

Instruments	**Gobi Wind and Wind Speed**					
Cluster Level	**Individual**		**School**		**Living District**	
Spring Pollution	PM10	CO	PM10	CO	PM10	CO
7th graders	139.64	696.05	51.795	157.26	16.377	43.420
10th graders	140.93	645.23	91.515	285.50	16.609	51.284
Instruments	**West Wind and Wind Speed**					
Cluster Level	**Individual**		**School**		**Living District**	
Spring Pollution	PM10	CO	PM10	CO	PM10	CO
7th graders	291.24	347.87	92.162	78.789	27.690	21.129
10th graders	256.87	391.37	171.25	143.83	29.387	23.293

First stage F-statistics (individual level regression) using *Gobi Wind* and *Wind Speed* as instruments are bigger than 130. Moreover, no matter the level of clustering, estimated F-statistics remain bigger than 16. The estimated F-statistics when using *West Wind* and *Wind Speed* as instruments are also high and well above the conventional threshold. Consequently, I have established *Gobi Wind* and *Wind Speed* as well as *West Wind* and *Wind Speed* as valid instruments in the 2SLS regression model. The full first-stage regressions can be found in Appendix Tables 1.14 and 1.15.

Although the instruments are highly correlated with air pollution levels as can be seen in the first stage F-statistics, they are uncorrelated with factors that can be associated with adolescent mental health such as gross regional domestic product, employment rate, population, and precipitation (see Appendix Table 1.16). I regressed ***Gobi Wind*** and ***Wind Speed*** on regional GDP, employment rate, population, and precipitation, but none of them shows statistical significance. This strengthens the credibility of the instruments that support the causal interpretation of the observed effects of air pollution on the mental health statuses of adolescents.

I now turn to the main findings. Table 1.9 displays the estimated 2SLS air pollution effects (PM10 and CO) on depressive symptoms index variable (DEPINDEX) of seventh-grade and tenth-grade students using *Gobi Wind* and *Wind Speed* as instruments.

Table 1.9: Regression Results Using Depressive Symptoms Index Variable
(Instruments: *Gobi Wind* and *Wind Speed*)

Model	**7th Grade**		**10th Grade**	
	PM10	**CO**	**PM10**	**CO**
2SLS	-0.0240	-1.1586	-0.0137	-0.8197
	(0.0071)***	(0.3982)***	(0.0075)*	(0.4224)*
P>|t|	0.001	0.004	0.068	0.052
	[0.0073]***	[0.4005]***	[0.0082]*	[0.4639]*
P>|t|	0.001	0.004	0.095	0.077
	{0.0069}***	{0.4256}***	{0.0079}*	{0.4399}*
P>|t|	0.001	0.006	0.084	0.062
Oprobit 2SLS	-0.0112**	-2.1575***	-0.0248*	-1.4914**
	(0.0044)	(0.6975)	(0.0129)	(0.7348)
P>|t|	0.011	0.002	0.054	0.042
Stand. Coef (2SLS)	-0.2205	-0.1032	-0.1208	-0.0718
Stand. Coef (Oprobit)	-0.2797	-0.1315	-0.1477	-0.0882
Observations	1,776	1,776	1,655	1,655
[School]	550	550	787	787
{Living District}	128	128	125	125

Note: Standard errors in (parentheses). Standard errors using school level clustering in [brackets]. Standard errors using living district level clustering in {braces}. *,**,*** denote statistical significance at the 10%, 5% and 1% level.

An increase in PM10 levels adversely affects both seventh-grade and tenth-grade students' depressive symptoms. Moreover, the estimated effects are more pronounced among seventh-grade students. Similarly, an increase in CO levels also adversely affects adolescent depressive symptoms. The magnitude of coefficients for PM10 and CO cannot be directly compared because of different units. To better compare the magnitudes of the estimated coefficients in the PM10 and CO regressions, I check the standardized coefficients. A one standard deviation increase in PM10 levels worsens the depressive symptoms approximately by 0.22 standard deviations for seventh-grade students and 0.12 standard deviations for tenth-grade students. Moreover, a one standard deviation increase in CO levels worsens the depressive symptoms approximately by 0.10 standard deviations for seventh-grade students and 0.07 standard deviations for tenth-grade students. The standardized coefficients show that PM10 have more adverse effects than CO. I also check the 2SLS ordered probit regression results as the outcome variable is defined categorically. The ordered probit regression results generally point in the same direction with the main regression results. Finally, to check which of the eleven depressive symptoms are affected, I also run separate regressions on each of the eleven depressive symptoms. 2SLS regression results are shown in Appendix Table 1.17; 2SLS regression results with different levels of clusters, in Appendix Table 1.18; 2SLS ordered probit regression results, in Appendix Table 1.19; standardized coefficients in Appendix Table 1.20. As Appendix Tables 1.17, 1.18, 1.19, and 1.20 show, nearly all aspects of depressive symptoms (Unproductive: “I am not productive and don’t have energy”, Depressed: “I am depressed and sad”, Anxious: “I am anxious”, Suicide: “I talk about committing suicide”, Cry: “I often cry”, Remorse: “Wrong things are caused by me”, Lonely: “I am lonely”, Unmotivated: “I am not motivated”, Pessimistic: “I am not optimistic”, Tough: “Everything is tough”, and Insomnia: “Cannot easily fall asleep or wake up”). Reassuringly, the estimated main effects (β_1) in Table 1.9 all fall within the ranges given by the corresponding coefficients found in the various depressive symptoms regressions and are significant throughout. The index variable regressions hence reconfirm the robustness of the results.

An increase in spring season PM10 and CO levels not only increases depressive symptoms of adolescents but also aggravates aggressive symptoms. Table 1.10 displays the 2SLS regression results for aggressive symptoms index variable (ARGIN) using *Gobi Wind* and *Wind Speed* as instruments. The general structure of the table follows the presentation in Table 1.9, i.e. regression results is shown by displaying only the main coefficients (β_1).

Table 1.10: Regression Results Using Aggressive Symptoms Index Variable
(Instruments: *Gobi Wind* and *Wind Speed*)

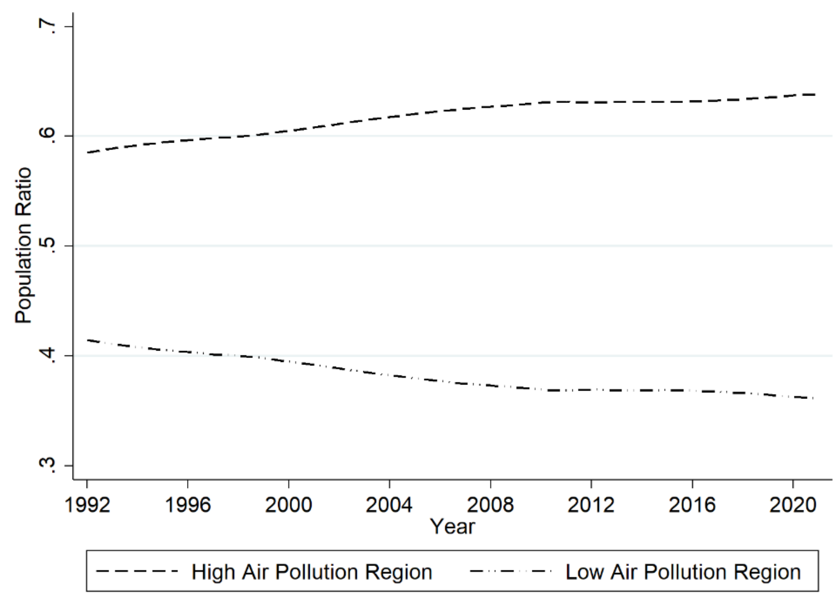
Model	**7th Grade**		**10th Grade**	
	PM10	**CO**	**PM10**	**CO**
2SLS	-0.0163	-0.7482	-0.0165	-1.3552
	(0.0075)**	(0.4211)*	(0.0077)**	(0.4368)***
P>|t|	0.029	0.076	0.033	0.002
	[0.0083]**	[0.4735]	[0.0093]*	[0.5318]**
P>|t|	0.049	0.114	0.077	0.011
	{0.0073}**	{0.4138}*	{0.0101}	{0.6086}**
P>|t|	0.026	0.071	0.103	0.026
Oprobit 2SLS	-0.0260**	-1.1928*	-0.0277**	-2.2094***
	(0.0115)	(0.6658)	(0.0127)	(0.7169)
P>|t|	0.024	0.073	0.029	0.002
Stand. Coef (2SLS)	-0.1418	-0.0630	-0.1421	-0.1160
Stand. Coef (Oprobit)	-0.1633	-0.0727	-0.1653	-0.1307
Observations	1,776	1,776	1,655	1,655
[School]	550	550	787	787
{Living District}	128	128	125	125

Note: Standard errors in (parentheses). Standard errors using school level clustering in [brackets]. Standard errors using living district level clustering in {braces}. *,**,*** denote statistical significance at the 10%, 5% and 1% level.

The increase in springtime PM10 and CO levels also affects aggressive symptoms. A one standard deviation increase in PM10 levels worsens the aggressive symptoms approximately by 0.14 standard deviations for seventh-grade students and tenth-grade students. Moreover, a one standard deviation increase in CO levels worsens the aggressive symptoms approximately by 0.06 standard deviations for seventh-grade students and 0.12 standard deviations for tenth-grade students. Similar to the depressive symptoms regressions, standardized coefficients show that PM10 has a bigger effect on aggressive symptoms than CO. The ordered probit regression results re-confirm the main findings as well. Finally, to check which of the five aggressive symptoms are affected, I also run separate regressions on each of the five aggressive symptoms. 2SLS regression results are shown in Appendix Table 1.21; 2SLS regression results with different levels of clusters, in Appendix Table 1.22; 2SLS ordered probit regression results, in Appendix Table 1.23; standardized coefficients in Appendix Table 1.24. As Appendix Tables 1.21, 1.22, 1.23, and 1.24 show, all aggressive symptoms (Irritable: “I become irritated even on small things", Disturb: "I disturb or annoy friends", Aggress: "I become aggressive if I cannot do things in my way", Fight: "I often fight on trivial matters", and Angry: "I am angry all the time") of adolescents are affected negatively.

One possible threat to the identification strategy is the possibility of families moving from more to less polluted regions. To check this, I first separated the seventeen South Korean administrative divisions into two regions with high respectively low air pollution levels. Then, I tracked the population ratios of the two regions from 1992 to 2021 as shown in Figure 1.2.[18]

Figure 1.2: Population Ratios of Regions with High and Low Air Pollution Levels



Note: The detailed regional information is shown in Appendix Table 1.25.

Interestingly, as Figure 1.2 shows, the share of the overall Korean population exposed to more air pollution gradually increased between 1992 and 2021. Thus, I argue that avoidance behavior such as pollution induced migration did not play a substantial role during that period. I also checked the distributions of students across the two regions in the KCYPS dataset between 2010

[18] Figure 2 displays the shares of the population living in the two regions based on the Resident Registry Population dataset provided by the Korean Statistical Information Service. The dataset is publicly available at: https://kosis.kr/statHtml/statHtml.do?orgId=202&tblId=DT_202N_B4&conn_path=I3 [last accessed: January 2nd, 2023].

and 2016 (see Appendix Table 1.25). Reassuringly, I do not observe any trend favoring the region with lower air pollution. Another possible threat to identification is avoidance behavior through long-distance traveling during the spring season. Endogeneity issues caused by long-distance traveling should not be a problem in the case of Korea as secondary schools start their academic year in March, and students have spring semester midterm examinations in mid-May and finals in late June on ten or more subjects. Considering the competitiveness of Korean education (for high-ranked high schools and university entrance) long-distance travelling during the spring season is unlikely (Anderson & Kohler, 2012; Lee, 2005).

I also conduct a number of additional robustness checks. First, as mentioned in the methodology section, I also check the robustness of the results when using *West Wind* and *Wind Speed* as alternative instruments. Reassuringly, Appendix Table 1.26 in the Appendix show that the estimated magnitudes of coefficients and statistical significance levels are similar to the main regressions as displayed in Tables 1.9 and 1.10. Likewise the regression results for each eleven depressive symptoms and five aggressive symptoms also reconfirm the robustness of the results when using the alternative pair of instruments (see Appendix Table 1.27). Second, I include observations from 2015 to construct a panel which allows me to control for unobserved time-invariant explanatory variables such as distinct local differences with respect to urbanization or regional identity.[19] Using the panel, I analyze the effect of year-by-province-level variations of springtime PM10 and CO levels on student mental health outcomes. In a first step, I rerun the cross-section regressions using 2015 instead of 2016 data. Then I estimate three panel specifications as follows: Panel 0 does not consider local fixed effects, Panel 1 controls for

[19] Students (and parents) may make different lifestyle choices and perform different leisure activities based on how urbanized their environments are. In addition, Koreans have strong regional identities that are rooted in the country's long history. During the Chosŏn dynasty, present day South Korea was divided into six regions with similar cultures, dialects, and social attitudes. These regional peculiarities remain to the present day and are revealed, for example, in nation-wide elections but are not restricted to voting behavior.

urbanization dummies, and Panel 2 uses region fixed effects (see Appendix Tables 1.28 to 1.35).[20] Once more the results do not change substantially, regardless of the chosen model.[21] Third, I adjust the obtained p-values for the false discovery rates (FDR). In fact testing multiple hypotheses one by one leads to an increased likelihood of false rejections. To reduce this likelihood, I follow Anderson (2008) and calculate the sharpened q-values for the main coefficients of all mental health symptoms regressions (see Appendix Tables 1.36 and 1.37). The sharpened q-values obtained reconfirm the original findings that air pollution levels (PM10 and CO) indeed have negative effects on adolescent mental health, in particular for seventh-grade students. Lastly, I use the depressive and aggressive symptoms index variables to expand the panel analyses. Appendix Tables 1.38 and 1.39 present panel regression results when controlling for a growing number of fixed effects (Panel 0, 1, and 2) and assuming bigger clusters of standard errors (Panel 3 and 4). A number of regressions using panel and cross-sectional settings exhibit similar results as they are mainly driven by the regional differences in air pollution levels (Kwak et al., 2022). Therefore, choosing the 2016 cross-section for the baseline regressions seems more relevant as it requires a weaker identifying assumptions.

1.5 Discussion

My analysis reveals a strong effect of air pollution on the mental health of South Korean adolescents. As Tables 1.9 and 1.10 show, springtime PM10 and CO levels negatively affect the

[20] The urbanization dummies of regression model Panel 1 reflect three levels of urbanization: first the most urbanized capital city Seoul, second the other metropolitan cities Incheon, Daejeon, Sejong, Gwangju, Daegu, Ulsan, and Busan, and third rural provinces. The region fixed effects of model Panel 2 refer to the following six historical regions. (1) **Capital region:** Seoul, Incheon, Gyeonggi; (2) **Chungcheong:** South Chungcheong, North Chungcheong, Sejong, Daejeon; (3) **Jeolla:** South Jeolla, North Jeolla, Gwangju; (4) **Gyeongsang:** South Gyeongsang, North Gyeongsang, Busan, Ulsan, Daegu; (5) **Gangwon:** Gangwon; (6) **Jeju**: Jeju.

[21] The aforementioned stability of the yellow dust dispersion in Korea results in only small variations of PM10 and CO levels over time (Kwak et al., 2022). In the panel analysis it is therefore advisable not to use too many fixed effects (such as first-differenced models or individual fixed effects). Instead the regional variations of PM10 and CO levels can and should be exploited. In other words, using first differences of springtime pollution levels or regional fixed effects for all 17 regions is no reasonable identification strategy.

mental health of secondary students with high statistical significance. I also show that the effect is more pronounced among seventh-grade students (relative to tenth-grade students) which is in line with previous results in the literature. In general, younger people tend to show stronger adverse psychological responses (Armstrong et al., 2000; Power et al., 2020). Although a lack of data does not allow me to analyze the effect of air pollution on children, regression results hint that children are more likely to show stronger responses to air pollution.

Possible explanations for the adverse effect of air pollution on adolescent depressive and aggressive symptoms might be medical in nature. For example, it is well established that Diesel-exhaust particles trigger the release of proinflammatory factors and reactive oxygen species, which subsequently can cause severe mental health problems (Block et al., 2004; Sui et al., 2018). More generally, air pollution increases the probability of systemic oxidative stress (Kelly 2003; Risom et al., 2005), which directly raises the risk for depression (Ng et al., 2008; Yanik et al., 2004). Because most of the outcome variables (i.e., Unproductive, Depressed, Anxious, Suicide, Remorse, Unmotivated, Pessimistic, Tough, and Insomnia) all correlate with the Hamilton Depression Rating Scale, my results are consistent with the explanations by Ng et al. (2008) and Yanik et al. (2004).

Another possible explanation of how air pollution may harm adolescent mental health is via its negative impact on sleep quality (Becker et al., 2017; Hayashino et al., 2010). Again, the regression results for Insomnia ("Cannot easily fall asleep or wake up") support this view. Last but not least, high levels of air pollution reduce the white surface matter in the left hemisphere of the brain of children and adolescents (Binter et al., 2022; Lopuszanska & Samardakiewicz, 2020; Peterson et al., 2015; Roberts et al., 2019) which causes cognitive as well as behavioral problems among the affected including attention-deficit/hyperactivity disorder symptoms and conduct disorder problems (Peterson et al., 2015). Common features of conduct disorder problems are being irresponsible, skipping school, stealing (or violating the rights of others), and physically harming other people or animals (Johns Hopkins Medicine, 2023). The findings for Disturb ("I disturb or annoy friends") and for Fight ("I often fight on trivial matters") fit squarely into this literature. Moreover, a number of very recent studies points into the same direction by establishing a connection between exposure to air pollution and crimes (Bondy et al., 2020; Burkhardt et al., 2019; Herrnstadt et al., 2021). Finally, mental health problems caused by air pollution might be amplified through peer effects (Hill, 2002; Quinton et al., 1993).

An alternate mechanism, distinct from the medical literature, may also negatively affect adolescent mental health. Serious air pollution exerts an impact on the cognitive abilities of individuals (Peterson et al., 2015), thereby potentially leading to diminished academic achievements among students (Balakrishnan & Tsaneva, 2021; Lu et al., 2021; Shier et al., 2019). The decline in academic performance can subsequently exert adverse effects on the mental well-being of students, inducing feelings of discouragement or frustration (Miles & Stipek, 2006; Paris et al., 1991). While the exploration of these intricate pathways falls beyond the scope of this paper, the presence of robust and statistically significant regression results on depressive and aggressive symptoms indeed underscores the necessity for further investigation.

My findings contribute to a growing body of literature that identifies air pollution effects on the health outcomes of adolescents. In particular, this study adds the mental health of adolescents to the spectrum of outcome variables. Although the effects on physical health (Chay & Greenstone, 2003; Chay et al., 2003; Currie & Neidell, 2005; Currie et al., 2009; Currie & Walker, 2011; Chen et al., 2013; Sanders & Stoecker, 2015; Deschênes et al., 2017; Deryugina et al., 2019) and adult mental health (Bishop et al., 2017; Buoli et al., 2018; Chen et al., 2018; Gu et al., 2020; Bakolis et al., 2020) have been investigated, research on adolescent mental health is lacking. This study fills this gap by producing evidence that air pollution does indeed adversely affect the mental health of young South Koreans.

Finally, this study complements an already rich body of literature about the effects of air pollution on health outcomes in Western countries by looking into a country in highly polluted East Asia. As I have argued above, this addition to the literature is particularly interesting because polluted air affects health outcomes most likely in a non-linear way (Arceo et al., 2016). Hence my findings offer a first glimpse into the health effects of air pollution in highly polluted areas.

1.6 Conclusion

South Korea has unique seasonal and remarkably stable wind patterns. For centuries, the dominating west winds during the spring season have been known to carry Yellow Dust from the Gobi Desert of China and Mongolia (Chun, 2004). With the beginning of the industrialization of East Asia, springtime winds started to carry air pollutants from the Chinese East coast and from the Korean West coast to the inner Korean peninsula. My analysis exploits the seasonal and regional consistency of wind patterns in Korea by using wind directions and wind speeds as instruments for regional air pollution levels. The instrumental variable approach resolves possible endogeneity issues caused by the correlation between the air pollution levels (in particular PM10 and CO levels) and the economic situation. Since the associated F-statistics are exceptionally high and well above conventional thresholds for weak instruments, I have indeed found strong instruments. Moreover, I also conduct a number of robustness checks that corroborate the results. In summary, my analysis delivers reliable findings that might be helpful to researchers and policymakers alike.

A possible direction for future research is to conduct a mediation analysis. Although air pollution exerts a direct effect on mental health (Bishop et al., 2017; Buoli et al., 2018; Chen et al., 2018; Gu et al., 2020; Bakolis et al., 2020), there might also be indirect effects of air pollution on mental health especially when considering adolescents. If fine dust levels are high, local governments and school boards may – and often do – restrict outdoor activities for students. Without sufficient outdoor activities and physical education, however, students are more likely to develop depressive symptoms (McHale et al., 2001). Thus, the estimated coefficients likely capture a combination of the direct and indirect effects of air pollution on adolescent mental health. Mediation analysis could help to separate the two effects, assuming of course that a valid mediator can be found.

Another possible direction for future research would be to study the effect of air pollution on health outcomes among the very young (i.e., elementary school and kindergarten children). The difference between the estimated coefficients in the seventh and tenth-grade regressions suggests that the mental health effects of air pollution is more severe for younger children. Of course, it requires a rigorous econometric analysis and appropriate data to back up such a claim. As primary school students and kindergarten children may have difficulties to report their mental health statuses accurately, the information should ideally come from precise medical reports or hospital records. So far a lack of data prevents me from conducting such an analysis.

Further investigation into avoidance behavior could enhance the precision of the upper and lower bounds of estimated coefficients. The estimated coefficients are the overall impact, encompassing avoidance behaviors such as wearing facial masks and staying indoors. Therefore, the true direct effect of air pollution on the mental health of adolescents may, in fact, exceed the magnitude indicated by the estimated coefficients. Furthermore, as mentioned in the discussion section, the empirical analysis of air pollution effects on academic outcomes may enrich the current discourse by exploring potential mechanisms of air pollution affecting adolescent mental health.

This research analyzes a relatively underexplored cost of pollution, with a specific emphasis on its influence on the mental health of adolescents. Considering that adolescents are at a pivotal stage of neurological development, the consequence of air pollution on the mental well-being of this demographic may potentially be more profound than those observed in adults (Binter et al., 2022; Lamb & Murphy, 2013; Lopuszanska & Samardakiewicz, 2020; Paul et al., 2013; Peterson et al., 2015; Roberts et al., 2019; Slack & Webber, 2007). It is also noteworthy to acknowledge that adolescent mental health plays a central role in shaping the mental well-being of adults (Johnson et al., 2018; McLeod et al., 2016). Thus air pollution may have longer-run effects on academic achievement and accumulation of human capital. Indeed, adolescent mental health affected by air pollution should no longer be neglected.

1.7 Appendix

1.7.1 Additional Figures and Results

Figure 1.3: Regional Pollution Level and Wind Patterns

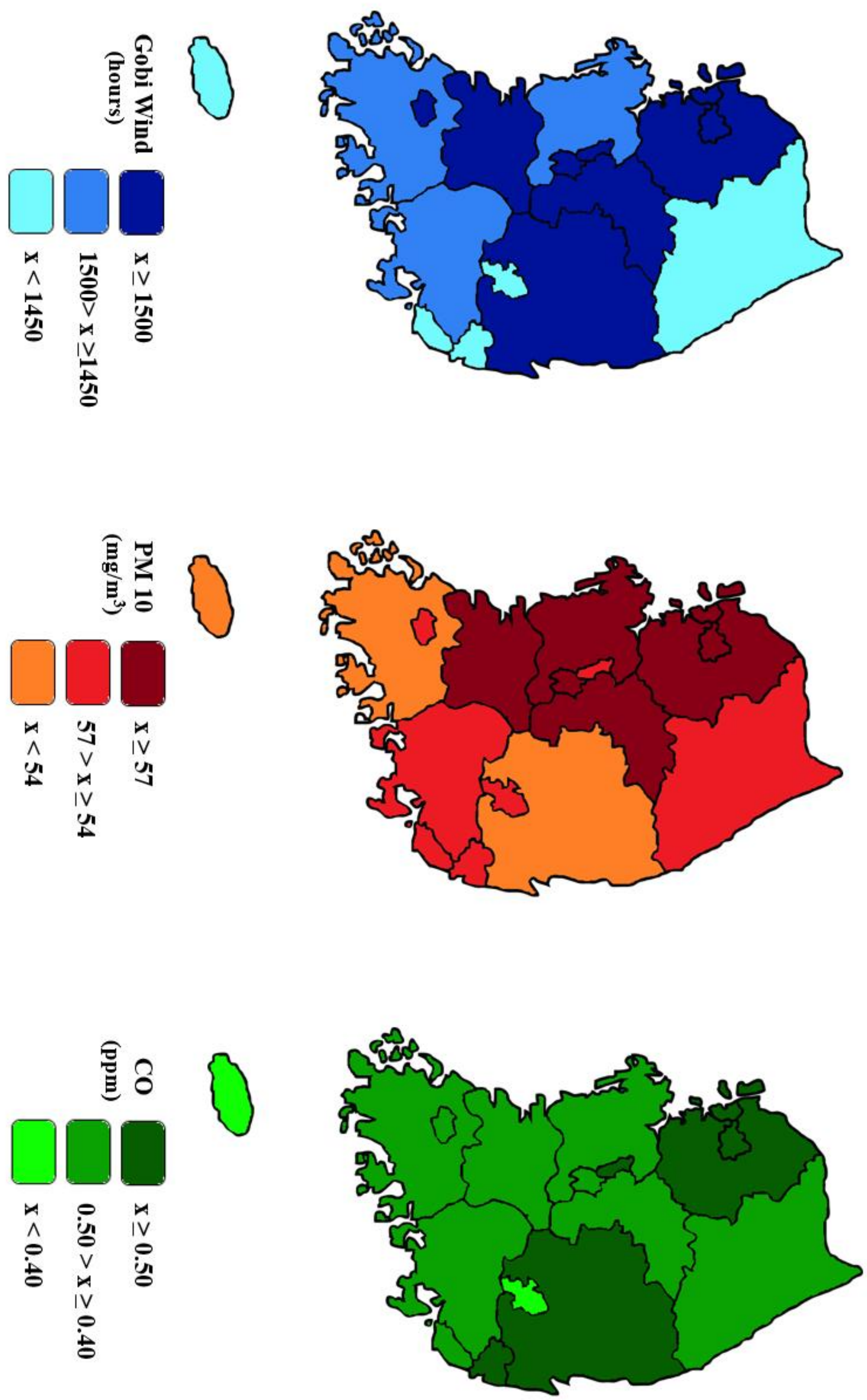


Table 1.11: Survey Questions on Depressive and Aggressive Symptoms

Q. These are questions about a student's usual behavior. Please respond to each item that applies to you.

I am…	**Strong Yes**	**Yes**	**No**	**Strong No**
I am not productive and don't have energy	1	2	3	4
I am depressed and sad	1	2	3	4
I am anxious	1	2	3	4
I talk about committing suicide	1	2	3	4
I often cry	1	2	3	4
I think wrong things caused by me	1	2	3	4
I am lonely	1	2	3	4
I am not motivated	1	2	3	4
I am not optimistic	1	2	3	4
I feel everything is tough	1	2	3	4
I cannot easily fall asleep or wake up	1	2	3	4
I become irritated even on small things	1	2	3	4
I disturb or annoy friends	1	2	3	4
I become aggressive if I cannot do things in my way	1	2	3	4
I often fight on trivial matters	1	2	3	4
I am angry all the time	1	2	3	4

Notes: survey questions are translated by the author.

Table 1.12: Endogeneity Tests Using Durbin and Wu-Hausman Tests

	7th Grade		**10th Grade**	
Variable	PM10	CO	PM10	CO
Unproductive	(18.870)***	(9.0446)***	(0.9849)	(0.3795)
	[18.943]***	[9.0294]***	[0.9783]	[0.3768]
Depressed	(2.7849)*	(1.0051)	(0.9122)	(0.3760)
	[2.7704]*	[0.9989]	[0.9061]	[0.3734]
Anxious	(3.4061)*	(4.4264)**	(0.4998)	(0.6962)
	[3.3896]*	[4.4075]**	[0.4964]	[0.6914]
Suicide	(0.8980)	(1.4163)	(1.9892)	(0.0473)
	[0.8924]	[1.4078]	[1.9771]	[0.0470]
Cry	(3.2813)*	(2.5162)	(4.9557)**	(0.7940)
	[3.2652]*	[2.5027]	[4.9346]**	[0.7887]
Remorse	(2.4485)	(0.4794)	(5.4904)**	(5.8098)**
	[2.4353]	[0.4763]	[5.4688]**	[5.7880]**
Lonely	(1.8587)	(0.0547)	(2.8313)*	(0.9820)
	[1.8481]	[0.0544]	[2.8156]*	[0.9755]
Unmotivated	(5.0888)**	(0.9395)	(0.6672)	(0.1069)
	[5.0690]**	[0.9336]	[0.6626]	[0.1061]
Pessimistic	(5.0418)**	(0.7914)	(3.7119)*	(1.1972)
	[5.0220]**	[0.7864]	[3.6932]*	[1.1893]
Tough	(4.4521)**	(0.3763)	(4.6322)**	(2.7416)*
	[4.4331]**	[0.3738]	[4.6115]**	[2.7262]*
Insomnia	(4.6491)**	(3.7242)*	(1.0231)	(0.0423)
	[4.6298]**	[3.7068]*	[1.0164]	[0.0420]
Irritable	(4.3635)**	(4.0903)**	(0.7314)	(2.4446)
	[4.3447]**	[4.0720]**	[0.7265]	[2.4305]
Disturb	(7.2816)***	(3.2540)*	(1.7488)	(3.0969)*
	[7.2622]***	[3.2380]*	[1.7379]	[3.0802]*
Aggress	(3.6879)*	(0.8107)	(1.7949)	(1.5244)
	[3.6706]*	[0.8056]	[1.7838]	[1.5147]
Fight	(3.8252)*	(4.9059)**	(3.3201)*	(6.0748)**
	[3.8075]*	[4.8862]**	[3.3027]*	[6.0529]**
Angry	(3.2248)*	(0.2763)	(3.8197)*	(4.5607)**
	[3.2089]*	[0.2744]	[3.8008]*	[4.5402]**
Observations	1776	1776	1655	1655

Note: Durbin Scores ~ $\chi^2(1)$ in (parentheses) and Wu-Hausman test statistics ~ F(1, Observations-12) in [brackets], testing the null hypotheses that variables are exogenous. *** denote p-value <0.01; **, p-value < 0.05; *, p-value < 0.1.

Table 1.13: OLS Regression Results with Spring PM10 and CO

Model	**7th Grade**		**10th Grade**	
	PM10	**CO**	**PM10**	**CO**
		Depressive Symptoms		
OLS	-0.0068	-0.5835	-0.0036	-0.4758
	(0.0026)***	(0.2649)**	(0.0029)	(0.2810)*
P>|t|	0.009	0.028	0.203	0.091
	[0.0027]**	[0.3107]*	[0.0028]	[0.2714]*
P>|t|	0.011	0.061	0.192	0.080
	{0.0027}**	{0.3002}*	{0.0031}	{0.2641}*
P>|t|	0.014	0.054	0.242	0.074
Oprobit	-0.0112**	-0.9886**	-0.0064	-0.8489*
	(0.0044)	(0.4556)	(0.0049)	(0.4853)
P>|t|	0.011	0.030	0.196	0.080
Stand.Coef (OLS)	-0.0627	-0.0520	-0.0322	-0.0417
Stand.Coef (Oprobit)	-0.0704	-0.0602	-0.0380	-0.0502
		Aggressive Symptoms		
OLS	-0.0012	-0.2358	-0.0035	-0.4304
	(0.0027)	(0.2802)	(0.0030)	(0.2898)
P>|t|	0.667	0.400	0.236	0.138
	[0.0029]	[0.3445]	[0.0028]	[0.3051]
P>|t|	0.681	0.494	0.217	0.159
	{0.0031}	{0.3095}	{0.0032}	{0.3708}
P>|t|	0.705	0.448	0.273	0.248
Oprobit	-0.0018	-0.3440	-0.0061	-0.7181
	(0.0043)	(0.4382)	(0.0049)	(0.4768)
P>|t|	0.678	0.432	0.206	0.132
Stand.Coef (OLS)	-0.0103	-0.0199	-0.0302	-0.0368
Stand.Coef (Oprobit)	-0.0112	-0.0210	-0.0366	-0.0425
Observations	1,776	1,776	1,655	1,655
[School]	550	550	787	787
{Living District}	128	128	125	125

Note: Standard errors in (parentheses). Standard errors using school level clustering in [brackets]. Standard errors using living district level clustering in {braces}. *,**,*** denote statistical significance at the 10%, 5% and 1% level.

Table 1.14: First Stage Regression Results (with *Gobi Wind*)

Variable	PM10		CO	
	7th Grade	10th Grade	7th Grade	10th Grade
Wind Speed	0.3290	0.7060	0.0368	0.0407
	(0.2745)	(0.2602)	(0.0022)	(0.0022)
	[0.4992]	[0.3169]	[0.0029]	[0.0022]
	{0.9537}	{0.7214}	{0.0052}	{0.0050}
Gobi Wind	0.0214	0.0201	0.00039	0.00036
	(0.0013)	(0.0012)	(0.00001)	(0.00001)
	[0.0022]	[0.0015]	[0.00003]	[0.00002]
	{0.0039}	{0.0036}	{0.00006}	{0.00005}
lnIncome	0.2585	-1.3013	0.0051	-0.0028
	(0.3436)	(0.3337)	(0.0027)	(0.0028)
	[0.4403]	[0.3324]	[0.0033]	[0.0031]
	{0.4733}	{0.4029}	{0.0039}	{0.0041}
lnAllow	-0.5780	0.7681	0.0023	0.0091
	(0.2194)	(0.2430)	(0.0017)	(0.0020)
	[0.2660]	[0.2833]	[0.0029]	[0.0028]
	{0.3327}	{0.4128}	{0.0038}	{0.0041}
Study	0.0025	0.0075	-0.00001	0.00002
	(0.0016)	(0.0014)	(0.00001)	(0.00001)
	[0.0018]	[0.0017]	[0.00001]	[0.00001]
	{0.0018}	{0.0025}	{0.00001}	{0.00002}
Game	0.0038	0.0146	-0.00007	0.00002
	(0.0020)	(0.0021)	(0.00002)	(0.00002)
	[0.0021]	[0.0024]	[0.00002]	[0.00002]
	{0.0027}	{0.0034}	{0.00003}	{0.00002}
Momeduc				
2	-0.2574	0.5454	-0.0019	0.0042
	(1.4751)	(0.8963)	(0.0116)	(0.0074)
	[1.5957]	[0.7660]	[0.0082]	[0.0086]
	{1.4217}	{0.8231}	{0.0087}	{0.0090}
3	-0.9115	0.8029	-0.0039	0.0011
	(1.4991)	(0.9713)	(0.0118)	(0.0080)
	[1.6077]	[0.8655]	[0.0085]	[0.0089]
	{1.4596}	{0.9833}	{0.0097}	{0.0095}
4	0.4596	1.6142	-0.0006	0.0033
	(1.4920)	(0.9270)	(0.0118)	(0.0077)
	[1.6217]	[0.8135]	[0.0083]	[0.0088]
	{1.5505}	{0.9151}	{0.0097}	{0.0097}
5	1.5736	3.6443	0.0012	0.0119
	(1.5804)	(1.0836)	(0.0125)	(0.0089)
	[1.7303]	[0.9499]	[0.0097]	[0.0100]
	{1.6651}	{1.1795}	{0.0111}	{0.0117}
Gender	0.2600	-0.0374	-0.0006	-0.0005
	(0.2772)	(0.2636)	(0.0022)	(0.0022)
	[0.3840]	[0.3725]	[0.0034]	[0.0038]
	{0.2082}	{0.2419}	{0.0016}	{0.0020}
Constant	24.803	35.341	-0.2238	-0.1448
	(3.8547)	(3.4700)	(0.0304)	(0.0286)
	[5.2443]	[3.7460]	[0.0526]	[0.0375]
	{7.7180}	{6.9465}	{0.0971}	{0.0763}
Observations	1776	1655	1776	1655
[School]	550	787	550	787
{Living District}	128	125	128	125

Note: Standard errors in (parentheses). Standard errors using school level clustering in [brackets]. Standard errors using living district level clustering in {braces}.

Table 1.15: First Stage Regression Results (with *West Wind*)

Variable	**PM10**		**CO**	
	7th Grade	10th Grade	7th Grade	10th Grade
Wind Speed	-2.4495	-1.9152	-0.0046	0.0008
	(0.2447)	(0.2345)	(0.0023)	(0.0023)
	[0.4821]	[0.3044]	[0.0045]	[0.0032]
	{0.8547}	{0.6877}	{0.0082}	{0.0074}
West Wind	0.0215	0.0192	0.00023	0.00022
	(0.0009)	(0.0009)	(0.00001)	(0.00001)
	[0.0018]	[0.0011]	[0.00002]	[0.00001]
	{0.0032}	{0.0027}	{0.00004}	{0.00003}
lnIncome	0.0999	-1.5722	0.0026	-0.0075
	(0.3206)	(0.3149)	(0.0031)	(0.0030)
	[0.4022]	[0.3185]	[0.0033]	[0.0032]
	{0.4597}	{0.3959}	{0.0039}	{0.0052}
lnAllow	-0.1368	0.7417	0.0060	0.0091
	(0.2060)	(0.2296)	(0.0020)	(0.0022)
	[0.2993]	[0.2862]	[0.0029]	[0.0031]
	{0.3786}	{0.4610}	{0.0030}	{0.0054}
Study	0.0019	0.0060	-0.00001	0.00002
	(0.0015)	(0.0014)	(0.00001)	(0.00001)
	[0.0016]	[0.0016]	[0.00001]	[0.00002]
	{0.0016}	{0.0022}	{0.00001}	{0.00001}
Game	0.0029	0.0121	-0.00007	0.00001
	(0.0019)	(0.0020)	(0.00002)	(0.00002)
	[0.0021]	[0.0021]	[0.00003]	[0.00002]
	{0.0029}	{0.0031}	{0.00004}	{0.00002}
Momeduc				
2	-0.6458	1.1364	-0.0065	0.0128
	(1.3765)	(0.8467)	(0.0132)	(0.0081)
	[1.4359]	[0.7857]	[0.0081]	[0.0106]
	{1.2827}	{0.7579}	{0.0102}	{0.0117}
3	-1.2370	1.5403	-0.0091	0.0085
	(1.3988)	(0.9185)	(0.0134)	(0.0088)
	[1.4652]	[0.8824]	[0.0086]	[0.0106]
	{1.3341}	{0.8945}	{0.0121}	{0.0111}
4	-0.3148	1.9706	-0.0080	0.0085
	(1.3927)	(0.8757)	(0.0133)	(0.0084)
	[1.4631]	[0.8343]	[0.0079]	[0.0107]
	{1.4136}	{0.8591}	{0.0101}	{0.0121}
5	0.3025	3.4713	-0.0095	0.0126
	(1.4763)	(1.0238)	(0.0141)	(0.0098)
	[1.5555]	[0.9706]	[0.0099]	[0.0120]
	{1.5298}	{1.1743}	{0.0118}	{0.0146}
Gender	0.2203	-0.0566	-0.0014	-0.0009
	(0.2587)	(0.2490)	(0.0025)	(0.0024)
	[0.3769]	[0.3838]	[0.0044]	[0.0048]
	{0.1811}	{0.2239}	{0.0013}	{0.0022}
Constant	36.060	48.0995	0.1690	0.2096
	(3.1110)	(2.7934)	(0.0298)	(0.0268)
	[4.0504]	[2.8627]	[0.0340]	[0.0347]
	{5.0451}	{4.4186}	{0.0496}	{0.0681}
Observations	1776	1655	1776	1655
[School]	550	787	550	787
{Living District}	128	125	128	125

Note: Standard errors in (parentheses). Standard errors using school level clustering in [brackets]. Standard errors using living district level clustering in {braces}.

Table 1.16: Correlation with Instrument Variables and
Other Variables (Weather, Economic Activities, and Demographics)

Variables	Rain Fall	Reg. Employment	Reg. GDP	Population
		2015		
Gobi Wind	-0.0002	-0.0001	0.0004	0.0007
(hour)	(0.0007)	(0.0001)	(0.0021)	(0.0022)
	Retain θ_1=0	Retain θ_1=0	Retain θ_1=0	Retain θ_1=0
Wind Speed	0.0041	-0.0006	0.1000	0.1275
(km/hour)	(0.0365)	(0.0048)	(0.1022)	(0.1058)
	Retain θ_2=0	Retain θ_2=0	Retain θ_2=0	Retain θ_2=0
		2016		
Gobi Wind	-0.0007	$5\times(0.1)^6$	0.0015	0.0008
(hour)	(0.0005)	(0.0001)	(0.0022)	(0.0022)
	Retain θ_1=0	Retain θ_1=0	Retain θ_1=0	Retain θ_1=0
Wind Speed	-0.0139	0.0014	0.0969	0.1172
(km/hour)	(0.0278)	(0.0059)	(0.1168)	(0.1175)
	Retain θ_2=0	Retain θ_2=0	Retain θ_2=0	Retain θ_2=0
		2015 – 2016		
Gobi Wind	-0.0006	$5\times(0.1)^5$	0.0009	0.0007
(hour)	(0.0005)	(0.0001)	(0.0014)	(0.0014)
	Retain θ_1=0	Retain θ_1=0	Retain θ_1=0	Retain θ_1=0
Wind Speed	-0.0118	$8\times(0.1)^6$	0.0925	0.1197
(km/hour)	(0.0254)	(0.0036)	(0.0722)	(0.0735)
	Retain θ_2=0	Retain θ_2=0	Retain θ_2=0	Retain θ_2=0

Note: Retain the null hypothesis ($\theta_i = 0$) if p>0.10 after regressing $y = \theta_0 + \theta_1$ Gobi Wind + θ_2 Wind Speed + ε where y is either ln(Rain Fall), ln(Employment Rate) ln(Population), or ln(Regional GDP) in the first-tier administrative division level. The first two panels show the results after regressing year 2015 and year 2016 separately, while the third panel shows the results after appending the two years. Therefore, the number of observations of the first two panels is seventeen; the third panel is thirty four. The statistical test results on (Retaining $\theta_i = 0$) if p>0.10 holds even if the outcomes are in levels.

Table 1.17: Depressive Symptoms 2SLS Results

	7th Grade		10th Grade	
Variable	PM10	CO	PM10	CO
Unproductive	-0.0394***	-2.1034***	-0.0106	-0.6147
	(0.0081)	(0.4445)	(0.0081)	(0.4594)
P>|t|	0.000	0.000	0.191	0.181
Depressed	-0.0168**	-0.6579	-0.0110	-0.6260
	(0.0075)	(0.4218)	(0.0079)	(0.4471)
P>|t|	0.024	0.119	0.164	0.161
Anxious	-0.0262***	-1.0633**	-0.0120	-0.4825
	(0.0093)	(0.5274)	(0.0102)	(0.5798)
P>|t|	0.000	0.044	0.242	0.405
Suicide	-0.0137**	-0.7511**	-0.0141**	-0.5025
	(0.0060)	(0.3422)	(0.0068)	(0.3827)
P>|t|	0.023	0.028	0.037	0.189
Cry	-0.0162*	-0.3944	-0.0119	-0.5173
	(0.0086)	(0.4855)	(0.0086)	(0.4848)
P>|t|	0.059	0.417	0.168	0.286
Remorse	-0.0161*	-0.6003	-0.0246***	-1.4531***
	(0.0084)	(0.4730)	(0.0091)	(0.5124)
P>|t|	0.054	0.204	0.007	0.005
Lonely	-0.0160**	-0.6876	-0.0153*	-0.8928*
	(0.0081)	(0.4569)	(0.0092)	(0.5200)
P>|t|	0.047	0.132	0.097	0.086
Unmotivated	-0.0227***	-1.2667***	-0.0159**	-0.7701*
	(0.0070)	(0.3912)	(0.0074)	(0.4170)
P>|t|	0.001	0.001	0.030	0.065
Pessimistic	-0.0227***	-1.1538***	-0.0249***	-1.3195***
	(0.0076)	(0.4294)	(0.0091)	(0.5146)
P>|t|	0.003	0.007	0.006	0.010
Tough	-0.0240***	-1.0921***	-0.0213**	-1.3760***
	(0.0073)	(0.4138)	(0.0085)	(0.4754)
P>|t|	0.001	0.008	0.012	0.004
Insomnia	-0.0221**	-0.9326*	-0.0175**	-1.0805**
	(0.0087)	(0.4890)	(0.0088)	(0.4951)
P>|t|	0.011	0.057	0.045	0.029
Observations	1776	1776	1655	1655

Note: Standard errors in parentheses. *,**,*** denote statistical significance at the 10%, 5% and 1% level.

Table 1.18: Depressive Symptoms Using Different Level of Clusters

	7th Grade		10th Grade	
Variable	**PM10**	**CO**	**PM10**	**CO**
Unproductive	-0.0394 [0.0079]*** {0.0089}***	-2.1034 [0.4488]*** {0.5243}***	-0.0106 [0.0084] {0.0088}	-0.6147 [0.4918] {0.4887}
Depressed	-0.0168 [0.0076]** {0.0068}**	-0.6579 [0.4024] {0.4078}	-0.0110 [0.0079] {0.0076}	-0.6260 [0.4514] {0.4546}
Anxious	-0.0262 [0.0092]*** {0.0092}***	-1.0633 [0.5284]** {0.5906}*	-0.0120 [0.0104] {0.0110}	-0.4825 [0.5956] {0.6733}
Suicide	-0.0137 [0.0068]** {0.0067}**	-0.7511 [0.3768]** {0.3781}**	-0.0141 [0.0066]** {0.0082}*	-0.5025 [0.3712] {0.4400}
Cry	-0.0162 [0.0091]* {0.0096}*	-0.3944 [0.4969] {0.6510}	-0.0119 [0.0085] {0.0086}	-0.5173 [0.4830] {0.4718}
Remorse	-0.0161 [0.0089]* {0.0090}*	-0.6003 [0.4898] {0.5209}	-0.0246 [0.0104]** {0.0105}**	-1.4531 [0.5780]** {0.6167}**
Lonely	-0.0160 [0.0072]** {0.0068}**	-0.6876 [0.4143]* {0.4093}*	-0.0153 [0.0106] {0.0118}	-0.8928 [0.6171] {0.6240}
Unmotivated	-0.0227 [0.0067]*** {0.0054}***	-1.2667 [0.3843]*** {0.3202}***	-0.0159 [0.0083]* {0.0093}*	-0.7701 [0.4756] {0.4846}
Pessimistic	-0.0227 [0.0079]*** {0.0082}***	-1.1538 [0.4514]** {0.5108}**	-0.0249 [0.0098]** {0.0114}**	-1.3195 [0.5503]** {0.6142}**
Tough	-0.0240 [0.0071]*** {0.0067}***	-1.0921 [0.3999]*** {0.4079}***	-0.0213 [0.0102]** {0.0108}**	-1.3760 [0.5762]** {0.6049}**
Insomnia	-0.0221 [0.0089]** {0.0079}***	-0.9326 [0.4748]** {0.4317}**	-0.0175 [0.0092]* {0.0104}*	-1.0805 [0.5234]** {0.5870}*
[School]	550	550	787	787
{Living District}	128	128	125	125

Note: Standard errors using school level clustering in [brackets]. Standard errors using living district level clustering in {braces} *,**,*** denote statistical significance at the 10%, 5% and 1% level.

Table 1.19: 2SLS with Ordered Probit on Depressive Symptoms

	7th Grade		10th Grade	
Variable	PM10	CO	PM10	CO
Unproductive	-0.0617***	-3.4016***	-0.0173	-0.9383
	(0.0105)	(0.6693)	(0.0126)	(0.7173)
P>|t|	0.000	0.000	0.170	0.191
Depressed	-0.0302**	-1.1943	-0.0178	-0.9866
	(0.0119)	(0.6917)	(0.0129)	(0.7314)
P>|t|	0.011	0.084	0.167	0.177
Anxious	-0.0343***	-1.4223**	-0.0147	-0.5889
	(0.0112)	(0.6490)	(0.0121)	(0.6884)
P>|t|	0.002	0.028	0.225	0.392
Suicide	-0.0329**	-1.8284**	-0.0298**	-1.0834
	(0.0132)	(0.7592)	(0.0139)	(0.7988)
P>|t|	0.013	0.016	0.032	0.175
Cry	-0.0235**	-0.6176	-0.0173	-0.7619
	(0.0116)	(0.6654)	(0.0126)	(0.7244)
P>|t|	0.043	0.353	0.169	0.293
Remorse	-0.0229**	-0.8612	-0.0349***	-2.0496***
	(0.0115)	(0.6582)	(0.0121)	(0.6985)
P>|t|	0.046	0.191	0.004	0.003
Lonely	-0.0264**	-1.1761*	-0.0233*	-1.2955*
	(0.0118)	(0.6825)	(0.0124)	(0.7134)
P>|t|	0.026	0.085	0.060	0.069
Unmotivated	-0.0424***	-2.3434***	-0.0291**	-1.4115*
	(0.0119)	(0.7031)	(0.0131)	(0.7446)
P>|t|	0.000	0.001	0.026	0.058
Pessimistic	-0.0402***	-2.0389***	-0.0121**	-1.8450***
	(0.0116)	(0.6880)	(0.0048)	(0.7149)
P>|t|	0.001	0.003	0.012	0.010
Tough	-0.0411***	-1.8598***	-0.0339***	-2.1417***
	(0.0118)	(0.6916)	(0.0124)	(0.7173)
P>|t|	0.000	0.007	0.007	0.003
Insomnia	-0.0332***	-1.4483**	-0.0276**	-1.6541**
	(0.0115)	(0.6713)	(0.0128)	(0.7286)
P>|t|	0.004	0.031	0.031	0.023
Observations	1776	1776	1655	1655

Note: Standard errors in parentheses. *,**,*** denote statistical significance at the 10%, 5% and 1% level.

Table 1.20: Depressive Symptoms Standardized Coefficients

		7th Grade		10th Grade	
Variable	Regression Type	PM10	CO	PM10	CO
Unproductive	2SLS	-0.3266***	-0.1691***	-0.0874	-0.0501
	OProbit2SLS	-0.3879***	-0.2073***	-0.1031	-0.0555
Depressed	2SLS	-0.1466**	-0.0557	-0.0925	-0.0523
	OProbit2SLS	-0.1899**	-0.0728*	-0.1061	-0.0584
Anxious	2SLS	-0.1823***	-0.0716**	-0.0775	-0.0310
	OProbit2SLS	-0.2155***	-0.0867**	-0.0877	-0.0348
Suicide	2SLS	-0.1472**	-0.0783**	-0.1390**	-0.0490
	OProbit2SLS	-0.2068**	-0.1114**	-0.1776**	-0.0641
Cry	2SLS	-0.1201*	-0.0284	-0.0874	-0.0377
	OProbit2SLS	-0.1479**	-0.0376	-0.1033	-0.0451
Remorse	2SLS	-0.1246*	-0.0450	-0.1807***	-0.1060***
	OProbit2SLS	-0.1441**	-0.0525	-0.2082***	-0.1212***
Lonely	2SLS	-0.1279**	-0.0532	-0.1109*	-0.0640*
	OProbit2SLS	-0.1661**	-0.0717*	-0.1391*	-0.0766*
Unmotivated	2SLS	-0.2125***	-0.1149***	-0.1430**	-0.0686*
	OProbit2SLS	-0.2668***	-0.1428***	-0.1735**	-0.0835*
Pessimistic	2SLS	-0.1941***	-0.0958***	-0.1819***	-0.0957***
	OProbit2SLS	-0.2527***	-0.1243***	-0.2051**	-0.1091***
Tough	2SLS	-0.2131***	-0.0938***	-0.1686**	-0.1082***
	OProbit2SLS	-0.2583***	-0.1133***	-0.2018***	-0.1267***
Insomnia	2SLS	-0.1671**	-0.0682*	-0.1340**	-0.0819**
	OProbit2SLS	-0.2088***	-0.0883**	-0.1647**	-0.0978**

Note: *,**,*** denote statistical significance at the 10%, 5% and 1% level.

Table 1.21: Aggressive Symptoms 2SLS Results

	7th Grade		10th Grade	
Variable	PM10	CO	PM10	CO
Irritable	-0.0180**	-0.9819**	-0.0086	-0.8892*
	(0.0082)	(0.4617)	(0.0083)	(0.4723)
P>|t|	0.028	0.033	0.301	0.06
Disturb	-0.0186**	-0.9948**	-0.0073	-0.4516
	(0.0079)	(0.4408)	(0.0080)	(0.4543)
P>|t|	0.018	0.024	0.363	0.32
Aggress	-0.0183**	-0.8832**	-0.0129	-1.0063**
	(0.0078)	(0.4406)	(0.0084)	(0.4743)
P>|t|	0.019	0.045	0.125	0.034
Fight	-0.0124	-0.3448	-0.0166**	-1.0900**
	(0.0079)	(0.4453)	(0.0079)	(0.4462)
P>|t|	0.116	0.439	0.036	0.015
Angry	-0.0169**	-0.5619	-0.0229***	-1.6437***
	(0.0084)	(0.4718)	(0.0085)	(0.4805)
P>|t|	0.043	0.234	0.007	0.001
Observations	1776	1776	1655	1655

Note: Standard errors in parentheses. *,**,*** denote statistical significance at the 10%, 5% and 1% level.

Table 1.22: Aggressive Symptoms Using Different Level of Clusters

	7th Grade		10th Grade	
Variable	**PM10**	**CO**	**PM10**	**CO**
Irritable	-0.0180	-0.9819	-0.0086	-0.8892
	[0.0098]*	[0.5559]*	[0.0112]	[0.6344]
P>|t|	0.066	0.077	0.44	0.161
	{0.0096}*	{0.5370}*	{0.0097}	{0.6256}
P>|t|	0.060	0.067	0.372	0.155
Disturb	-0.0186	-0.9948	-0.0073	-0.4516
	[0.0085]**	[0.4653]**	[0.0098]	[0.5364]
P>|t|	0.028	0.033	0.457	0.400
	{0.0086}**	{0.4671}**	{0.0087}	{0.4893}
P>|t|	0.032	0.033	0.402	0.356
Aggress	-0.0183	-0.8832	-0.0129	-1.0063
	[0.0085]**	[0.4867]*	[0.0096]	[0.5599]*
P>|t|	0.031	0.070	0.180	0.072
	{0.0079}**	{0.4738}*	{0.0099}	{0.6338}
P>|t|	0.021	0.062	0.194	0.112
Fight	-0.0124	-0.3448	-0.0166	-1.0900
	[0.0086]	[0.4773]	[0.0101]*	[0.5762]*
P>|t|	0.151	0.470	0.100	0.059
	{0.0099}	{0.5615}	{0.0117}	{0.6862}
P>|t|	0.211	0.539	0.158	0.112
Angry	-0.0169	-0.5619	-0.0229	-1.6437
	[0.0091]*	[0.4946]	[0.0093]**	[0.5505]***
P>|t|	0.062	0.256	0.013	0.003
	{0.0073}**	{0.3998}	{0.0118}*	{0.6659}**
P>|t|	0.020	0.160	0.051	0.014
[School]	550	550	787	787
{Living District}	128	128	125	125

Note: Standard errors in parentheses. *,**,*** denote statistical significance at the 10%, 5% and 1% level.

Table 1.23: 2SLS with Ordered Probit on Aggressive Symptoms

	7th Grade		10th Grade	
Variable	PM10	CO	PM10	CO
Irritable	-0.0252**	-1.3854**	-0.0129	-1.2915*
	(0.0112)	(0.6450)	(0.0125)	(0.6990)
P>|t|	0.024	0.032	0.301	0.065
Disturb	-0.0281**	-1.5266**	-0.0117	-0.7218
	(0.0112)	(0.6515)	(0.0125)	(0.7065)
P>|t|	0.012	0.019	0.349	0.307
Aggress	-0.0275**	-1.3302**	-0.0195	-1.4857**
	(0.0114)	(0.6593)	(0.0124)	(0.7031)
P>|t|	0.016	0.044	0.116	0.035
Fight	-0.0197*	-0.5772	-0.0269**	-1.7256**
	(0.0115)	(0.6587)	(0.0125)	(0.7136)
P>|t|	0.086	0.381	0.032	0.016
Angry	-0.0258**	-0.8881	-0.0345***	-2.4158***
	(0.0115)	(0.6643)	(0.0126)	(0.7155)
P>|t|	0.025	0.181	0.006	0.001
Observations	1776	1776	1655	1655

Note: Standard errors in parentheses. *,**,*** denote statistical significance at the 10%, 5% and 1% level.

Table 1.24: Aggressive Symptoms Standardized Coefficients

		7th Grade		10th Grade	
Variable	Regression Type	PM10	CO	PM10	CO
Irritable	2SLS	-0.1432**	-0.0757**	-0.0688	-0.0703*
	OProbit2SLS	-0.1585**	-0.0844**	-0.0768	-0.0764*
Disturb	2SLS	-0.1536**	-0.0798**	-0.0605	-0.0371
	OProbit2SLS	-0.1766**	-0.0930**	-0.0695	-0.0427
Aggress	2SLS	-0.1529**	-0.0714**	-0.1026	-0.0795**
	OProbit2SLS	-0.1729**	-0.0811**	-0.1163	-0.0879**
Fight	2SLS	-0.1015	-0.0274	-0.1405**	-0.0918**
	OProbit2SLS	-0.1240*	-0.0352	-0.1603**	-0.1021**
Angry	2SLS	-0.1315**	-0.0423	-0.1796***	-0.1279***
	OProbit2SLS	-0.1625**	-0.0541	-0.2059***	-0.1429***

Note: *,**,*** denote statistical significance at the 10%, 5% and 1% level.

Table 1.25: Air Pollution and Student / Population Ratio

Panel 1: Classification of Regions According to Air Pollution Levels

Pollution Rank	High Air Pollution Region	Location
1 (PM10: 68.1)	Gyeonggi	North – West
2 (PM10: 66.2)	Seoul	North – West
3 (PM10: 65.5)	North Jeolla	Central – West
4 (PM10: 63.8)	Incheon	North – West
5 (PM10: 62.3)	North Chungcheong	Central – West
6 (PM10: 59.4)	South Chungcheong	Central – West
7 (PM10: 57.4)	Daejeon	Central – West

Pollution Rank	Low Air Pollution Region	Location
8 (PM10: 56.9)	Gangwon	North – East
9 (PM10: 56.7)	Gwangju	South – West
10 (PM10: 55.5)	Ulsan	South – East
11 (PM10: 55.4)	Daegu	South – East
12 (PM10: 55.4)	Sejong	Central – West
13 (PM10: 55.4)	Busan	South – East
14 (PM10: 54.4)	South Gyeongsang	South – East
15 (PM10: 53.4)	Jeju	South – West
16 (PM10: 50.3)	North Gyeongsang	Central – East
17 (PM10: 48.3)	South Jeolla	South – West

Panel 2: Distribution of Respondents in KCYPS Dataset

Year	High Air Pollution Region	Low Air Pollution Region
2010	53.9%	46.1%
2011	53.5%	46.5%
2012	53.4%	46.6%
2013	53.0%	47.0%
2014	53.7%	46.3%
2015	53.6%	46.4%
2016	53.5%	46.4%

Table 1.26: Depressive and Aggressive Symptoms Index Robustness Check
(Instruments: *West Wind* and *Wind Speed*)

	7th Grade		10th Grade	
Model	**PM10**	**CO**	**PM10**	**CO**
		DEPINDEX		
2SLS	-0.0151	-1.1829	-0.0112	-0.9829
	(0.0052)***	(0.4973)**	(0.0059)*	(0.4935)**
P>|t|	0.004	0.017	0.056	0.046
	[0.0053]***	[0.4812]**	[0.0060]***	[0.5251]*
P>|t|	0.004	0.014	0.063	0.061
	{0.0055}***	{0.5272}**	{0.0064}*	{0.5405}*
P>|t|	0.006	0.025	0.078	0.069
Oprobit 2SLS	-0.0271***	-2.1011**	-0.0200**	-1.7555**
	(0.0089)	(0.8551)	(0.0101)	(0.8481)
P>|t|	0.002	0.014	0.047	0.038
Stand. Coef (2SLS)	-0.1384	-0.1054	-0.0990	-0.0861
Stand. Coef (Oprobit)	-0.1704	-0.1281	-0.1192	-0.1038
		AGRINDEX		
2SLS	-0.0087	-0.6121	-0.0151	-1.7324
	(0.0055)	(0.5256)	(0.0061)**	(0.5116)***
P>|t|	0.114	0.244	0.013	0.001
	[0.0058]	[0.5637]	[0.0066]**	[0.5740]***
P>|t|	0.134	0.278	0.022	0.003
	{0.0061}	{0.5649}	{0.0071}**	{0.6783}**
P>|t|	0.150	0.279	0.034	0.011
Oprobit 2SLS	-0.0132	-0.8992	-0.0257***	-2.8400***
	(0.0086)	(0.8256)	(0.0099)	(0.8238)
P>|t|	0.127	0.276	0.010	0.001
Stand. Coef (2SLS)	-0.0756	-0.0516	-0.0990	-0.0861
Stand. Coef (Oprobit)	-0.0827	-0.0548	-0.1530	-0.1680
Observations	1,776	1,776	1,655	1,655
[School]	550	550	787	787
{Living District}	128	128	125	125

Note: Standard errors in (parentheses). Standard errors using school level clustering in [brackets]. Standard errors using living district level clustering in {braces}. *,**,*** denote statistical significance at the 10%, 5% and 1% level.

Table 1.27: Mental Health Symptoms Robustness Check
(Instruments: *West Wind* and *Wind Speed*)

	7th Grade		10th Grade	
Variable	**PM10**	**CO**	**PM10**	**CO**
Unproductive	-0.0200***	-1.7063***	-0.0117*	-0.9828*
	(0.0058)	(0.5539)	(0.0064)	(0.5370)
P>|t|	0.001	0.002	0.068	0.067
Depressed	-0.0070	-0.3216	-0.0111*	-0.9277*
	(0.0055)	(0.5265)	(0.0062)	(0.5224)
P>|t|	0.202	0.541	0.073	0.076
Anxious	-0.0186***	-1.2804*	-0.0134*	-0.8915
	(0.0069)	(0.6593)	(0.0080)	(0.6776)
P>|t|	0.007	0.052	0.096	0.188
Suicide	-0.0114**	-1.0411**	-0.0088*	-0.3937
	(0.0045)	(0.4279)	(0.0053)	(0.4469)
P>|t|	0.011	0.015	0.098	0.378
Cry	-0.0076	-0.1309	-0.0036	-0.1173
	(0.0063)	(0.6059)	(0.0067)	(0.5659)
P>|t|	0.231	0.82	0.597	0.836
Remorse	-0.0116*	-0.7499	-0.0200***	-1.7332***
	(0.0062)	(0.5908)	(0.0071)	(0.5992)
P>|t|	0.06	0.204	0.005	0.004
Lonely	-0.0105*	-0.7369	-0.0080	-0.6884
	(0.0060)	(0.5706)	(0.0072)	(0.6070)
P>|t|	0.079	0.197	0.269	0.257
Unmotivated	-0.0148***	-1.3589***	-0.0152***	-1.1151**
	(0.0051)	(0.4886)	(0.0058)	(0.4870)
P>|t|	0.004	0.005	0.008	0.022
Pessimistic	-0.0163***	-1.3817***	-0.0231***	-1.8209***
	(0.0056)	(0.5365)	(0.0072)	(0.6019)
P>|t|	0.004	0.01	0.001	0.002
Tough	-0.0168***	-1.2763**	-0.0149**	-1.4315***
	(0.0054)	(0.5170)	(0.0066)	(0.5552)
P>|t|	0.002	0.014	0.024	0.01
Insomnia	-0.0124*	-0.8022	-0.0207***	-1.8221***
	(0.0064)	(0.6104)	(0.0069)	(0.5792)
P>|t|	0.051	0.189	0.003	0.002
Irritable	-0.0109*	-0.9722*	-0.0123*	-1.4748***
	(0.0060)	(0.5765)	(0.0066)	(0.5531)
P>|t|	0.071	0.092	0.06	0.008
Disturb	-0.0137**	-1.2204**	-0.0065	-0.5801
	(0.0058)	(0.5510)	(0.0063)	(0.5307)
P>|t|	0.018	0.027	0.304	0.274
Aggress	-0.0102*	-0.7741	-0.0095	-1.0994**
	(0.0058)	(0.5501)	(0.0066)	(0.5539)
P>|t|	0.077	0.159	0.149	0.047
Fight	-0.0040	0.0288	-0.0140**	-1.3370***
	(0.0058)	(0.5553)	(0.0062)	(0.5218)
P>|t|	0.492	0.959	0.023	0.01
Angry	-0.0092	-0.4255	-0.0254***	-2.5010***
	(0.0062)	(0.5892)	(0.0067)	(0.5644)
P>|t|	0.134	0.47	0.000	0.000
Observations	1776	1776	1655	1655

Note: Standard errors in parentheses. *,**,*** denote statistical significance at the 10%, 5% and 1% level.

Table 1.28: Panel Analysis of Spring PM10 on Depressive Symptoms (7th Grade)

	2016	2015	Panel 0	Panel 1	Panel 2
Unproductive	-0.0394***	-0.0377***	-0.0381***	-0.0360***	-0.0354**
	(0.0081)	(0.0129)	(0.0075)	(0.0091)	(0.0175)
P>|t|	0.000	0.004	0.000	0.000	0.044
Std.Coef	-0.3266	-0.2188	-0.2997	-0.2836	-0.2785
Depressed	-0.0168**	-0.0455***	-0.0279***	-0.0253***	-0.0362**
	(0.0075)	(0.0129)	(0.0072)	(0.0087)	(0.0167)
P>|t|	0.024	0.000	0.000	0.004	0.030
Std.Coef	-0.1466	-0.2674	-0.2267	-0.2061	-0.2944
Anxious	-0.0262***	-0.0498***	-0.0352***	-0.0332***	-0.0408**
	(0.0093)	(0.0157)	(0.0093)	(0.0111)	(0.0203)
P>|t|	0.000	0.002	0.000	0.003	0.044
Std.Coef	-0.1823	-0.2354	-0.2290	-0.2161	-0.2656
Suicide	-0.0137**	-0.0234**	-0.0167***	-0.0141**	-0.0024
	(0.0060)	(0.0111)	(0.0059)	(0.0070)	(0.0141)
P>|t|	0.023	0.035	0.004	0.043	0.863
Std.Coef	-0.1472	-0.1567	-0.1605	-0.1358	-0.0235
Cry	-0.0162*	-0.0404***	-0.0258***	-0.0280***	-0.0673***
	(0.0086)	(0.0150)	(0.0083)	(0.0099)	(0.0197)
P>|t|	0.059	0.007	0.002	0.005	0.001
Std.Coef	-0.1201	-0.1975	-0.1765	-0.1912	-0.4605
Remorse	-0.0161*	-0.0328**	-0.0226***	-0.0207**	-0.0636***
	(0.0084)	(0.0143)	(0.0080)	(0.0096)	(0.0191)
P>|t|	0.054	0.022	0.005	0.032	0.001
Std.Coef	-0.1246	-0.1682	-0.1619	-0.1482	-0.4551
Lonely	-0.0160**	-0.0305**	-0.0215***	-0.0151*	-0.0374**
	(0.0081)	(0.0130)	(0.0075)	(0.0089)	(0.0171)
P>|t|	0.047	0.019	0.004	0.092	0.029
Std.Coef	-0.1279	-0.1738	-0.1646	-0.1153	-0.2863
Unmotivated	-0.0227***	-0.0310***	-0.0255***	-0.0195***	-0.0194
	(0.0070)	(0.0112)	(0.0063)	(0.0075)	(0.0148)
P>|t|	0.001	0.005	0.000	0.009	0.189
Std.Coef	-0.2125	-0.2078	-0.2287	-0.1752	-0.1747
Pessimistic	-0.0227***	-0.0398***	-0.0291***	-0.0252***	-0.0403**
	(0.0076)	(0.0123)	(0.0072)	(0.0086)	(0.0170)
P>|t|	0.003	0.001	0.000	0.003	0.017
Std.Coef	-0.1941	-0.2428	-0.2386	-0.2065	-0.3302
Tough	-0.0240***	-0.0293**	-0.0260***	-0.0203**	-0.0398**
	(0.0073)	(0.0121)	(0.0069)	(0.0082)	(0.0157)
P>|t|	0.001	0.015	0.000	0.013	0.011
Std.Coef	-0.2131	-0.1803	-0.2175	-0.1700	-0.3331
Insomnia	-0.0221**	-0.0255*	-0.0244***	-0.0222**	-0.0396*
	(0.0087)	(0.0153)	(0.0087)	(0.0102)	(0.0203)
P>|t|	0.011	0.094	0.005	0.030	0.051
Std.Coef	-0.1671	-0.1232	-0.1671	-0.1517	-0.2707
DEPINDEX	-0.0240***	-0.0408***	-0.0305***	-0.0281***	-0.0445***
	(0.0071)	(0.0120)	(0.0068)	(0.0081)	(0.0162)
P>|t|	0.001	0.001	0.000	0.001	0.006
Std.Coef	-0.2205	-0.2548	-0.2628	-0.2414	-0.3826
Region FE	No	No	No	No	Yes
Urbanization	No	No	No	Yes	No
Year FE	No	No	Yes	Yes	Yes
Observations	1776	1761	3537	3537	3537
Clusters (ID)	No	No	1929	1929	1929
First Stg F-Stat	139.643	104.604	331.44	217.519	104.827

Note: Column 2016 repeats the baseline results from Table 1.9. Column 2015 presents the results when using observations from 2015 instead of 2016. Columns Panel 0, 1 and 2 show the results from three different specifications of the panel model. Standard errors in parentheses. *,**,*** denote statistical significance at the 10%, 5% and 1% level.

Table 1.29: Panel Analysis of Spring CO on Depressive Symptoms (7^{th} Grade)

	2016	2015	Panel 0	Panel 1	Panel 2
Unproductive	-2.1034***	-1.4781***	-1.8089***	-1.8697***	-2.2687**
	(0.4445)	(0.5412)	(0.3813)	(0.4589)	(0.9029)
P>|t|	0.000	0.006	0.000	0.000	0.012
Std.Coef	-0.1691	-0.1362	-0.1567	-0.1620	-0.1965
Depressed	-0.6579	-1.8133***	-1.2275***	-1.1581***	-1.4437*
	(0.4218)	(0.5368)	(0.3636)	(0.4374)	(0.8705)
P>|t|	0.119	0.001	0.001	0.008	0.097
Std.Coef	-0.0557	-0.1693	-0.1099	-0.1037	-0.1293
Anxious	-1.0633**	-1.9345***	-1.4845***	-1.3234**	-1.4477
	(0.5274)	(0.6587)	(0.4704)	(0.5645)	(1.1245)
P>|t|	0.044	0.003	0.002	0.019	0.198
Std.Coef	-0.0716	-0.1454	-0.1062	-0.0947	-0.1036
Suicide	-0.7511**	-1.1907**	-0.9260***	-0.9377***	-0.0940
	(0.3422)	(0.4661)	(0.3040)	(0.3666)	(0.7339)
P>|t|	0.028	0.011	0.002	0.011	0.898
Std.Coef	-0.0783	-0.1268	-0.0981	-0.0993	-0.0100
Cry	-0.3944	-0.6035	-0.5593	-0.6989	-1.4398
	(0.4855)	(0.6326)	(0.4219)	(0.4971)	(1.0322)
P>|t|	0.417	0.340	0.185	0.160	0.163
Std.Coef	-0.0284	-0.0469	-0.0421	-0.0526	-0.1084
Remorse	-0.6003	-0.4501	-0.5825	-0.5273	-0.3921
	(0.4730)	(0.6077)	(0.4074)	(0.4896)	(0.9890)
P>|t|	0.204	0.459	0.153	0.281	0.692
Std.Coef	-0.0450	-0.0368	-0.0459	-0.0415	-0.0309
Lonely	-0.6876	-1.1873**	-0.9195**	-0.6409	-0.5080
	(0.4569)	(0.5463)	(0.3826)	(0.4515)	(0.9163)
P>|t|	0.132	0.030	0.016	0.156	0.579
Std.Coef	-0.0532	-0.1075	-0.0774	-0.0539	-0.0428
Unmotivated	-1.2667***	-1.3952***	-1.3365***	-1.2691***	-1.1675
	(0.3912)	(0.4681)	(0.3296)	(0.3947)	(0.7708)
P>|t|	0.001	0.003	0.000	0.001	0.130
Std.Coef	-0.1149	-0.1485	-0.1322	-0.1255	-0.1154
Pessimistic	-1.1538***	-1.5825***	-1.3876***	-1.3032***	-1.4223*
	(0.4294)	(0.5135)	(0.3636)	(0.4414)	(0.8480)
P>|t|	0.007	0.002	0.000	0.003	0.093
Std.Coef	-0.0958	-0.1534	-0.1250	-0.1174	-0.1281
Tough	-1.0921***	-0.9741*	-1.0529***	-0.8279**	-0.4443
	(0.4138)	(0.5110)	(0.3578)	(0.4206)	(0.8389)
P>|t|	0.008	0.057	0.003	0.049	0.596
Std.Coef	-0.0938	-0.0952	-0.0971	-0.0763	-0.0410
Insomnia	-0.9326*	-0.9901	-1.0185**	-1.2211**	-0.9749
	(0.4890)	(0.6504)	(0.4556)	(0.5491)	(1.0875)
P>|t|	0.057	0.128	0.025	0.026	0.370
Std.Coef	-0.0682	-0.0760	-0.0767	-0.0920	-0.0734
DEPINDEX	-1.1586***	-1.4209***	-1.3222***	-1.3515***	-1.4694*
	(0.3982)	(0.5011)	(0.3514)	(0.4193)	(0.8248)
P>|t|	0.004	0.005	0.000	0.001	0.075
Std.Coef	-0.1032	-0.1410	-0.1252	-0.1280	-0.1391
Region FE	No	No	No	No	Yes
Urbanization	No	No	No	Yes	No
Year FE	No	No	Yes	Yes	Yes
Observations	1776	1761	3537	3537	3537
Clusters (ID)	No	No	1929	1929	1929
First Stg F-Stat	696.053	258.499	985.956	788.846	193.251

Note: Column 2016 repeats the baseline results from Table 1.9. Column 2015 presents the results when using observations from 2015 instead of 2016. Columns Panel 0, 1 and 2 show the results from three different specifications of the panel model. Standard errors in parentheses. *,**,*** denote statistical significance at the 10%, 5% and 1% level.

Table 1.30: Panel Analysis of Spring PM10 on Aggressive Symptoms (7th Grade)

	2016	**2015**	**Panel 0**	**Panel 1**	**Panel 2**
Irritable	-0.0180**	-0.0291**	-0.0215***	-0.0202**	-0.0332*
	(0.0082)	(0.0133)	(0.0078)	(0.0094)	(0.0181)
P>|t|	0.028	0.028	0.006	0.031	0.068
Std.Coef	-0.1432	-0.1608	-0.1619	-0.1523	-0.2497
Disturb	-0.0186**	-0.0319**	-0.0228***	-0.0232**	-0.0288
	(0.0079)	(0.0128)	(0.0076)	(0.0091)	(0.0176)
P>|t|	0.018	0.012	0.003	0.011	0.102
Std.Coef	-0.1536	-0.1827	-0.1784	-0.1814	-0.2251
Aggress	-0.0183**	-0.0395***	-0.0257***	-0.0255***	-0.0453**
	(0.0078)	(0.0132)	(0.0074)	(0.0089)	(0.0178)
P>|t|	0.019	0.003	0.000	0.004	0.011
Std.Coef	-0.1529	-0.2234	-0.2007	-0.1990	-0.3537
Fight	-0.0124	-0.0444***	-0.0241***	-0.0284***	-0.0364**
	(0.0079)	(0.0132)	(0.0075)	(0.0089)	(0.0177)
P>|t|	0.116	0.001	0.001	0.001	0.040
Std.Coef	-0.1015	-0.2519	-0.1864	-0.2195	-0.2819
Angry	-0.0169**	-0.0366**	-0.0244***	-0.0290***	-0.0336*
	(0.0084)	(0.0146)	(0.0085)	(0.0103)	(0.0189)
P>|t|	0.043	0.012	0.004	0.005	0.076
Std.Coef	-0.1315	-0.1877	-0.1746	-0.2077	-0.2408
AGRINDEX	-0.0163**	-0.0387***	-0.0242***	-0.0261***	-0.0331*
	(0.0075)	(0.0125)	(0.0072)	(0.0087)	(0.0173)
P>|t|	0.029	0.002	0.001	0.003	0.056
Std.Coef	-0.1418	-0.2288	-0.1970	-0.2127	-0.2696
Region FE	No	No	No	No	Yes
Urbanization	No	No	No	Yes	No
Year FE	No	No	Yes	Yes	Yes
Observations	1776	1761	3537	3537	3537
Clusters (ID)	No	No	1929	1929	1929
First Stg F-Stat	139.643	104.604	331.44	217.519	104.827

Note: Column 2016 repeats the baseline results from Table 1.10. Column 2015 presents the results when using observations from 2015 instead of 2016. Columns Panel 0, 1 and 2 show the results from three different specifications of the panel model. Standard errors in parentheses. *,**,*** denote statistical significance at the 10%, 5% and 1% level.

Table 1.31: Panel Analysis of Spring CO on Aggressive Symptoms (7th Grade)

	2016	**2015**	**Panel 0**	**Panel 1**	**Panel 2**
Irritable	-0.9819**	-0.9641*	-0.9621**	-1.2441***	-1.5562*
	(0.4617)	(0.5653)	(0.3964)	(0.4782)	(0.9343)
P>|t|	0.033	0.088	0.015	0.009	0.096
Std.Coef	-0.0757	-0.0846	-0.0797	-0.1031	-0.1289
Disturb	-0.9948**	-1.0946**	-1.0319***	-1.3249***	-2.5071***
	(0.4408)	(0.5409)	(0.3827)	(0.4643)	(0.9044)
P>|t|	0.024	0.043	0.007	0.004	0.006
Std.Coef	-0.0798	-0.0996	-0.0887	-0.1139	-0.2156
Aggress	-0.8832**	-1.5509***	-1.1792***	-1.5005***	-1.7218**
	(0.4406)	(0.5553)	(0.3736)	(0.4508)	(0.8758)
P>|t|	0.045	0.005	0.002	0.001	0.049
Std.Coef	-0.0714	-0.1395	-0.1012	-0.1288	-0.1478
Fight	-0.3448	-1.4112***	-0.8565**	-1.3950***	-2.4044***
	(0.4453)	(0.5504)	(0.3821)	(0.4621)	(0.9062)
P>|t|	0.439	0.010	0.025	0.003	0.008
Std.Coef	0.0274	-0.1272	-0.0730	-0.1188	-0.2048
Angry	-0.5619	-1.4711**	-0.9946**	-1.4111***	-1.6382
	(0.4718)	(0.6135)	(0.4271)	(0.5212)	(1.0240)
P>|t|	0.234	0.016	0.020	0.007	0.110
Std.Coef	-0.0423	-0.1200	-0.0784	-0.1112	-0.1291
AGRINDEX	-0.7482*	-1.4697***	-1.0846***	-1.5153***	-1.8965**
	(0.4211)	(0.5274)	(0.3686)	(0.4483)	(0.8790)
P>|t|	0.076	0.005	0.003	0.001	0.031
Std.Coef	-0.0630	-0.1382	-0.0972	-0.1358	-0.1699
Region FE	No	No	No	No	Yes
Urbanization	No	No	No	Yes	No
Year FE	No	No	Yes	Yes	Yes
Observations	1776	1761	3537	3537	3537
Clusters (ID)	No	No	1929	1929	1929
First Stg F-Stat	696.053	258.499	985.956	788.846	193.251

Note: Column 2016 repeats the baseline results from Tables 1.10. Column 2015 presents the results when using observations from 2015 instead of 2016. Columns Panel 0, 1 and 2 show the results from three different specifications of the panel model. Standard errors in parentheses. *,**,*** denote statistical significance at the 10%, 5% and 1% level.

Table 1.32: Panel Analysis of Spring PM10 on Depressive Symptoms (10th Grade)

	2016	2015	Panel 0	Panel 1	Panel 2
Unproductive	-0.0106	-0.0258*	-0.0153*	-0.0192*	-0.0175
	(0.0081)	(0.0152)	(0.0083)	(0.0104)	(0.0217)
P>|t|	0.191	0.090	0.064	0.067	0.420
Std.Coef	-0.0874	-0.1402	-0.1184	-0.1483	-0.1355
Depressed	-0.0110	0.0013	-0.0063	-0.0096	-0.0214
	(0.0079)	(0.0153)	(0.0082)	(0.0104)	(0.0200)
P>|t|	0.164	0.931	0.439	0.354	0.284
Std.Coef	-0.0925	0.0071	-0.0492	-0.0747	-0.1663
Anxious	-0.0120	-0.0135	-0.0132	-0.0113	-0.0190
	(0.0102)	(0.0188)	(0.0104)	(0.0132)	(0.0269)
P>|t|	0.242	0.474	0.205	0.394	0.479
Std.Coef	-0.0775	-0.0580	-0.0807	-0.0689	-0.1162
Suicide	-0.0141**	-0.0298**	-0.0198***	-0.0250***	-0.0404**
	(0.0068)	(0.0131)	(0.0067)	(0.0085)	(0.0173)
P>|t|	0.037	0.023	0.003	0.003	0.019
Std.Coef	-0.1390	-0.1897	-0.1812	-0.2294	-0.3700
Cry	-0.0119	-0.0099	-0.0107	-0.0172	-0.0371
	(0.0086)	(0.0165)	(0.0090)	(0.0113)	(0.0233)
P>|t|	0.168	0.550	0.238	0.130	0.112
Std.Coef	-0.0874	-0.0470	-0.0729	-0.1176	-0.2543
Remorse	-0.0246***	-0.0050	-0.0167*	-0.0210*	-0.0097
	(0.0091)	(0.0170)	(0.0093)	(0.0116)	(0.0239)
P>|t|	0.007	0.770	0.072	0.071	0.685
Std.Coef	-0.1807	-0.0240	-0.1146	-0.1440	-0.0669
Lonely	-0.0153*	0.0043	-0.0082	-0.0109	-0.0146
	(0.0092)	(0.0170)	(0.0093)	(0.0117)	(0.0244)
P>|t|	0.097	0.803	0.378	0.348	0.549
Std.Coef	-0.1109	0.0205	-0.0558	-0.0749	-0.1000
Unmotivated	-0.0159**	-0.0135	-0.0154**	-0.0165*	-0.0279
	(0.0074)	(0.0140)	(0.0074)	(0.0094)	(0.0198)
P>|t|	0.030	0.334	0.037	0.079	0.159
Std.Coef	-0.1430	-0.0794	-0.1293	-0.1388	-0.2346
Pessimistic	-0.0249***	-0.0285*	-0.0242***	-0.0223**	-0.0234
	(0.0091)	(0.0161)	(0.0088)	(0.0113)	(0.0230)
P>|t|	0.006	0.077	0.006	0.047	0.310
Std.Coef	-0.1819	-0.1460	-0.1711	-0.1581	-0.1655
Tough	-0.0213**	-0.0264*	-0.0218***	-0.0250**	-0.0097
	(0.0085)	(0.0152)	(0.0084)	(0.0108)	(0.0214)
P>|t|	0.012	0.082	0.009	0.020	0.649
Std.Coef	-0.1686	-0.1436	-0.1659	-0.1896	-0.0740
Insomnia	-0.0175**	-0.0335*	-0.0224**	-0.0264**	-0.0521**
	(0.0088)	(0.0174)	(0.0089)	(0.0114)	(0.0230)
P>|t|	0.045	0.053	0.012	0.020	0.023
Std.Coef	-0.1340	-0.1604	-0.1560	-0.1840	-0.3636
DEPINDEX	-0.0137*	-0.0125	-0.0125	-0.0158	-0.0210
	(0.0075)	(0.0141)	(0.0079)	(0.0100)	(0.0196)
P>|t|	0.068	0.377	0.114	0.113	0.284
Std.Coef	-0.1208	-0.0722	-0.1030	-0.1310	-0.1734
Region FE	No	No	No	No	Yes
Urbanization	No	No	No	Yes	No
Year FE	No	No	Yes	Yes	Yes
Observations	1655	1,711	3,366	3,366	3,366
Clusters (ID)	No	No	1,838	1,838	1,838
First Stg F-Stat	140.928	81.1844	343.651	214.383	62.5228

Note: Column 2016 repeats the baseline results from Table 1.9. Column 2015 presents the results when using observations from 2015 instead of 2016. Columns Panel 0, 1 and 2 show the results from three different specifications of the panel model. Standard errors in parentheses. *,**,*** denote statistical significance at the 10%, 5% and 1% level.

Table 1.33: Panel Analysis of Spring CO on Depressive Symptoms (10^{th} Grade)

	2016	2015	Panel 0	Panel 1	Panel 2
Unproductive	-0.6147	-0.8204	-0.7072*	-0.8083	-1.3391
	(0.4594)	(0.5507)	(0.4085)	(0.5357)	(0.9438)
P>|t|	0.181	0.136	0.083	0.131	0.156
Std.Coef	-0.0501	-0.0745	-0.0614	-0.0701	-0.1162
Depressed	-0.6260	0.1922	-0.2044	-0.2373	-0.5746
	(0.4471)	(0.5548)	(0.3955)	(0.5140)	(0.9416)
P>|t|	0.161	0.729	0.605	0.644	0.542
Std.Coef	-0.0523	0.0172	-0.0178	-0.0207	-0.0501
Anxious	-0.4825	-0.5440	-0.5316	-0.5929	-0.6377
	(0.5798)	(0.6849)	(0.5040)	(0.6590)	(1.1853)
P>|t|	0.405	0.427	0.291	0.368	0.591
Std.Coef	-0.0310	-0.0392	-0.0364	-0.0406	-0.0437
Suicide	-0.5025	-0.9693**	-0.7395**	-0.9155**	-1.4395*
	(0.3827)	(0.4686)	(0.3256)	(0.4294)	(0.7904)
P>|t|	0.189	0.039	0.023	0.033	0.069
Std.Coef	-0.0490	-0.1032	-0.0760	-0.0940	-0.1479
Cry	-0.5173	0.0010	-0.2383	-0.4959	-1.2115
	(0.4848)	(0.5985)	(0.4441)	(0.5717)	(1.0457)
P>|t|	0.286	0.999	0.592	0.386	0.247
Std.Coef	-0.0377	0.0001	-0.0183	-0.0381	-0.0930
Remorse	-1.4531***	0.0100	-0.7143	-1.0036*	-1.5372
	(0.5124)	(0.6191)	(0.4566)	(0.5941)	(1.0619)
P>|t|	0.005	0.987	0.118	0.091	0.148
Std.Coef	-0.1060	0.0008	-0.0551	-0.0774	-0.1185
Lonely	-0.8928*	0.2897	-0.3015	-0.4627	-0.5018
	(0.5200)	(0.6190)	(0.4630)	(0.6031)	(1.0807)
P>|t|	0.086	0.640	0.515	0.443	0.642
Std.Coef	-0.0640	0.0234	-0.0231	-0.0355	-0.0385
Unmotivated	-0.7701*	-0.5036	-0.6705*	-0.7412	-0.5567
	(0.4170)	(0.5089)	(0.3709)	(0.4857)	(0.8610)
P>|t|	0.065	0.322	0.071	0.127	0.518
Std.Coef	-0.0686	-0.0494	-0.0632	-0.0699	-0.0525
Pessimistic	-1.3195***	-1.0870*	-1.1543***	-1.0791*	-1.5329
	(0.5146)	(0.5812)	(0.4320)	(0.5629)	(1.0093)
P>|t|	0.010	0.061	0.008	0.055	0.129
Std.Coef	-0.0957	-0.0930	-0.0916	-0.0856	-0.1216
Tough	-1.3760***	-0.9664*	-1.1538***	-1.3033**	-1.4006
	(0.4754)	(0.5507)	(0.4151)	(0.5493)	(0.9684)
P>|t|	0.004	0.079	0.005	0.018	0.148
Std.Coef	-0.1082	-0.0878	-0.0982	-0.1109	-0.1192
Insomnia	-1.0805**	-0.8155	-0.9425**	-1.2216**	-3.0691***
	(0.4951)	(0.6247)	(0.4310)	(0.5584)	(1.0072)
P>|t|	0.029	0.192	0.029	0.029	0.002
Std.Coef	-0.0819	-0.0652	-0.0737	-0.0955	-0.2399
DEPINDEX	-0.8197*	-0.2391	-0.5117	-0.6641	-0.9509
	(0.4224)	(0.5140)	(0.3828)	(0.4989)	(0.8937)
P>|t|	0.052	0.642	0.181	0.183	0.287
Std.Coef	-0.0718	-0.0231	-0.0474	-0.0616	-0.0882
Region FE	No	No	No	No	Yes
Urbanization	No	No	No	Yes	No
Year FE	No	No	Yes	Yes	Yes
Observations	1655	1,711	3,366	3,366	3,366
Clusters (ID)	No	No	1,838	1,838	1,838
First Stg F-Stat	645.227	259.349	1023.89	729.645	165.428

Note: Column 2016 repeats the baseline results from Table 1.9. Column 2015 presents the results when using observations from 2015 instead of 2016. Columns Panel 0, 1 and 2 show the results from three different specifications of the panel model. Standard errors in parentheses. *,**,*** denote statistical significance at the 10%, 5% and 1% level.

Table 1.34: Panel Analysis of Spring PM10 on Aggressive Symptoms (10th Grade)

	2016	2015	Panel 0	Panel 1	Panel 2
Irritable	-0.0086	-0.0341**	-0.0166*	-0.0179*	0.0397
	(0.0083)	(0.0163)	(0.0086)	(0.0108)	(0.0235)
P>|t|	0.301	0.037	0.053	0.098	0.092
Std.Coef	-0.0688	-0.1716	-0.1216	-0.1313	0.2903
Disturb	-0.0073	-0.0108	-0.0083	-0.0142	-0.0020
	(0.0080)	(0.0155)	(0.0081)	(0.0104)	(0.0213)
P>|t|	0.363	0.486	0.305	0.171	0.925
Std.Coef	-0.0605	-0.0568	-0.0636	-0.1079	-0.0153
Aggress	-0.0129	-0.0355**	-0.0190**	-0.0241**	0.0197
	(0.0084)	(0.0161)	(0.0084)	(0.0106)	(0.0232)
P>|t|	0.125	0.027	0.025	0.023	0.394
Std.Coef	-0.1026	-0.1824	-0.1404	-0.1781	0.1461
Fight	-0.0166**	-0.0150	-0.0140*	-0.0167*	-0.0105
	(0.0079)	(0.0150)	(0.0079)	(0.0101)	(0.0216)
P>|t|	0.036	0.318	0.078	0.099	0.628
Std.Coef	-0.1405	-0.0818	-0.1099	-0.1312	-0.0824
Angry	-0.0229***	-0.0188	-0.0202**	-0.0224**	-0.0056
	(0.0085)	(0.0159)	(0.0080)	(0.0101)	(0.0218)
P>|t|	0.007	0.239	0.011	0.027	0.797
Std.Coef	-0.1796	-0.0963	-0.1483	-0.1642	-0.0412
AGRINDEX	-0.0165**	-0.0128	-0.0140*	-0.0153	0.0259
	(0.0077)	(0.0146)	(0.0077)	(0.0098)	(0.0212)
P>|t|	0.033	0.381	0.069	0.118	0.222
Std.Coef	-0.1421	-0.0714	-0.1118	-0.1228	0.2068
Region FE	No	No	No	No	Yes
Urbanization	No	No	No	Yes	No
Year FE	No	No	Yes	Yes	Yes
Observations	1655	1,711	3,366	3,366	3,366
Clusters (ID)	No	No	1,838	1,838	1,838
First Stage F-Stat	140.928	81.1844	343.651	214.383	62.5228

Note: Column 2016 repeats the baseline results from Table 1.10. Column 2015 presents the results when using observations from 2015 instead of 2016. Columns Panel 0, 1 and 2 show the results from three different specifications of the panel model. Standard errors in parentheses. *,**,*** denote statistical significance at the 10%, 5% and 1% level.

Table 1.35: Panel Analysis of Spring CO on Aggressive Symptoms (10th Grade)

	2016	2015	Panel 0	Panel 1	Panel 2
Irritable	-0.8892*	-1.5756***	-1.2201***	-1.5447***	-1.9514*
	(0.4723)	(0.5930)	(0.4327)	(0.5651)	(1.0184)
P>|t|	0.060	0.008	0.005	0.006	0.055
Std.Coef	-0.0703	-0.1326	-0.1000	-0.1266	-0.1600
Disturb	-0.4516	-0.1922	-0.3328	-0.7172	-1.5334
	(0.4543)	(0.5622)	(0.4035)	(0.5278)	(0.9753)
P>|t|	0.320	0.732	0.410	0.174	0.116
Std.Coef	-0.0371	-0.0169	-0.0284	-0.0613	-0.1310
Aggress	-1.0063**	-1.4061**	-1.1667***	-1.6565***	-2.9062***
	(0.4743)	(0.5817)	(0.4250)	(0.5570)	(0.9719)
P>|t|	0.034	0.016	0.006	0.003	0.003
Std.Coef	-0.0795	-0.1208	-0.0968	-0.1375	-0.2412
Fight	-1.0900**	-0.1457	-0.5791	-0.8386	-1.9064**
	(0.4462)	(0.5429)	(0.3908)	(0.5160)	(0.9124)
P>|t|	0.015	0.788	0.138	0.104	0.037
Std.Coef	-0.0918	-0.0133	-0.0510	-0.0739	-0.1680
Angry	-1.6437***	-0.3570	-1.0059**	-1.1332**	-1.9008**
	(0.4805)	(0.5784)	(0.3941)	(0.5089)	(0.9124)
P>|t|	0.001	0.537	0.011	0.026	0.037
Std.Coef	-0.1279	-0.0306	-0.0827	-0.0931	-0.1562
AGRINDEX	-1.3552***	-0.4253	-0.8882**	-1.0600**	-1.9171**
	(0.4368)	(0.5318)	(0.3828)	(0.5002)	(0.8752)
P>|t|	0.002	0.424	0.020	0.034	0.028
Std.Coef	-0.1160	-0.0396	-0.0796	-0.0950	-0.1719
Region FE	No	No	No	No	Yes
Urbanization	No	No	No	Yes	No
Year FE	No	No	Yes	Yes	Yes
Observations	1655	1,711	3,366	3,366	3,366
Clusters (ID)	No	No	1,838	1,838	1,838
First Stage F-Stat	645.227	259.349	1023.89	729.645	165.428

Note: Column 2016 repeats the baseline results from Table 1.10. Column 2015 presents the results when using observations from 2015 instead of 2016. Columns Panel 0, 1 and 2 show the results from three different specifications of the panel model. Standard errors in parentheses. *,**,*** denote statistical significance at the 10%, 5% and 1% level.

Table 1.36: Anderson (2008) Sharpened False Discovery Rate Q-Values on PM10

Variables	2016	2015	Panel 0	Panel 1	Panel 2
Unproductive	(0.001)***	(0.009)***	(0.001)***	(0.001)***	(0.056)*
	[0.134]	[0.246]	[0.077]*	[0.107]	[1.000]
Depressed	(0.028)**	(0.001)***	(0.001)***	(0.009)***	(0.055)*
	[0.127]	[0.916]	[0.160]	[0.169]	[1.000]
Anxious	(0.001)***	(0.007)***	(0.001)***	(0.009)***	(0.056)*
	[0.161]	[0.574]	[0.128]	[0.169]	[1.000]
Suicide	(0.028)**	(0.020)**	(0.003)***	(0.018)**	(0.132)
	[0.068]*	[0.246]	[0.040]**	[0.051]*	[0.226]
Cry	(0.040)**	(0.010)***	(0.003)***	(0.009)***	(0.009)***
	[0.127]	[0.615]	[0.128]	[0.136]	[0.645]
Remorse	(0.040)**	(0.015)**	(0.003)***	(0.017)**	(0.009)***
	[0.039]**	[0.916]	[0.077]*	[0.107]	[1.000]
Lonely	(0.038)**	(0.015)**	(0.003)***	(0.030)**	(0.055)*
	[0.099]*	[0.916]	[0.145]	[0.169]	[1.000]
Unmotivated	(0.004)***	(0.009)***	(0.001)***	(0.011)**	(0.081)*
	[0.068]*	[0.430]	[0.062]*	[0.107]	[0.803]
Pessimistic	(0.008)***	(0.006)***	(0.001)***	(0.009)***	(0.047)**
	[0.039]**	[0.246]	[0.040]**	[0.095]*	[1.000]
Tough	(0.004)***	(0.013)**	(0.001)***	(0.011)**	(0.041)**
	[0.041]**	[0.246]	[0.040]**	[0.089]*	[1.000]
Insomnia	(0.021)**	(0.026)**	(0.003)***	(0.017)**	(0.058)*
	[0.073]*	[0.246]	[0.040]**	[0.089]*	[0.226]
Irritable	(0.029)**	(0.017)**	(0.004)***	(0.017)**	(0.067)*
	[0.192]	[0.246]	[0.072]**	[0.110]	[0.645]
Disturb	(0.027)**	(0.012)**	(0.003)***	(0.011)**	(0.074)*
	[0.209]	[0.574]	[0.128]	[0.152]	[1.000]
Aggress	(0.027)**	(0.008)***	(0.001)***	(0.009)***	(0.041)**
	[0.112]	[0.246]	[0.049]**	[0.089]*	[1.000]
Fight	(0.043)**	(0.006)***	(0.002)***	(0.008)***	(0.056)*
	[0.068]*	[0.430]	[0.077]*	[0.110]	[1.000]
Angry	(0.038)**	(0.012)**	(0.003)***	(0.009)***	(0.069)*
	[0.039]**	[0.368]	[0.040]**	[0.089]*	[1.000]

Note: Based on the regression results reported in Tables 1.28, 1.30, 1.32, and 1.34. Sharpened q-values for regressions related to 7th graders in parentheses. Sharpened q-values for regressions related to 10th graders in [brackets]. *,**,*** denote statistical significance at the 10%, 5% and 1% level.

Table 1.37: Anderson (2008) Sharpened False Discovery Rate Q-Values on CO

Variables	2016	2015	Panel 0	Panel 1	Panel 2
Unproductive	(0.001)***	(0.013)**	(0.001)***	(0.001)***	(0.069)*
	[0.131]	[0.339]	[0.091]*	[0.136]	[0.206]
Depressed	(0.087)*	(0.013)**	(0.004)***	(0.011)**	(0.218)
	[0.131]	[1.000]	[0.252]	[0.252]	[0.294]
Anxious	(0.064)*	(0.013)**	(0.004)***	(0.017)**	(0.276)
	[0.180]	[0.783]	[0.205]	[0.236]	[0.294]
Suicide	(0.060)*	(0.017)**	(0.004)***	(0.012)**	(0.405)
	[0.131]	[0.223]	[0.045]**	[0.084]*	[0.174]
Cry	(0.198)	(0.100)*	(0.030)**	(0.045)**	(0.244)
	[0.160]	[1.000]	[0.252]	[0.236]	[0.260]
Remorse	(0.104)	(0.122)	(0.030)**	(0.065)*	(0.383)
	[0.026]**	[1.000]	[0.119]	[0.129]	[0.206]
Lonely	(0.087)*	(0.025)**	(0.012)**	(0.045)**	(0.353)
	[0.085]*	[1.000]	[0.252]	[0.236]	[0.294]
Unmotivated	(0.008)***	(0.013)**	(0.001)***	(0.006)***	(0.232)
	[0.084]*	[0.674]	[0.087]*	[0.136]	[0.294]
Pessimistic	(0.029)**	(0.013)**	(0.001)***	(0.008)***	(0.218)
	[0.034]**	[0.263]	[0.034]**	[0.097]*	[0.206]
Tough	(0.029)**	(0.035)**	(0.005)***	(0.024)**	(0.353)
	[0.026]**	[0.263]	[0.034]**	[0.084]*	[0.206]
Insomnia	(0.074)*	(0.059)*	(0.013)**	(0.017)**	(0.301)
	[0.057]*	[0.433]	[0.048]**	[0.084]*	[0.025]**
Irritable	(0.060)*	(0.050)**	(0.012)**	(0.011)**	(0.218)
	[0.084]*	[0.147]	[0.034]**	[0.051]*	[0.174]
Disturb	(0.060)*	(0.030)**	(0.007)***	(0.009)***	(0.069)*
	[0.160]	[1.000]	[0.234]**	[0.151]	[0.206]
Aggress	(0.064)*	(0.013)**	(0.004)***	(0.006)***	(0.190)
	[0.057]**	[0.147]	[0.034]**	[0.051]*	[0.025]**
Fight	(0.198)	(0.017)**	(0.013)**	(0.008)***	(0.069)*
	[0.041]**	[1.000]	[0.128]	[0.131]	[0.149]
Angry	(0.112)	(0.019)**	(0.013)**	(0.011)**	(0.218)
	[0.017]**	[0.936]	[0.034]**	[0.084]*	[0.149]

Note: Based on the regression results reported in Tables 1.29, 1.31, 1.33, and 1.35. Sharpened q-values for regressions related to 7th graders in parentheses. Sharpened q-values for regressions related to 10th graders in [brackets]. *,**,*** denote statistical significance at the 10%, 5% and 1% level.

Table 1.38: Panel Analysis Using Index Variables (7^{th} Grade)

	Panel 0	**Panel 1**	**Panel 2**	**Panel 3**	**Panel 4**
			Pollutant: PM10		
DEPINDEX	-0.0305***	-0.0281***	-0.0445***	-0.0305***	-0.0305***
	(0.0068)	(0.0081)	(0.0162)	(0.0089)	(0.0116)
P>|t|	0.000	0.001	0.006	0.001	0.009
Std.Coef	-0.2628	-0.2414	-0.3826	-0.2628	-0.2628
AGRINDEX	-0.0242***	-0.0261***	-0.0331*	-0.0242***	-0.0242*
	(0.0072)	(0.0087)	(0.0173)	(0.0094)	(0.0133)
P>|t|	0.001	0.003	0.056	0.010	0.068
Std.Coef	-0.1970	-0.2127	-0.2696	-0.1970	-0.1970
			Pollutant: CO		
DEPINDEX	-1.3222***	-1.3515***	-1.4694*	-1.3222***	-1.3222***
	(0.3514)	(0.4193)	(0.8248)	(0.4344)	(0.3994)
P>|t|	0.000	0.001	0.075	0.002	0.001
Std.Coef	-0.1252	-0.1280	-0.1391	-0.1252	-0.1252
AGRINDEX	-1.0846***	-1.5153***	-1.8965**	-1.0846**	-1.0846**
	(0.3686)	(0.4483)	(0.8790)	(0.4573)	(0.5514)
P>|t|	0.003	0.001	0.031	0.018	0.049
Std.Coef	-0.0972	-0.1358	-0.1699	-0.0972	-0.0972
Observations	3,537	3,537	3,537	3,537	3,537
Clustering Level	Individual	Individual	Individual	Living District	Year-Location
Clusters	1929	1929	1929	136	34
First Stg F-stat (PM10)	331.44	217.519	104.827	14.7148	4.36166
First Stg F-stat (CO)	985.956	788.846	193.251	41.7635	13.242
Region FE	No	No	Yes	No	No
Urbanization	No	Yes	No	No	No
Year FE	Yes	Yes	Yes	Yes	Yes

Note: Standard errors in parentheses. *,**,*** denote statistical significance at the 10%, 5% and 1% level.

Table 1.39: Panel Analysis Using Index Variables (10^{th} Grade)

Variable	Panel 0	Panel 1	Panel 2	Panel 3	Panel 4
			Pollutant: PM10		
DEPINDEX	-0.0125	-0.0158	-0.0210	-0.0125	-0.0125*
	(0.0079)	(0.0100)	(0.0196)	(0.0081)	(0.0067)
P>|t|	0.114	0.113	0.284	0.126	0.064
Std.Coef	-0.1030	-0.1310	-0.1734	-0.1030	-0.1030
AGRINDEX	-0.0140*	-0.0153	0.0259	-0.0140	-0.0140**
	(0.0077)	(0.0098)	(0.0212)	(0.0087)	(0.0069)
P>|t|	0.069	0.118	0.222	0.109	0.044
Std.Coef	-0.1118	-0.1228	0.2068	-0.1118	-0.1118
			Pollutant: CO		
DEPINEX	-0.5117	-0.6641	-0.9509	-0.5117	-0.5117
	(0.3828)	(0.4989)	(0.8937)	(0.3908)	(0.3238)
P>|t|	0.181	0.183	0.287	0.190	0.114
Std.Coef	-0.0474	-0.0616	-0.0882	-0.0474	-0.0474
AGRINDEX	-0.8882**	-1.0600**	-1.9171**	-0.8882**	-0.8882**
	0.3828	(0.5002)	(0.8752)	(0.4497)	(0.3965)
P>|t|	0.020	0.034	0.028	0.048	0.025
Std.Coef	-0.0796	-0.0950	-0.1719	-0.0796	-0.0796
Observations	3,366	3,366	3,366	3,366	3,366
Clustering Level	Individual	Individual	Individual	Living District	Year-Location
Clusters	1,838	1,838	1,838	131	34
First Stg F-stat (PM10)	343.651	214.383	62.5228	17.8243	4.79215
First Stg F-stat (CO)	1023.89	729.645	165.428	44.1498	13.034
Region FE	No	No	Yes	No	No
Urbanization	No	Yes	No	No	No
Year FE	Yes	Yes	Yes	Yes	Yes

Note: Standard errors in parentheses. *,**,*** denote statistical significance at the 10%, 5% and 1% level.

CHAPTER 2

Gangnam Style and the Housing Market in the Eponymous District: How a Global Pop Culture Phenomenon Boosted Property Prices

2.1 Introduction

After its release in 2012, the song “Gangnam Style” very quickly went viral.[22] The song's popularity made Gangnam, a southern district in Seoul, a household name around the world. Even the slightly mocking undertone of the lyrics satirizing the district's artificial self-image and its hip and trendy residents did not detract from Gangnam’s sudden fame. In fact during the years after the song’s debut the district became an attraction for domestic and international tourists.

Expectations that “Gangnam Style” would have a major impact on tourism in Korea and the Korean economy in general were expressed very early on, and even international observers quickly recognized the song’s potential economic benefits. Already on November 16, 2012, Voice of America was reporting about the potential the song had for boosting the South Korean tourism industry. According to this reporting, local businesses and merchants in Gangnam were benefitting from “an increase in foreign shoppers.” Soon after, on January 24, 2013, CNBC voiced a similar opinion. In an article about the South Korean economy, the news channel described the Korean tourism boom as a “bright spot” in the midst of the economic struggles of 2012, and went on to say that “a large part of the credit for this boom in tourism can go to ‘Gangnam Style’.” The Korea Tourist Organization (KTO) also received early signals about an increased interest in Korea as a travel destination. According to a KTO survey conducted in Los

[22] Korean singer Psy released his song “Gangnam Style” on July 15, 2012. It became the first Korean song to make it onto the Billboard Hot 100 Chart on September 22 that same year. Just two weeks later, it claimed second place on the chart. On November 24, the song's music video became the most viewed on YouTube and reached one billion views on December 21.

Angeles on October 6, 2012, more than 90 percent of the participants had viewed the “Gangnam Style” music video and over 60 percent wished to visit Gangnam in order to experience Korean food and shopping (Korea Tourism Organization, 2012). In short, already in 2012 many observers expected the song to have a positive impact on the number of visitors to Korea in general and Gangnam in particular.

The primary goal of this paper is to use the popularity of the song “Gangnam Style” and the attendant rise in Gangnam’s profile in 2012 as an experiment to assess if and how a sudden increase in a location’s popularity may affect its housing market. In particular, the virality of the hit song may have altered the expectations of participants in the district's residential property market. Apart from these (possibly) altered expectations, and taking into account the sluggishness of real estate markets, it is safe to say that the release of the song did initially not alter any other driving force of Gangnam’s housing market. For the medium to long-run, however, the district gained the prospect of developing into a tourist magnet and therefore coming to have a greater need for more and/or better public infrastructure (e.g. new subway stations or bus lines) as well as tourism-related services (e.g. hotels, restaurants, or cafés), which could eventually drive up the value of land in Gangnam. In anticipation of such intensified land use, it is likely that owners of real estate in Gangnam immediately began asking for higher prices for their properties. In other words I think that the residential property market in Gangnam experienced an immediate contraction of supply when the song went viral.

In the years following the release of “Gangnam Style”, South Korea’s tourism industry indeed picked up speed. The number of international arrivals with a touristic purpose increased from 8.7 million individuals in 2012 to 14.4 million in 2019.[23] Survey data suggest that Seoul’s

[23] All non-resident foreign nationals entering South Korea are required to submit an arrival card where they must state the purpose of their visit. The KTO gathers this information and publishes summary statistics on its website: https://know.tour.go.kr/stat/entryTourStatDis19Re.do.

Gangnam district benefitted disproportionately from this increase. Between 2014 and 2019, the share of international visitors that intended to visit Gangnam increased from about 22.9 to 28.2 percent. Interestingly, Seoul as a whole bucked this trend and experienced a slight decline from 80.4 to 76.4 percent over the same time span.[24]

The regression results support the view that the housing market in Gangnam experienced a contraction of supply when "Gangnam Style" went viral. Using a difference-in-differences (DID) design, I show that residential property prices per square meter in Gangnam increased by 1,679,614 KRW (ca. 1,400 USD[25]) compared to other districts in Seoul in the period since the song was released. The logarithmic specification of the house price model reveals that the growth rate of square meter prices increased by an average of 4.8 percent compared to other districts in Seoul. By using dong[26]-level transaction quantity as the dependent variable, I can identify which side of the market has seen the bigger change. Besides the significant increase in housing prices in Gangnam I also find a significant decrease of about 18 transactions per square kilometer (i.e. more than one sixth of the district's mean) relative to other districts. Thus, I conclude that a contraction of supply dominated the residential property market in Gangnam following the release of the song.

My analysis also reveals that Gangnam experienced a significant increase in the number of available hotel rooms compared to other districts in Seoul. Interestingly, the time paths of the

[24] Since Gangnam was previously not perceived as a tourist destination, no figures were collected for the district before 2014. I refer to visitor numbers up to and including 2019, the last year before the COVID-19 pandemic. Data is publicly available at Data Seoul: https://data.seoul.go.kr/dataList/10944/S/2/datasetView.do. (last accessed on March 4th, 2023).
[25] Throughout this paper I convert 1,200 KRW into one US dollar.
[26] A 'dong' (often translated as 'neighborhood') constitutes a sub-division of a district in Seoul. Confusingly, two different types of dong are recognized, so-called legal-status dongs and administrative dongs. For the purposes legal-status dongs are of primary interest because they are used for land registries and also for the traditional address system. Administrative dongs, in contrast, are areas which are under the purview of local governments, within whose parameters they coordinate administrative tasks, provide services to the public and organize elections. The metropolitan city of Seoul has a total of 467 legal-status dongs, while the district of Gangnam has 14 legal-status dongs (see also Online Appendix Figure A1).

effects of “Gangnam Style” on the number of available hotel rooms and on residential property prices look similar, with the latter taking place slightly in advance of the former. This observation supports the view that an anticipated increase in visitor numbers led owners of residential properties in Gangnam to ask for higher prices. I therefore argue that rapidly adjusting expectations are most likely the dominant driving force behind the supply contraction found in the regressions.

This paper contribute to the housing market literature in three distinct ways. First, my analysis benefits from a clear and simple econometric design. The treatment is precise in terms of time and regional dimensions; thus, I do not need to worry about issues surrounding the timing and localization of the exogenous variation. Second, this study explains a possible link between cultural assets, tourism, and residential property prices. Instead of physical (or tangible) cultural amenities and infrastructure, however, I focus on fame – an intangible cultural asset – as a means of stimulating tourism and as a consequence the housing market. Third, this research generates initial insights into the notoriously under-researched housing markets of megacities. Real estate markets in such places have various traits and exhibit different dynamics compared to typical cities, towns or rural areas and therefore constitute a separate area of research.

2.2 Related Literature

Hedonic house pricing is the workhorse model in empirical studies of real estate markets, and dates back decades (see Ball, 1973, for an overview of the early literature). In essence, the model views a residential property as the sum of individual characteristics that cannot be sold separately (Montero & Fernández-Avilés, 2014). The objective of empirical investigations is to estimate the contribution of each characteristic to the price of a residential property.

The relevant literature is so extensive that I can only selectively mention a few contributions. I start with articles that have unraveled the relationship between various building characteristics and residential property prices, particularly dwelling size and lot size (for example Murphy, 2018, but also many other studies cited below); types of building (Richardson et al., 1974), building age (e.g. Coulson et al., 2019; Goodman & Thibodeau, 2003), number of (bed-)rooms, bathrooms and kitchens (e.g. Cebula, 2009; Ebru & Eban, 2011; Kain & Quigley, 1970; Zietz et al., 2008), to name a few.

Extending the hedonic model beyond building characteristics, previous research has also identified a number of housing price determinants related to neighborhood attributes. Examples include urban design characteristics such as imageability (i.e. the memorability and recognizability of a place, Hamidi et al., 2020), cultural amenities (such as theaters, libraries, museums, etc., Borgoni et al, 2018), street connectivity (Bresson & Hsiao, 2011; Shen & Karimi, 2017), density of the surrounding area (Fesselmeyer & Seah, 2018), availability of parking places (Bakis et al., 2019; Park et al., 2017), proximity to schools and public transportation (Lee & Choi, 2016; Lieske et al., 2021), quality of nearby schools (Black, 1999; Downes & Zabel, 2002), air pollution (e.g. Chay & Greenstone, 2005; Harrison & Rubinfeld, 1978; Ridker & Henning, 1967; Smith & Huang, 1993), quality and attributes of the local environment (Currie et al., 2015; Greenstone & Gallagher, 2008; Luttik, 2000), level and perception of crime (Gibbons, 2004; Linden & Rockoff, 2008; Lynch & Rasmussen, 2001; Pope, 2008), existence of retail marijuana stores in the neighborhood (Cheng et al., 2018; Conklin et al., 2020), distance to nuclear power plants (e.g. Clark et al, 1997; Bauer et al., 2017), and risk level with respect to a rise in sea levels (Beck & Lin, 2020; Bernstein et al., 2019).

The hedonic model does however have an innate weakness due to potential “garden variety” (Bartik, 1987) identification issues when unobserved attributes correlate with the main attribute of interest (Banzhaf, 2019; Greenstone, 2017). For this reason Black (1999) uses the hedonic model and takes a regression discontinuity approach to measure parents’ valuation of school quality. Greenstone & Gallagher (2008) also employ a regression discontinuity strategy and compare house prices in areas surrounding hazardous waste sites chosen for a public cleanup program to those in areas that narrowly missed qualification for these cleanups. Other empirical applications of the hedonic model employ even more complex econometric methods related to heteroskedasticity and spatial auto-correlation (Fletcher et al., 2000; Forrest, 1991; Goodman & Thibodeau, 1995; Kim et al., 2003; Osland, 2020). Zietz et al. (2008) depart further from the typical use of the model and apply a quantile regression approach. They show that lot size, number of bathrooms, and floor type (e.g. hardwood) have greater effects on selling prices than the presence of garages, exterior sidings, sprinkler systems, and distance to the city center (Zietz et al., 2008)

Another (possible) weakness in this context is that those who employ the hedonic model often simply ignore the macroeconomic determinants of residential property prices like income fluctuations due to the business cycle, changes in taxation rules, or interest rate movements. However, a few papers have evaluated the effect of macroeconomic variables on the volatility of residential property prices. Using vector autoregressive models, Hossain & Latif (2009) investigate the effect of GDP growth rates, Hendershott et al. (2021) examine the impact of changes in taxation and Borowiecki (2009) studies the effect of interest rate changes.

A number of recent papers have employed event study designs (or DID approaches) to identify potential determinants of residential property prices that were previously overlooked. Currie et al. (2015) examine the effect of environmental health risks (on the basis of the openings and closings of 1,600 industrial plants) on property values. They show that toxic plant openings (within 0.5 miles) lead to an 11 percent decrease in property values (Currie et al., 2015). Other studies analyze the effects of a flood event (Skantz & Strickland, 1987), bank lending policies (Jung & Lee, 2017; Park, 2019), property condition disclosure laws (Nanda & Ross, 2012), foreclosure policies (Campbell et al., 2011), and the introduction of recreational marijuana sales by medical marijuana stores (Conklin et al., 2020). The event study (or DID) design is an efficient tool that allows researchers to overcome possible endogeneity issues related to links between residential property prices and variables of interest.

This paper contributes in three ways to the ongoing debate on the characteristics of residential property markets in general and the determinants of residential property prices in particular. First, my analysis benefits from a very clear and simple econometric design. Regarding the time dimension, there is a sharp distinction between the periods before and after the release of "Gangnam Style", which is important when using the DID design. In contrast to studies that use the implementation or change of a housing market policy (see, for example, Campbell et al., 2011; Jung & Lee, 2017; Park, 2019), I do not have to worry about any anticipatory effects. Similarly, I am also not concerned about slow-acting treatment effects as can come about in relation to slowly emerging environmental problems.[27] As I have already mentioned, the song "Gangnam Style" went viral very quickly after its release. Regarding the regional dimension, there is a precisely defined district that is treated, namely Gangnam. Thus, complicated calculations of the average effects in the cases of various regions being treated with heterogeneous timing are not necessary in this setting. The basic research design, therefore, simplifies the interpretation of the estimated regression coefficients.

Second, this paper contribute to the new and still relatively small body of literature that analyzes culture and tourism as potential factors influencing residential property prices (and residential property markets). Borgoni et al. (2018) use the hedonic house price model to show that cultural amenities (cultural facilities and infrastructure) may make neighborhoods more vibrant and enjoyable, which generally increases property prices. Most of the economic literature

[27] Interestingly, Currie et al. (2015) avoid this issue by using information about toxic plant openings and closings to estimate the effect of environmental health risks on housing values.

on property markets, however, does not analyze the effects of culture and tourism on prices. This study attempts to fill this gap by analyzing the link between pop culture, tourism, and the residential property market, finding that sudden fame attracts tourists to a district and thereby creates opportunities for business owners to meet touristic demands (by providing services such as hotel beds and restaurant meals), which directly affects the local value of land.

Third, this paper examine the housing market of a megacity. The city of Seoul has a population of about 9.4 million people; Gangnam – one of the city's 25 districts – has about half a million residents. Housing markets in megacities may have different characteristics and may also undergo different dynamics compared to typical cities, towns or rural areas. Interestingly, there is a lack of in-depth investigation of housing markets in megacities, probably due to a dearth of data. In this study, I use a dataset that contains information on all real estate purchases registered to the Seoul Metropolitan Government between 2009 and 2022 in order to address this gap.

2.3 Data and Methodology

Under Korean law, any transfer of the ownership of land or a part of it must be registered in the local land registry. In the empirical analysis, I use the registrations of house purchases completed by the Seoul Metropolitan Government between 2009 and 2022 as the main data source. The dataset contains not only price information for ALL residential property purchases conducted in Seoul during this time span, as the registry also provides very detailed information on property and building characteristics such as year of construction, housing type (apartment, multi-household house, regular house, and dual-purpose buildings for commercial and residential use), floor location, size in square meters (m^2), as well as district and legal-status dong location.

My analysis aims to identify the impact that the release of the song "Gangnam Style" had on the housing market and on the provision of tourism services in Gangnam district. Table 2.1 describes the research design.

Table 2.1: Summary of Research Design

Assignment	Districts	Timing
Treated District	Gangnam	Treated: 2012 – 2022 Controlled: 2009 – 2011
Control Districts 1	Seoul Metropolitan City (excluding Gangnam)	Controlled: 2009 – 2022
Control Districts 2	Seoul to the South of the Han River (excluding Gangnam)	Controlled: 2009 – 2022
Control Districts 3	Neighboring Districts (Seocho and Songpa)	Controlled: 2009 – 2022

Note: Please refer to Appendix Figure 2.4 for the location of districts in Seoul.

The first goal is to estimate the effect that the release of the song had on property prices. To get an outcome measure that is independent of the size of a residential property, I have divided every observed transaction price by the relevant residential property size to obtain the square meter price. I have also used the natural logarithm of the square meter price to estimate the song's effect on the growth rate of residential property prices. Nearby schools and means of public transportation have the potential to affect residential property prices (Lee & Choi, 2016; Lieske et al., 2021). Therefore, I have manually collected additional information about the number of local elementary, middle, and high schools as well as about the availability of nearby subway stations and lines for every legal-status dong and every observation year. The collected information has then been used to complement the main data source from the Seoul Metropolitan Government registry. Equation (2.1) summarizes the empirical strategy I employed to identify the effect of "Gangnam Style" on residential property prices using the DID design illustrated in Table 2.1:

$$y_{idst} = \beta_0 + \beta_1 DID_{it} + \theta^T BldChar_i + \alpha^T NbhAttr_{st} + \gamma_t + \delta_d + (\mu_s) + \varepsilon_{idst} \tag{2.1}$$

The outcome variable y_{idst} refers to the square meter price (or the logarithm thereof) observed for purchase i of a residential property located in district d (and legal-status dong s) in year t. The indicator variable DID_{it} equals one only when transaction i was conducted in the district of Gangnam in 2012 and subsequently – the year of the release of “Gangnam Style”. $BldChar_i$ denotes a vector consisting of detailed information about purchase i: the specific housing type, its floor location, and the age of the building. $NbhAttr_{st}$ is a vector containing neighborhood attributes like the number of elementary, middle, and high schools as well as information about the public transport connectivity of the residential property’s legal-status dong s at time t. I also control for a number of fixed effects – namely year (γ_t), district (δ_d), and in some specifications also legal-status dong fixed effects (μ_s). Since the treatment occurred on the district level, I cluster the standard errors ε_{idst} on the year-district level.[28] To ensure the robustness of the results, I run four different specifications of equation (2.1). In Models 1 and 2, I simply count the number of subway stations and lines in a legal-status dong at time t to measure public transport connectivity. In Models 3 and 4, I use a set of 17 dummy variables, each indicating whether a legal-status dong is connected to one of Seoul’s 17 subway lines at the time a purchase is observed. Legal-status dong fixed effects are excluded in Models 1 and 3 whereas in Models 2 and 4 they are included to control for heterogeneity on this subdivision level of a district.

The dataset contains information on 2,129,882 residential property purchases conducted in Seoul over a time span of 14 years (i.e. from 2009 to 2022). Table 2.2 summarizes the descriptive statistics for the main variables in the housing price regression analysis.

[28] Year-district level clustering assumes regional (district) correlation within a specific year. For example residential property prices in Gaepo-dong (which is located in Gangnam) in 2010 are assumed to be correlated with residential property prices in Samseong-dong (also in Gangnam) in 2010, but not in any other year.

Table 2.2: Descriptive Statistics for Housing Price Regression Analysis

Variables	Obs	Mean	Definition
Square Meter Price*	2,129,882	688.6707 (441.4354)	Housing Price (in 10,000 KRW) / Size of Residential Property (in m^2)
Housing Type	2,129,882	1.7637 (0.9587)	Transactions by Housing Type Apartments: 1,083,001 Multi-Household Houses: 673,906 Regular Houses: 166,190 Dual Purpose Buildings: 206,785
Age of Building	2,069,845	16.1072 (11.8165)	Transaction Year – Construction Year
Floor	2,129,882	6.3756 (5.7973)	Floor on which residential property is located
Elementary Schools	2,129,882	4.6539 (3.6487)	Number of elementary schools in residential property's legal-status dong in year of transaction
Middle Schools	2,129,882	2.8187 (2.3192)	Number of middle schools in residential property's legal-status dong in year of transaction
High Schools	2,129,882	2.2703 (2.3114)	Number of high schools in residential property's legal-status dong in year of transaction
Subways†	2,129,882	2.0570 (1.8373)	Number of subway stations & lines in residential property's legal-status dong in year of transaction

Notes: Standard deviations in parentheses. *Annual averages of square meter prices can be found in Appendix Table 2.12. †*Subways* is a count variable that reflects the number of subway stations and subway lines in a legal-status dong in a given year. Occasionally this number is zero or one, but often a legal-status dong is home to more than one subway station (e.g. Godeok-dong in Gangdong District hosts Line-5-stations Godeok and Sangil, so *Subways* equals two in this case). Similarly, it is possible for a legal-status dong to host a single station where people can transfer between several subway lines. In this case, *Subways* reflects the number of subway lines serving the station (e.g. Oksu-dong in Seongdong District hosts Oksu Station on Line 3 and Gyeongui-Jungang Line, so that *Subways* in this case also equals two). Moreover, there are cases in which a legal-status dong hosts two or more subway stations and at least one of these stations is on more than one line (e.g. Gaehwa-dong in Gangseo District hosts Gaehwa Station on Line 9 but also Gimpo Airport Station on Line 5, Line 9, the Gimpo-Gold Line, and the Airport Line. Here *Subways* equals five). The maximum value of ten is observed for Nonhyeon-dong in Gangnam (hosting Nonhyeon Station on Line 7 and the New Bundang Line, Eonju Station on Line 9, Hakdong Station on Line 7, Sinnonhyeon on Line 9 and the New Bundang Line, Sinsa Station on Line 3 and the New Bundang Line, and Gangnam District Office Station on Line 7 and the Bundang Line).

As can be seen in Table 2.2, apartments were the subject of more than half of all transactions. This is not surprising as they are the housing type that most people reside in, and wish to reside in (Hong & Lim, 2018). Also, I can confirm the general notion that residential property prices in Seoul are very high. Take for example the most popular apartment size of about $84m^2$. According to the table, the average price of such an apartment was 578,483,388 KRW (0.482 million USD). Considering the standard deviation of prices per square meter, an apartment can easily cost more than one million USD. Moreover, the data also shows that residential property prices increased by about 78 percent between 2009 and 2022 (see Appendix Table 2.12). The average building age at the time of a purchase was a little over 16 years, which may be due to the fact that buildings older than 30 years are usually demolished, to be replaced by a new-build. The average purchased residential property is located higher than the sixth floor, a reflection of Seoul's relatively high population density. Finally, the average numbers of schools and subways indicate short distances to the nearest schools (particularly to elementary schools) and also a good level of connectivity.

The second goal of the empirical analysis is to estimate the effect of "Gangnam Style" on the number of transactions in the residential property housing market. This part of the analysis is crucial in determining whether supply or demand side forces have driven the market. Given an increase in residential property prices, a reduction in transactions would indicate that a contraction of supply was the dominant factor. Likewise, an increase in transactions would point to an expansion of demand as the dominant force. Given a decrease in residential property prices, the reverse patterns would hold true.

To address the level of transactions econometrically, I change the perspective slightly and consider the number of transactions per legal-status dong. To this end, I have created a three-dimensional panel (Year×Dong×HousingType) as follows: for every year and every legal-status dong I have counted the number of transactions separately for each of the four housing types. To correct for a potential legal-status-dong-size effect on the outcome variable, I have divided the numbers of transactions per year and dong by the size of the dong measured in square kilometers (km^2).[29] Maintaining the DID design, equation (2.2) reflects the slightly different econometric strategy employed to identify the effect of "Gangnam Style" on the number of transactions in the housing market:

$$y_{hdst} = \beta_0 + \beta_1 DID_{st} + \theta^T BldChar_{hst} + \alpha^T NbhAttr_{st} + \gamma_t + \delta_d + \mu_s + \varepsilon_{dst} \tag{2.2}$$

The outcome variable y_{hdst} refers to housing type h specific numbers of transactions per square kilometer in district d and legal-status dong s in year t. The indicator variable DID_{st} equals one only when the outcome variable refers to a legal-status dong located in the district of Gangnam and counts transactions in 2012 and subsequently. $BldChar_{hst}$ is a vector containing the averages of the floor locations and ages of the purchased buildings by housing type, legal-status dong and year. $NbhAttr_{st}$ is unchanged from the price regressions (see equation (2.1) and the following discussion). Also, I control for year fixed effects (γ_t), district fixed effects (δ_d), and legal-status dong fixed effects (μ_s). Standard errors ε_{dst} are clustered on the year-district level, which allows contemporaneous regional (district) correlation. Similar to the price regressions, I either use the count variable *Subways* or the set of 17 subway line dummies (Models 5 and 7 versus Models 6 and 8). Regarding the outcome variable, I use transactions for all housing types and for apartments only (Models 5 and 6 versus Models 7 and 8; see also footnote 8). Thus, I consider four different specifications of equation (2.2). Table 2.3 shows the descriptive statistics for variables used in the transaction regression analysis.

---

[29] Using transactions per km^2 as the dependent variable to measure the housing market response in the quantity dimension is justified for the following three reasons. First, residential properties (particularly apartments) purchased in Seoul are relatively homogeneous in terms of their square footage (see Appendix Table 2.13). Second, the number of apartments per km^2 is comparable throughout Seoul. Using three stock-takings from the years 2010, 2015 and 2020, I have calculated the number of apartments per km^2 in 422 out of 426 administrative dongs. As a result, I obtain similar means with relatively small standard deviations for *Seoul Metropolitan City*, *Seoul to the South of the Han River*, as well as *Gangnam, Seocho, and Songpa* (see Appendix Table 2.14). Third, the growth rates of the numbers of apartments in various districts are statistically not different from zero with relatively high confidence levels, particularly for *Seoul to the South of the Han River* and for the neighboring districts *Gangnam, Seocho, and Songpa* (see Appendix Table 2.15). Thus any remaining differences in apartment density can be captured by dong fixed effects.

Table 2.3: Descriptive Statistics for Transaction Regression Analysis

Variables	Obs	Mean	Definition
Transactions	16,456	105.4969 (211.5426)	Number of transactions per km^2 by Year, Dong, and Housing type Apartments: 183.8779 (284.7648) Multi-Household Houses: 98.8985 (131.0371) Regular Houses: 37.7546 (49.1405) Dual Purpose Buildings: 109.6636 (283.7505)
Avg. Age of Buildings	16,391	20.6233 (16.1787)	Average (transaction year minus construction year) by Year, Dong, and Housing Type Apartments: 15.2000 (8.7844) Multi-Household Houses: 14.1278 (6.5037) Regular Houses: 38.0546 (18.1347) Dual Purpose Buildings: 10.0439 (5.6227)
Avg. Floor Location	16,456	4.7365 (4.0155)	Average number of floors by Year, Dong, and Housing Type Apartments: 8.6904 (3.1950) Multi-Household Houses: 2.6305 (1.1629) Regular Houses: 1.0000 (0.0000) Dual Purpose Buildings: 7.7985 (3.2172)
Elementary Schools	16,456	1.8743 (2.5126)	Number of Elementary Schools by Dong and Year
Middle Schools	16,456	1.1788 (1.7039)	Number of Middle Schools by Dong and Year
High Schools	16,456	0.9528 (1.6016)	Number of High Schools by Dong and Year
Subways	16,456	1.0601 (1.4480)	Number of Subway Stations by Dong and Year

Note: Standard deviations in parentheses. Unlike Table 2.2, this table shows the summary statistics of the variables after the construction of a Year×Dong×HousingType panel. During the observation period, an average of 310.4 legal-status dongs reported purchases of apartments. Regarding the other housing types, the annual average numbers of reporting dongs are 297.6 for multi-household houses, 343.9 for regular houses, and 223.6 for dual purpose buildings.

For the most part, Table 2.3 presents dong-level averages of variables already discussed in connection with Table 2.2, the notable exception being the transaction data. With more than two million transactions in just 14 years, Seoul's housing market must be described as very dynamic in nature. The annual average of more than 100 transactions per km^2 and dong in Table 2.3 underlines this fact.

The third goal of the empirical analysis is to estimate the effect "Gangnam Style" had on tourism in Gangnam. For this purpose, I shift the attention to the number of available hotel rooms in a district. Such data is offered by the Seoul Metropolitan Government and covers the period from 2009 to 2021 (i.e. one year less than the residential property registry data, which also includes 2022). For every district and year I observe the registered hotels and for every hotel the total number of rooms offered for rent.[30] From this data I have constructed an annual panel containing the total number of hotel rooms available in Seoul's 25 districts during the period from 2009 to 2021. The resulting variable is likely to contain information about the current number of tourists but also about the expected future number in a district at time t. Equation (2.3) summarizes the DID identification strategy:

$$y_{dt} = \beta_0 + \beta_1 DID_{dt} + \alpha^T NbhAttr_{dt} + \gamma_t + \delta_d + \varepsilon_{dt} \tag{2.3}$$

The outcome variable y_{dt} is the number of available hotel rooms in district d and year t. The indicator variable DID_{dt} equals one only when the outcome variable refers to Gangnam in 2012 or later. $NbhAttr_{dt}$ is a vector that refers to the number of elementary, middle, and high schools as well as subway stations and lines in district d in year t. As before, variables γ_t and δ_d represent year and district fixed effects. Again, I consider four different specifications of equation (2.3). On the one hand, I exclude the neighborhood attributes of schools and subways in Models 9 and 11, while in Models 10 and 12 they are included. On the other hand, I do not cluster the standard errors (ε_{dt}) on the respective district levels in Models 9 and 10, while in Models 11 and 12 I do. Given the sampling period of 13 years and the fact that Seoul has 25 districts, I no longer benefit from a large number of observations. Table 2.4 displays the descriptive statistics for the variables used in the hotel rooms regressions.

[30] Unfortunately, no dong-level hotel data is available.

Table 2.4: Descriptive Statistics for Hotel Rooms Regression Analysis

Variables	Obs	Mean	Definition
All Hotel Rooms	325	1680.9350 (2974.9450)	Number of all hotel rooms per district
Three Star Hotel Rooms	325	338.5785 (784.6191)	Number of three star hotel rooms per district
District Elementary Schools	325	23.9323 (7.4694)	Number of elementary schools per district
District Middle Schools	325	15.2923 (5.0562)	Number of middle schools per district
District High Schools	325	12.7385 (5.2689)	Number of high schools per district
District Subways	325	14.4615 (6.6585)	Number of subway stations per district

Note: Standard deviations in parentheses. Unlike Table 2.2, this table shows the summary statistics of the variables after using the dataset to construct a Year × District panel.

As can be seen in Table 2.4, the explanatory variables still display substantial variation, although district level aggregates are shown. Also, there are about 1,700 hotel rooms per district. Districts with many tourist attractions have more hotel rooms. For example Jung District has the most hotel rooms in Seoul as it is home to famous sights such as South Gate (Sungnyemun), East Gate (Dongdaemun), Dongdaemun History and Culture Park, Namsan Park with Seoul Tower, and the Myeong-dong shopping area. Districts with fewer tourist attractions (typically bedroom or commuter districts), in turn, have fewer hotel rooms (i.e. Seongdong, Yangcheon, and Gangbuk). Given that three-star hotels are most popular, I also use the available number of three-star hotel rooms as an outcome variable. Despite their popularity among domestic and international tourists, however, the three-star hotel rooms panel contains a few zero entries.

In addition to the twelve specifications of regression models (1), (2) and (3) introduced so far, I also perform a series of robustness checks. To this end, I make minor changes to the models. First, I modify the models to explicitly control for the intensity of treatment in a legal-status dong. To implement this, I have measured the *Distance*[31] from the Gangnam District Office to each legal-status dong and use this *Distance* measure in place of the *DID* indicator variables.[32] If the release of the song indeed disproportionately affected the housing market in Gangnam, and also the number of available hotel rooms, then it is likely that locations close to Gangnam have been more affected in terms of property prices and number of hotel rooms, and those further away from Gangnam have been less affected. Second, I control for additional fixed effects and introduce year-district fixed effects. With a total of 299 fixed effects, this modification is particularly strong in controlling for unobserved district-level (time-varying) changes, such as local housing loan policies and official land use policy changes.[33] Third, I reconsider the clustering levels of standard errors. Dong-level clustering is more robust with regards to regional correlation. Two-way (year and district) clustering, in turn, enables me to adjust not only for district-level regional correlations but also for possible auto-correlations of the error terms. The robustness of the estimated coefficients is established when I am able to generate significant regression results for all clustering levels considered.

2.4 Results

I start by estimating the four different specifications of equation (2.1). Table 2.5 displays the regression results when using residential property prices per square meter as the outcome variable.

---

[31] Just like every district, each administrative dong typically has an administrative center (often called a town office). For the distance variable, I have simply measured the straight lineal distance between the Gangnam District Office and the administrative center in a legal-status dong (which are of interest here). If a particular legal-status dong does not have an administrative center, I have looked for a subway station's main exit 1, then a bus stop, and lastly a famous tourist destination.

[32] Note that the non-negative distance measure works somewhat in reverse to the DID indicator variable. If the observation year is 2012 or later, it assigns a (possibly large) positive number to a legal-status dong that belongs to the control group (where DID equals zero) but it assigns a small number to a legal-status dong that belongs to the treatment group (where DID equals one). If the observation year is before 2012, it assigns zero distance.

[33] Seoul has a total of four land use zonings (residential, commercial, industrial, and green belt).

In this analysis, I consider all 25 districts of ***Seoul Metropolitan City*** during the time span between 2009 and 2022. The treated district is Gangnam from 2012 to 2022.

Table 2.5: Regression Results for Square Meter Prices (Seoul Metropolitan City)

	Model 1	**Model 2**	**Model 3**	**Model 4**
DID	166.6478***	212.3474***	184.1798***	167.9614***
	(40.9001)	(49.3083)	(44.7163)	(48.7834)
P>|t|	0.000	0.000	0.000	0.001
Multi-Household House	-222.8450***	-193.8743***	-213.6120***	-194.3552***
	(10.9802)	(10.3256)	(10.9865)	(10.3077)
P>|t|	0.000	0.000	0.000	0.000
Regular House	-97.9366***	-74.1660***	-88.3712***	-75.4935***
	(12.4676)	(10.6113)	(12.5311)	(10.6271)
P>|t|	0.000	0.000	0.000	0.000
Dual Purpose Building	-296.4010***	-285.9439***	-288.0098***	-286.0096***
	(22.4162)	(18.6706)	(20.0886)	(18.6755)
P>|t|	0.000	0.000	0.000	0.000
Age of Building	-1.2406***	-2.3720***	-1.5005***	-2.3262***
	(0.4523)	(0.3995)	(0.4453)	(0.3982)
P>|t|	0.006	0.000	0.001	0.000
Floor	7.0654***	5.8218***	6.6888***	5.7981***
	(0.4397)	(0.3300)	(0.4217)	(0.3282)
P>|t|	0.000	0.000	0.000	0.000
Elementary Schools	-2.0720	-8.4460	7.1517***	-7.6208
	(1.9987)	(9.6329)	(1.8126)	(9.7028)
P>|t|	0.301	0.381	0.000	0.433
Middle Schools	-5.5076	1.6571	-14.4995***	-2.5878
	(4.2199)	(17.8486)	(3.0755)	(17.5683)
P>|t|	0.193	0.926	0.000	0.883
High Schools	10.3613***	40.9272*	10.5710***	38.5690*
	(2.9482)	(24.6721)	(2.6822)	(21.1672)
P>|t|	0.000	0.098	0.000	0.069
Subways	yes	yes	no	no
17 Subway Line Dummies	no	no	yes	yes
Observations	2,069,845	2,069,845	2,069,845	2,069,845
Year Fixed Effects	Yes	Yes	Yes	Yes
District Fixed Effects	Yes	Yes	Yes	Yes
Dong Fixed Effects	No	Yes	No	Yes
Clusters (year-district)	Yes (350)	Yes (350)	Yes (350)	Yes (350)

Note: Each column tabulates regression coefficients obtained from estimating the specifications of equation (2.1). Regression results related to subway-related coefficients are omitted for brevity here and can be found in Appendix Table 2.16. Standard errors are in parenthesis. *,**,*** denote statistical significance at the 10%, 5% and 1% level, respectively.

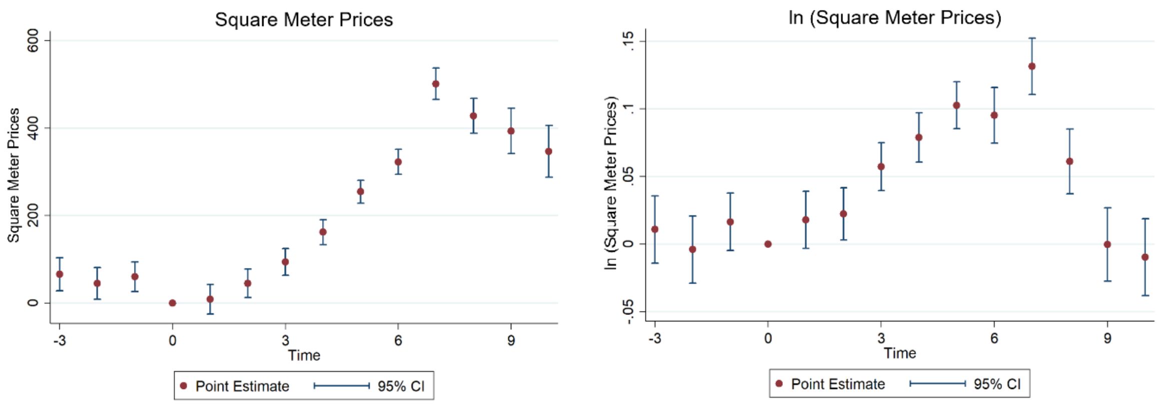
As can be seen, depending on the model specification average square meter prices of residential properties in Gangnam increased by 1,666,478 to 2,123,474 KRW (ca. 1,390 to 1,770 USD) after the release of Psy's viral hit. With regard to house characteristics I obtain clear results across all four specifications. Housing type exerts a significant influence on estimated prices per square meter. Most expensive are apartments followed by regular and multi-household houses. Dual-purpose buildings are the least expensive. Unsurprisingly, building age has a negative impact on prices, which is highly significant across all specifications. Similarly, apartments on higher floors tend to be significantly more expensive than those on lower floors.

With regard to neighborhood attributes, the four specifications do not produce uniform results. Take, for example, the measures for a legal-status dong's connectivity. The estimated coefficients for the Subways count variable (but also for the 17 subway line dummies) all depend on whether or not I control for legal-status dong fixed effects (tabulations of these estimates are shown in Appendix Table 2.16). Without dong fixed effects the number of subway stations and lines has a positive and significant relationship with residential property prices. An additional subway station increases local residential property prices by approximately 90,583 KRW (75.50 USD) per square meter. When controlling for dong fixed effects, however, this relationship turns negative with a p-value of about 12 percent. Most likely, the subway count variable picks up legal-status dong characteristics that have a positive impact on residential property prices which are commonly shared by dongs with good connectivity (such as good access to local shopping). It is for this reason that I prefer specifications that include legal-status-dong fixed effects. Moreover, measuring the connectivity effect on residential property prices is likely to be more accurate using the 17 subway line dummies rather than the Subways count variable. For instance, consider Line 9 and the Bundang Line (in Appendix Table 2.16: Line 10). Both subway lines are relatively new and both serve Gangnam. As can be expected, housing located near these two lines have higher average prices than those near other subway lines even after controlling for year, district, and dong fixed effects. Thus, I choose Model 4 as the preferred specification.

Next, I turn to school-related neighborhood attributes and start with the number of elementary and middle schools in a legal-status dong. In three specifications (including Model 4) I find negative relationships with residential property prices, a result that is consistent with the results produced by Lee & Choi (2016) who apply the hedonic house price model to a small legal-status dong (Ichon) in Seoul. Note, however, that the findings are distant from conventional significance levels. The number of high schools in a legal-status dong, in contrast, is positively related to residential property prices, which is also consistent with Lee & Choi (2016).

The left panel of Figure 2.1 shows how the estimated average treatment effect on residential property prices in Gangnam (i.e. the DID coefficient of Model 4) is distributed over the treated years. In addition, the panel also reveals the pre-existing trends of housing prices in Gangnam. Time 0 refers to the year 2012 when the song "Gangnam Style" was released. As can be seen, residential property prices in Gangnam kept increasing relative to other districts in Seoul until 2019 when they started to decline slowly while still being relatively high. As a result the "Gangnam Style" effect persisted even beyond 2020 when the COVID-19 pandemic hit the tourism industry (Park, 2022). The panel also illustrates that housing prices in Gangnam were not very different from prices in the control districts prior to 2012.

Figure 2.1: Annual Effects of Gangnam Style on Housing Prices in Gangnam



Note: This figure shows how the average treatment effects on residential property prices (left panel) and on the growth rates of residential property prices (right panel) are distributed over the treated years. Any pre-existing trends can also be inspected. Using the same covariates as in Model 4, the estimation of these annual effects follows the procedure suggested by Clarke and Tapia-Schythe (2021).

Still using the previously employed four specifications of equation (2.1), I next identify the effect of "Gangnam Style" on the growth rates of property prices by using the natural logarithm of square meter prices as the outcome variable. Table 2.6 shows the regression results.

Table 2.6: Regression Results for Log of Square Meter Prices (Seoul Metropolitan City)

	Model 1	**Model 2**	**Model 3**	**Model 4**
DID	0.0327	0.0746***	0.0452**	0.0467***
	(0.0233)	(0.0158)	(0.0206)	(0.0150)
P>|t|	0.161	0.000	0.029	0.002
Multi-Household House	-0.3453***	-0.3073***	-0.3344***	-0.3074***
	(0.0097)	(0.0096)	(0.0096)	(0.0096)
P>|t|	0.000	0.000	0.000	0.000
Regular House	-0.2223***	-0.1957***	-0.2121***	-0.1961***
	(0.0144)	(0.0132)	(0.0146)	(0.0132)
P>|t|	0.000	0.000	0.000	0.000
Dual Purpose Building	-0.3724***	-0.3723***	-0.3691***	-0.3721***
	(0.0167)	(0.0140)	(0.0154)	(0.0140)
P>|t|	0.000	0.000	0.000	0.000
Age of Building	-0.0044***	-0.0056***	-0.0046***	-0.0056***
	(0.0006)	(0.0005)	(0.0006)	(0.0005)
P>|t|	0.000	0.000	0.000	0.000
Floor	0.0092***	0.0078***	0.0087***	0.0078***
	(0.0004)	(0.0003)	(0.0003)	(0.0003)
P>|t|	0.000	0.000	0.000	0.000
Elementary Schools	-0.0029	-0.0062	0.0049***	-0.0044
	(0.0022)	(0.0076)	(0.0018)	(0.0077)
P>|t|	0.190	0.413	0.008	0.565
Middle Schools	-0.0163***	-0.0271	-0.0257***	-0.0300
	(0.0045)	(0.0197)	(0.0036)	(0.0194)
P>|t|	0.000	0.170	0.000	0.123
High Schools	0.0163***	0.0350*	0.0168***	0.0396**
	(0.0030)	(0.0180)	(0.0030)	(0.0188)
P>|t|	0.000	0.052	0.000	0.036
Subways	yes	yes	no	no
17 Subway Line Dummies	no	no	yes	yes
Observations	2,069,845	2,069,845	2,069,845	2,069,845
Year Fixed Effects	Yes	Yes	Yes	Yes
District Fixed Effects	Yes	Yes	Yes	Yes
Dong Fixed Effects	No	Yes	No	Yes
Clusters (year-district)	Yes (350)	Yes (350)	Yes (350)	Yes (350)

Note: Each column tabulates regression coefficients obtained from estimating the specifications of equation (2.1). Regression results related to subway-related coefficients are omitted for brevity. Standard errors are in parenthesis. *,**,*** denote statistical significance at the 10%, 5% and 1% level, respectively.

According to Model 4, residential property prices in Gangnam increased by approximately 4.8 percent compared to the control districts after the song's release. By and large, the patterns of signs and statistical significance for house characteristics and neighborhood attributes resemble the patterns in Table 2.5. The right panel in Figure 1 illustrates how the estimated average effect on residential property price growth rates in Gangnam is composed from annual treatment effects. As in the left panel, I can also inspect any pre-existing trends.

According to the right panel in Figure 1, there is no evidence for a pre-existing differential trend in housing price growth rates in Gangnam compared to the controlled districts. Only after the release of "Gangnam Style" in 2012 can I observe an accelerated growth of square meter prices in Gangnam relative to other districts. As in the left panel, the effects last until the start of the COVID-19 pandemic in 2020.

In the case where the COVID-19 pandemic indeed affected residential property prices, the interpretation of the estimated "Gangnam Style" effect would be unnecessarily complicated. I therefore conduct additional analyses using trimmed datasets. On the one hand I exclude observations from 2021 and 2022 and on the other hand I make incremental restrictions on the control districts (see also Table 2.1). Regarding the latter, I first rerun the regressions for the Metropolitan City of Seoul. Then, I restrict the attention to observations from Seoul to the South of the Han River. Finally, I only use the three neighboring districts of Gangnam, Seocho and Songpa.[34] Excluding observations from more and more districts renders the analyzed part of Seoul increasingly homogeneous such that I increasingly control for unobservable potential determinants of residential property prices.[35] Table 2.7 shows the regression results for the trimmed datasets.

[34] Gangnam is literally translated as "South of the River." Moreover, Seoulites often refer to the combined area of the three districts of Gangnam, Seocho and Songpa as the "Gangnam Region", reflecting the fact that Gangnam was split into three districts in 1988, thereby establishing the districts of Seocho and Songpa.
[35] Historically, Seoul was completely located to the North of the Han River. Only after the Second World War did the city start to expand South of the river, so that this part of Seoul is much less heterogeneous than the more traditional quarters to the North of the Han River. Likewise, the three neighboring districts of Gangnam, Seocho and Songpa resemble each other in terms of urban planning, public infrastructure, and job opportunities.

Table 2.7: Regression Results for (Log of) Square Meter Prices using Trimmed Datasets

	Model 1	Model 2	Model 3	Model 4
Seoul Metropolitan City from 2009 to 2020				
Square Meter Price	152.8086***	175.2363***	154.4379***	139.7650***
	(48.7512)	(52.5072)	(49.0295)	(52.2832)
P>|t|	0.002	0.001	0.002	0.008
Ln (Square Meter Price)	0.0608***	0.0855***	0.0606***	0.0549***
	(0.0144)	(0.0146)	(0.0163)	(0.0141)
P>|t|	0.000	0.000	0.000	0.000
Observations	1,714,881	1,714,881	1,714,881	1,714,881
Clusters (year-district)	Yes (300)	Yes (300)	Yes (300)	Yes (300)
Seoul to the South of the Han River from 2009 to 2020				
Square Meter Price	140.2927***	140.6592***	154.4719***	150.7932***
	(45.1001)	(48.1575)	(47.1723)	(49.3008)
P>|t|	0.002	0.004	0.001	0.003
Ln (Square Meter Price)	0.0499***	0.0695***	0.0679***	0.0657***
	(0.0152)	(0.0154)	(0.0195)	(0.0149)
P>|t|	0.001	0.000	0.001	0.000
Observations	864,630	864,630	864,630	864,630
Clusters (year-district)	Yes (132)	Yes (132)	Yes (132)	Yes (132)
Gangnam, Seocho and Songpa from 2009 to 2020				
Square Meter Price	75.0160**	89.0744**	79.0221**	87.1756**
	(29.1112)	(35.5372)	(37.0541)	(35.9103)
P>|t|	0.014	0.017	0.040	0.020
Ln (Square Meter Price)	0.0455**	0.0492**	0.0528*	0.0488**
	(0.0172)	(0.0215)	(0.0268)	(0.0184)
P>|t|	0.012	0.029	0.056	0.012
Observations	264,708	264,708	264,708	264,708
Clusters (year-district)	Yes (36)	Yes (36)	Yes (36)	Yes (36)
Year Fixed Effects	Yes	Yes	Yes	Yes
District Fixed Effects	Yes	Yes	Yes	Yes
Dong Fixed Effects	No	Yes	No	Yes

Note: Each entry is a result obtained from estimating equation (2.1). The control variables are the same as in the regressions reported in Tables 2.5 and 2.6. For simplicity and brevity, however, only the estimated DID coefficients are tabulated. Standard errors are in parentheses. *,**,*** denote statistical significance at the 10%, 5% and 1% levels, respectively.

Table 2.7 clearly illustrates that residential property prices in Gangnam outperformed prices in other districts, whichever group of control districts is chosen. When excluding observations from the pandemic years 2021 and 2022 but otherwise focusing the analysis on ***Seoul Metropolitan City***, I find slightly smaller effects of the release of "Gangnam Style" on per square meter residential property prices. This finding, of course, is simply a reflection of the fact that the annual treatment effects, while declining, are still above the average treatment effect in 2021 and 2022 (see the left panel in Figure 1). A similar logic explains the slightly greater effects of "Gangnam Style" on the logarithms of square meter prices. With negative growth rates during of the pandemic years (see the right panel in Figure 1), the estimated DID coefficient indeed increases after trimming the dataset. When restricting the analyses to ***Seoul to the South of the Han River***, I find larger price effects both in absolute terms and in terms of growth rates. Since the release of the song, residential property prices in Gangnam are estimated to have grown to be on average 1,507,932 KRW (1,257 USD) more expensive than in the remainder of the southern districts, and the average growth rate of prices in Gangnam is about 6.8 percent higher than them. Lastly, when restricting the analysis to ***Gangnam, Seocho and Songpa***, I find somewhat smaller effects. Compared to neighboring Seocho and Songpa, housing prices in Gangnam went up by 871,756 KRW (727 USD). The residential property price growth rate in Gangnam accelerated by about 5.0 percent compared to the two neighboring districts.

My next goal is to estimate the effect of "Gangnam Style" on the number of transactions in the housing market. As mentioned above, I have created a three-dimensional panel with a focus on transactions per legal-status dong for this purpose. Table 2.8 shows the regression results when estimating equation (2.2).

Table 2.8: Regression Results for Transactions per km^2 (Seoul Metropolitan City)

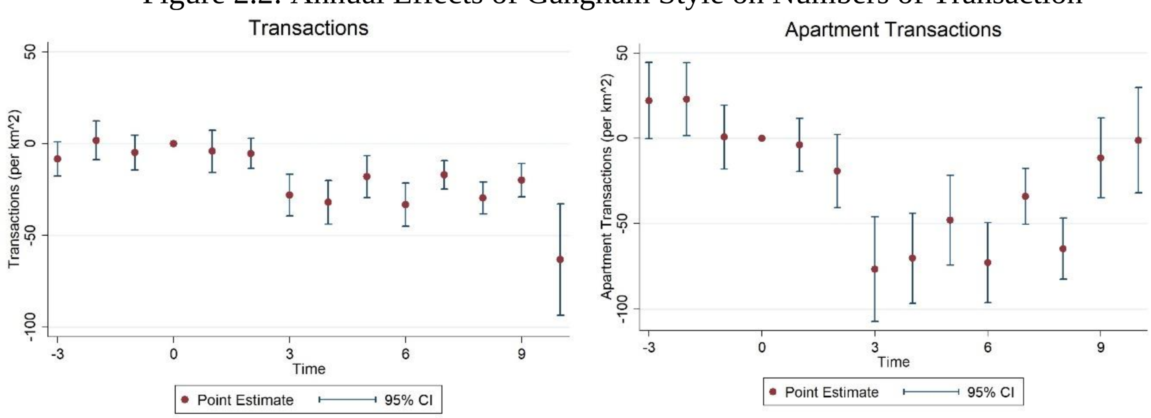
	All Housing Types		Only Apartments	
	Model 5	**Model 6**	**Model 7**	**Model 8**
DID	-19.6130***	-18.8039***	-58.6118***	-51.1292***
	(5.9257)	(6.5324)	(11.9893)	(11.3919)
P>|t|	0.001	0.004	0.000	0.000
Multi-Household House	1.9145	2.0317		
	(7.3025)	(7.3096)		
P>|t|	0.793	0.781		
Regular House	-32.2445***	-32.1594***		
	(6.3637)	(6.3713)		
P>|t|	0.000	0.000		
Dual purpose Building	-71.4948***	-71.5199***		
	(9.2138)	(9.2117)		
P>|t|	0.000	0.000		
Avg Age of Building	-0.2630*	-0.2604*	-5.2574***	-5.3438***
	(0.1458)	(0.1457)	(0.8152)	(0.8257)
P>|t|	0.072	0.075	0.000	0.000
Avg Floor Location	12.8097***	12.8299***	0.9105	0.9285
	(0.8843)	(0.8850)	(1.5496)	(1.5439)
P>|t|	0.000	0.000	0.557	0.548
Elementary Schools	-3.5659	-2.7263	-6.1885	-5.9563
	(4.2668)	(4.2114)	(7.8921)	(8.1354)
P>|t|	0.404	0.518	0.433	0.465
Middle Schools	-3.9072	-4.9427	-13.6108	-9.8163
	(10.4496)	(10.3114)	(18.4595)	(19.7145)
P>|t|	0.709	0.632	0.461	0.619
High Schools	-5.0226	-5.4222	-0.5191	-3.2515
	(11.8675)	(12.2614)	(28.9305)	(28.8969)
P>|t|	0.672	0.659	0.986	0.910
Subways	yes	no	yes	no
17 Subway Line Dummies	no	yes	no	yes
Observations	16,391	16,391	16,391	16,391
Year Fixed Effects	Yes	Yes	Yes	Yes
District Fixed Effects	Yes	Yes	Yes	Yes
Dong Fixed Effects	Yes	Yes	Yes	Yes
Clusters (year-district)	Yes (350)	Yes (350)	Yes (350)	Yes (350)

Note: Each column tabulates regression coefficients obtained from estimating the four specifications of equation (2.2). Estimates related to connectivity are omitted for brevity. Standard errors are in parentheses. *,**,*** denote statistical significance at the 10%, 5% and 1% levels, respectively.

As I speculated in the previous section, I indeed find negative DID coefficients. Regarding the all-house-types regressions, the number of transactions per square kilometer decreased by about

18.8 in Gangnam compared to the rest of Seoul after the release of the song in 2012.[36] Residential properties with favorable access to schools and public transportation are the subject of fewer transactions. Compared to apartments, multi-household houses are involved in transactions more often while regular houses and dual-purpose buildings are involved less often. Interestingly, the sign patterns related to the house characteristics building age and floor location is the same for the price and for the transaction regressions. Neighborhood attributes, in contrast, display an inconsistent pattern. With a decline of more than 50 transactions per square kilometer, the "Gangnam Style" effect is estimated to be even stronger when focusing on apartments only. The two panels in Figure 2.2 show the annual effects that constitute the pre-existing trends and make up the DID coefficients in Table 2.8. As can be seen, the number of transactions per km^2 in Gangnam relative to the control districts is not statistically different until 2012. After "Gangnam Style" was released, however, I see fewer transactions in Gangnam compared to the rest of Seoul, in particular from 2015 to 2020/21.

Figure 2.2: Annual Effects of Gangnam Style on Numbers of Transaction



Note: This figure shows how the average treatment effects on the numbers of transaction involving all housing types (left panel) and only apartments (right panel) are distributed over the treated years, and also any pre-existing trends. Using the same covariates as in Model 6 and Model 8, the estimation of these annual effects follows the procedure suggested by Clarke and Tapia-Schythe (2021).

[36] Such a decline in transactions per square kilometer corresponds to 17.8 percent compared to its mean.

As for the price regressions, I rerun the transaction regressions for increasingly homogeneous regions of Seoul (again I use ***Seoul to the South of the Han River*** and the three ***Neighboring Districts Gangnam, Seocho and Songpa***) between 2009 and 2020. Table 2.9 summarizes the results obtained from these regressions.

Table 2.9: Regression Results for Transactions per km^2 using Trimmed Datasets

	All Housing Types		**Only Apartments**	
	Model 5	**Model 6**	**Model 7**	**Model 8**
Seoul Metropolitan City from 2009 to 2020				
Transactions	-16.7804***	-16.3469***	-59.3909***	-56.6538***
	(5.0794)	(5.1479)	(12.9008)	(12.1235)
P>|t|	0.001	0.002	0.000	0.000
Observations	14,046	14,046	3,695	3,695
Clusters (year-district)	Yes (300)	Yes (300)	Yes (300)	Yes (300)
Seoul to the South of the Han River from 2009 to 2020				
Transactions	-18.9771***	-19.1710***	-51.7026***	-52.1717***
	(5.3081)	(5.4587)	(13.6617)	(13.3677)
P>|t|	0.000	0.001	0.000	0.000
Observations	4,573	4,573	1,279	1,279
Clusters (year-district)	Yes (132)	Yes (132)	Yes (132)	Yes (132)
Gangnam, Seocho and Songpa from 2009 to 2020				
Transactions	-7.7696	-8.5719	-14.2266	-13.9580
	(5.7082)	(5.8783)	(8.5578)	(9.1362)
P>|t|	0.182	0.154	0.105	0.136
Observations	1,414	1,414	398	398
Clusters (year-district)	Yes (36)	Yes (36)	Yes (36)	Yes (36)
Year Fixed Effects	Yes	Yes	Yes	Yes
District Fixed Effects	Yes	Yes	Yes	Yes
Dong Fixed Effects	Yes	Yes	Yes	Yes

Note: The table shows the estimated DID coefficients obtained from estimating the specifications of equation (2.2) when using transactions per km^2 as the outcome variable. The control variables are the same as in the regressions reported in Table 2.8. In particular, I use the Subways count variable in Models 5 and 7 and the 17 subway line dummies in Models 6 and 8. Standard errors are in parenthesis. *,**,*** denote statistical significance at the 10%, 5% and 1% levels, respectively.

Table 2.9 reconfirms that the number of transactions per km^2 decreases significantly after 2012. Compared to districts in the rest of ***Seoul to the South of the Han River***, the number of transactions per km^2 in Gangnam declines by about 19 units. Restricting the regressions even further to the wider Gangnam region, I still find a decline of eight transactions per km^2 in Gangnam compared to neighboring Seocho and Songpa (however with p-values between 15 and 20 percent this is not significant at conventional levels). In general, the results imply that the turnover in Gangnam's housing market falls following the release of ”Gangnam Style.” Moreover, the estimated “Gangnam Style” effects are stronger and more precisely estimated when restricting the attention to apartments alone.

My next goal is to estimate the effect of Psy's hit song on the available number of hotel rooms in Gangnam. Such a measure of supply is likely to reflect both the current number of visitors but also expected future visits. Table 2.10 summarizes the regression results.

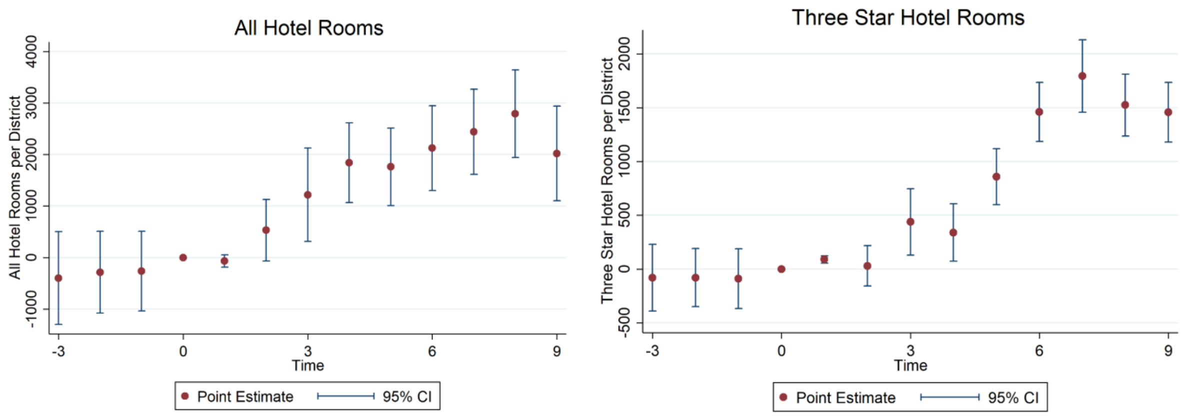
Table 2.10: Regression Results for Numbers of Hotel Rooms (Seoul Metropolitan City)

	Model 9	**Model 10**	**Model 11**	**Model 12**
DID	1,537.1280**	1,567.3140**	1,537.1280***	1,567.3140***
	(660.0613)	(700.3302)	(311.0699)	(513.8115)
P>|t|	0.021	0.026	0.000	0.006
District Elementary Schools	no	-35.2737	no	-35.2737
		(119.0672)		(160.5782)
P>|t|		0.767		0.828
District Middle Schools	no	-50.5744	no	-50.5744
		(187.6666)		(218.8025)
P>|t|		0.788		0.819
District High Schools	no	-992.6411***	no	-992.6411
		(174.1249)		(694.3681)
P>|t|		0.000		0.166
District Subways	no	-35.3707	no	-35.3707
		(43.9571)		(86.5783)
P>|t|		0.422		0.686
Observations	325	325	325	325
Year Fixed Effects	Yes	Yes	Yes	Yes
District Fixed Effects	Yes	Yes	Yes	Yes
Cluster (Districts)	No	No	Yes (25)	Yes (25)

Note: Each column tabulates regression coefficients obtained from estimating the specifications of equation (2.3) when using the available number of hotel rooms in general and of three star hotel rooms as outcome variables. Standard errors are in parentheses. *,**,*** denote statistical significance at the 10%, 5% and 1% levels, respectively.

As shown in Table 2.10, the number of hotel rooms in Gangnam increased significantly after “Gangnam Style” was released in 2012. To be precise: I estimate that the supply of hotel rooms increased by about 1,567 units (or by about 93 percent of the Seoul-wide average) compared to other districts in Seoul. The left panel in Figure 2.3 provides a graphical illustration of these results and shows how the estimated DID coefficient can be split across the treated years. The panel also displays the respective pre-treatment coefficients.

Figure 2.3: Annual Effects of Gangnam Style on the Available Numbers of Hotel Rooms



Note: This figure shows how the average treatment effects on numbers of available hotel rooms (left panel: all hotel rooms, right panel only three-star hotel rooms) are distributed over the treated years, and also any pre-existing trends. Using the same covariates as in Model 12 as well as district-level clustering, the estimation of these annual effects follows the procedure suggested by Clarke and Tapia-Schythe (2021).

As can be seen in the left panel in Figure 2.3, before the release of the song the difference between the number of hotel rooms in Gangnam and the other districts of Seoul is not statistically different from zero. After "Gangnam Style" was released, however, I see a drastic increase in the number of available hotel rooms compared to the control districts.

As before, I assess the robustness of the findings by re-estimating the regressions for more homogeneous regions of Seoul. To be precise, I first restrict the analysis to ***Seoul to the South of the Han River*** and then to the three districts ***Gangnam, Seocho and Songpa***. Table 2.11 shows the results of these estimations. In addition the table presents the findings when replacing hotel rooms with three star hotel rooms, as this is the most popular category among domestic and international tourists.

Table 2.11: Hotels Regression Results Using Various Control Regions

	Model 9	**Model 10**	**Model 11**	**Model 12**
	Seoul Metropolitan City from 2009 to 2021			
All Hotel Rooms	1537.1280**	1567.3140**	1537.1280***	1567.3140***
	(660.0613)	(700.3302)	(311.0699)	(513.8115)
	0.021	0.026	0.000	0.006
Three Star Hotel Rooms	821.2875***	782.8063***	821.2875***	782.8063***
	(270.0736)	(288.6862)	(114.7238)	(182.4975)
	0.003	0.007	0.000	0.000
Observations	325	325	325	325
Clusters (Districts)	No	No	Yes (25)	Yes (25)
	Seoul to the South of the Han River from 2009 to 2021			
All Hotel Rooms	1787.7870***	1022.3050***	1787.7870***	1022.3050*
	(307.2134)	(366.1069)	(187.2487)	(460.0699)
	0.000	0.006	0.000	0.051
Three Star Hotel Rooms	898.9833***	572.1942***	898.9833***	572.1942**
	(153.8642)	(187.3023)	(33.9799)	(254.3451)
	0.000	0.003	0.000	0.048
Observations	143	143	143	143
Clusters (Districts)	No	No	Yes (11)	Yes (11)
	Gangnam, Seocho and Songpa from 2009 to 2021			
All Hotel Rooms	1756.6000***	1938.8510***	1756.6000**	1938.8510***
	(444.4943)	(466.6917)	(199.6271)	(95.9296)
	0.001	0.001	0.013	0.002
Three Star Hotel Rooms	821.3167***	1240.8360***	821.3167***	1240.8360***
	(266.8771)	(275.5605)	(32.6713)	(99.9148)
	0.005	0.000	0.002	0.006
Observations	39	39	39	39
Clusters (Districts)	No	No	Yes (3)	Yes (3)
Year Fixed Effects	Yes	Yes	Yes	Yes
District Fixed Effects	Yes	Yes	Yes	Yes

Note: The table shows the estimated DID coefficients for the outcome variables hotel rooms and three star hotel rooms. In Models 10 and 12, I control for district-level numbers of elementary, middle and high schools as well as for the number of subway stations and lines in a district, using the Subways count variable. In Models 9 and 11, I do not. Standard errors are in parenthesis. *,**,*** denote statistical significance at the 10%, 5% and 1% levels, respectively.

Table 2.11 shows that the number of (three-star) hotel rooms in Gangnam increased by about 1,022 rooms (572 rooms) compared to other districts in Seoul to the South of the Han River. Moreover, compared to its two neighboring districts Seocho and Songpa hotel capacities in Gangnam expanded by approximately 1,939 rooms (1,240 rooms). Note that all results in Tables 2.10 and 2.11 are significant at conventional levels. Hence, I argue that the number of available hotel rooms in Gangnam, including three-star, significantly increased after the song was released in 2012. With respect to three-star hotel rooms, the right panel in Figure 3 shows how the DID coefficient can be distributed among annual treatment effects after the release of the song. Also, the panel suggests that there was no pre-existing trend.

Both panels in Figure 2.3 display similar patterns. The number of available hotel rooms has increased dramatically since 2012 and peaked in 2019/20 just before the COVID-19 pandemic. Considering that the ban on entry to Korea was the most drastic measure implemented by the Korean government in the fight against COVID-19, the drop in the number of available hotel rooms in 2021 is most likely a consequence of the pandemic. Interestingly, both panels in Figure 2.3 resemble the left panel in Figure 2.1 (square meter price regressions) in shape and slope. This similarity supports the hypothesis that the residential property price increase in Gangnam is likely to be driven by a current but also an expected future increase in domestic and international tourist numbers in Gangnam.

I begin additional robustness checks by using an alternative treatment measure, namely the distance between a legal-status dong and the Gangnam District Office. With the replacement of the DID coefficient by the distance measure, I effectively relax the assumption that Gangnam is the only treated district. The distance measure also allows me to assess whether legal-status dongs close to the Gangnam District Office experienced an increase in residential property prices but also a decrease in the number of property transactions. Appendix Table 2.17 shows that residential properties located far from Gangnam District Office are trading at a discount compared to those closer to the office. Moreover, this discount remains significant when restricting the analysis to increasingly homogeneous regions such as Seoul to the South of the Han River or to the three neighboring districts Gangnam, Seocho and Songpa. Regarding transactions, Appendix Table 2.18 shows that the further a legal-status dong is from the Gangnam

District Office, the more property purchases it reports (although for the Metropolitan City of Seoul this effect is not significant). Finally, the hotel room regressions (Appendix Table 2.19) generally show that the number of hotel rooms decreases the further away a district is from the Gangnam District Office.

In my second set of robustness checks, I additionally control for year–district dummies (i.e. the inner product of year and district dummies) to take the possibility into account that there might be (unobserved) time-varying changes at the district level such as population size, housing loan policy and land use zoning changes. Panel A in Appendix Table 2.20 shows that even after controlling for time-varying district-level characteristics, square meter prices in Gangnam are estimated to be 4,463,312 KRW (3719 USD) higher than in the rest of Seoul since the release of "Gangnam Style." Similarly, Panel B implies that the price markup in Gangnam after 2012 is about 11.4 percent. Again, the alternative treatment measure – the distance to the Gangnam District Office – shows that residential properties sell at a greater discount the further away they are from the Office.

In the last set of robustness checks, I consider two different ways to cluster standard errors. To begin with, I cluster them at the legal-status-dong level to adjust for possible regional correlations. Furthermore, I consider two-way clustering (year and district) to adjust for both regional correlations and auto-correlations. Despite using various levels of clusters and methods of clustering, the estimates generally remain statistically significant (see Appendix Tables 2.21, 2.22, and 2.23). I thereby reconfirm the "Gangnam Style" effect on the housing market in Seoul.

2.5 Discussion

My analysis reveals the strong impact of the song "Gangnam Style" on Seoul's housing market. As Tables 2.5, 2.6, and 2.7 show, residential properties prices in Gangnam have increased by

1,679,614 KRW (1,400 USD) per m^2 (or about 4.8 percent in terms of growth rates) compared to other districts in Seoul. Despite the increase in residential properties prices, the actual number of transactions has in fact decreased, as shown in Tables 2.8 and 2.9. The results also reveal that the number of hotel rooms that are available to accommodate tourists in Gangnam has simultaneously increased (see Tables 2.10 and 2.11). Moreover, the regression results prove robust when an alternative treatment measure (i.e. distance from Gangnam District Office), increasingly homogeneous estimation samples, various levels of standard error clustering and controlling for different sets of fixed effects are employed (see Appendix Tables 2.17, 2.18, 2.19, and 2.20). The findings are consistent with the hypothesis that the success of "Gangnam Style" has boosted property owners' asking prices and hence let to a contraction of supply.

By and large, the increase in residential property prices has been driven by apartment prices (see Tables 2.8 and 2.9). Further evidence for this hypothesis is presented by making use of the three dimensional structure of the panel dataset. In particular, I rerun the regressions after interacting the DID coefficients with house-type dummies. The results show that apartment prices (and their growth rates) have increased while prices of dual-purpose buildings (and their growth rates) have decreased. Besides the increase in residential property prices, the three dimensional panel also reconfirms that the numbers of the transactions have decreased (see Appendix Table 2.24).

One possible and direct explanatory pathway for the residential property price increase in Gangnam might be the influx of (international) tourists after the song was released in 2012. Although the share of international tourists visiting Gangnam relative to other parts of Seoul before 2014 is unknown, I do know that this share has increased since 2014 despite Seoul losing ground as a tourist destination relative to the rest of South Korea. Another indication that tourism played a key role in price increases is the fact that residential property price growth rates turned negative when the COVID-19 pandemic started. With a strict entry ban in place, the number of domestic and international tourists declined drastically. In Gangnam, this decline in tourist activity and residential property price growth rates show a strong positive relationship. The hotel room regression results further strengthen this point. Since hotel capacities are good proxy variables for tourism (Mckercher & Lau, 2008; Simmons, 1984; Tremblay, 1998), the increased number of available hotel rooms after 2012 also supports the possible connection between tourism and residential property prices in Gangnam. The number of local restaurants is also a good proxy variable for tourism (Cohen & Avieli, 2004; Erkuş-Öztürk & Terhorst, 2016; Min & Lee, 2014). Although not presented in the previous section, regression analyses using numbers of restaurants in legal-status dongs as the outcome variable leads to results that are comparable to the hotel regressions. Again, I find a significant and positive effect of "Gangnam Style", which stresses once more the role of tourism (see Appendix Table 2.25).

Another possible and indirect pathway for the contraction of supply to have occurred in Gangnam is the change of expectations among property owners. Anticipating more tourists visiting Gangnam in the foreseeable future, property owners may have immediately sought higher prices. The prospects of improved public transportation and more hotels and private businesses in the near future brought about by an increase in tourism (Gronau & Kagermeier, 2007; Polo-Peña et al., 2012; Smith, 1983) may have merely reinforced expectations among property owners. As a result, I find an increased hesitation to sell property. Appendix Tables 2.18 and 2.26 support the hypothesis of the expectations pathway with further evidence. According to Table 2.18, the number of transactions rises with increasing distance from the district office in Gangnam. This effect becomes more pronounced with smaller and increasingly homogeneous control groups while significance levels simultaneously improve. These findings are again consistent with the notion that owners of residential properties in Gangnam became more hesitant to sell their property after the release of the song. Table 2.26 in the Appendix shows that residential property prices (and their growth rates) have increased even after controlling for the number of transactions, which also points to the supply side, in particular rapidly adjusting expectations. Reassuringly, these results are most pronounced for apartments, the housing type with the highest turnover in Seoul’s real estate market.

2.6 Conclusion

The spontaneous and viral popularity of “Gangnam Style” was one of the first international successes of K-Pop culture and it has since become an exemplary moment in the so-called Korean Wave. The success of the song inspired domestic and international tourists to visit Gangnam, a district of Seoul previously not considered a main tourist destination. I suggest that the distinctive development of the housing market in Gangnam relative to the rest of Seoul since 2012 is due to the district’s sudden high profile. Increasing tourist numbers may have either directly driven up the value of land in Gangnam or at least fueled the expectations of property owners regarding future tourism-related improvements of the area.

The empirical investigation has revealed that the release of “Gangnam Style” has caused the following three effects (among others). Compared to the rest of Seoul, the district of Gangnam experienced an increase in house prices, a reduction in transactions, and a rise in the number of available hotel rooms. In identifying these ”Gangnam Style“ effects I benefited from a treatment that is precisely defined in terms of time (2012 and after), and location (Seoul’s Gangnam District). Given this structure, I employ the simple and clear event-study (or differences-in-differences) design.

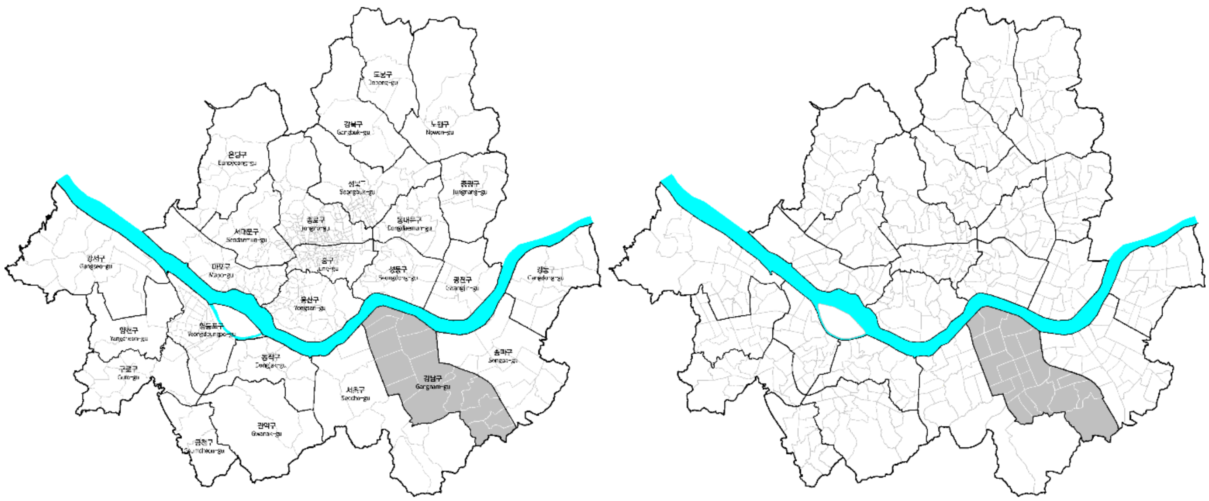
Furthermore, I examine the link between (popular) culture, tourism and housing markets in a quasi-experimental setting. How much a place is perceived as a tourist magnet usually depends on its place in history and its role in arts and culture. In other words, such attractiveness usually evolves slowly. Gangnam, in contrast, became a household name almost overnight. After the song went viral, more domestic and international tourists began visiting the district, which benefited the local hotel and restaurant businesses and hence the housing market. This leads to a novel argument in favor of investment in culture. If successful, a strengthening of cultural industries can attract international tourists, which directly benefits local service industries and property owners. In fact, I show that these benefits are not short-lived. The effect of "Gangnam Style" has persisted for more than ten years since the song's release.

Finally, I hope that my research produces additional insights into the housing markets of megacities. As I have seen, the real estate market in Seoul is characterized by the availability of excellent and safe public transportation as well as a great variety of nearby schools. Also, I have argued that residential property in Seoul is more or less homogeneous and that the market is generally very dynamic in nature. Only in a dynamic environment can prices and quantities respond to changes in expectations as quickly as they did in Gangnam in 2012.

2.7 Appendix

2.7.1 Additional Figures and Results

Figure 2.4: Seoul Metropolitan City



Note: These maps of Seoul show the city's 25 districts. Gangnam (in gray) is one of eleven districts located to the South of the Han River and is neighbored by Seocho and Songpa. The map on the left shows how each district is subdivided into legal-status dongs, of which Seoul has a total of 467. Gangnam has 14 legal-status dongs, Seocho and Songpa have 10 and 27 legal-status dongs, respectively. The map on the right shows how each district is subdivided into administrative dongs. In total there are 426 administrative dongs. Gangnam, Seocho, and Songpa have 22, 18, and 13 administrative dongs, respectively. Basis Geometries: The Seoul Research Data Service (http://data.si.re.kr)

Table 2.12: Average Square Meter Prices by Year

Year	Observations	Price per m^2	Year	Observations	Price per m^2
2009	143,352	534.2260	**2016**	209,129	605.4355
		(292.5730)			(303.5486)
2010	93,990	519.3619	**2017**	191,311	687.1998
		(278.3937)			(363.5468)
2011	107,888	508.3630	**2018**	157,923	718.2322
		(255.3693)			(407.9522)
2012	80,679	503.0488	**2019**	141,177	865.3173
		(255.5238)			(519.9097)
2013	113,674	511.8982	**2020**	175,661	894.1357
		(246.3851)			(537.2262)
2014	143,770	533.0107	**2021**	138,309	968.0319
		(256.8265)			(613.5148)
2015	213,090	553.5513	**2022**	219,929	948.2976
		(267.9335)			(593.2110)

Note: Standard deviations in parentheses.

Table 2.13: Median Residential Property Size Transacted by Year and Region

Year	Housing type	Gangnam	Seoul Metro-politan City*	Seoul to the South of the Han River*	Seocho & Songpa
2009	All Properties	62.46	60.00	64.08	75.29
	Observations	9,018	134,334	66,443	16,925
	Apartments	79.97	82.49	83.81	84.75
	Observations	6,551	72,249	38,929	11,523
2010	All Properties	59.82	59.95	59.99	74.30
	Observations	6,070	87,920	42,420	10,298
	Apartments	79.80	82.94	83.67	84.82
	Observations	3,942	43,297	21,907	6,629
2011	All Properties	59.96	59.97	60.74	71.20
	Observations	5,924	101,964	47,914	10,386
	Apartments	81.07	79.70	83.15	84.79
	Observations	3,962	53,260	26,046	6,535
2012	All Properties	59.97	59.94	59.98	71.49
	Observations	4,054	76,625	35,088	8,131
	Apartments	79.61	80.64	83.06	84.75
	Observations	2,753	40,389	19,664	5,472
2013	All Properties	78.05	60.41	64.99	80.27
	Observations	6,044	107,630	49,772	11,802
	Apartments	84.43	83.34	84.28	84.80
	Observations	4,788	66,077	32,152	8,718
2014	All Properties	79.25	63.24	66.72	82.61
	Observations	8,300	135,470	62,985	14,814
	Apartments	84.73	83.86	84.33	84.87
	Observations	6,194	84,136	40,752	10,697
2015	All Properties	74.40	64.53	66.69	79.47
	Observations	11,140	201,950	96,164	20,513
	Apartments	84.48	84.17	84.36	84.82
	Observations	7,662	119,608	58,446	13,772
2016	All Properties	76.79	63.49	66.41	79.70
	Observations	10,339	198,790	94,894	21,740
	Apartments	84.72	84.42	84.61	84.89
	Observations	7,092	113,176	55,304	14,035
2017	All Properties	80.67	62.37	66.63	81.23
	Observations	10,606	180,705	86,688	21,201
	Apartments	84.84	84.39	84.60	84.88
	Observations	7,589	106,169	52,056	14,244
2018	All Properties	72.96	59.98	59.99	67.95
	Observations	6,066	151,857	69,370	14,192
	Apartments	84.90	84.30	84.48	84.90
	Observations	3,543	80,658	36,572	7,953
2019	All Properties	76.79	59.93	59.96	75.71
	Observations	6,889	134,288	64,207	14,734
	Apartments	84.85	82.77	84.34	84.85
	Observations	4,654	72,541	34,827	9,192
2020	All Properties	59.96	59.64	59.57	59.87
	Observations	6,725	168,936	80,323	17,005
	Apartments	84.52	78.60	82.79	84.79
	Observations	3,779	81,666	38,103	8,416

Note: Columns with an asterisk (*) present **median** values excluding Gangnam. I also conducted 72 **mean** difference tests as follows. For each year I carried out t-tests between the mean square footage in the treated district of Gangnam versus the controlled districts in the Metropolitan City of Seoul. Then I did the same with Gangnam versus the other districts to the South of the Han River (excluding Gangnam). Lastly, Gangnam versus Seocho and Songpa. These tests were run for all housing types combined but also for apartments only. The results show that **all** mean differences lie within the 95 percent confidence intervals.

Table 2.14: Number of Apartments Per Km^2

Year	**Seoul Metropolitan City**	**Seoul to the South of the Han River**	**Gangnam, Seocho, and Songpa**
2010	3449.1740 (3269.0640)	3602.8440 (2802.9050)	3774.2120 (2447.3450)
Obs	419	200	65
2015	3756.7930 (3399.0230)	3781.5470 (2862.8600)	3717.8280 (2403.0510)
Obs	422	202	67
2020	4105.3570 (3500.0980)	4069.0450 (2956.1400)	3838.7130 (2391.3920)
Obs	422	202	67

Note: Housing stock data is publicly available at: https://data.seoul.go.kr/dataList/10585/S/2/datasetView.do. During the period of interest, Data Seoul published information about the housing stock at three points that are equidistant in time (2010, 2015, and 2020). The underlying geographical areas, however, are administrative (i.e. not legal-status) dongs. The tabulated values result from dividing the total number of apartments in an administrative dong by the administrative dong's area in square kilometers. Standard deviations are in parentheses.

Table 2.15: Annual Growth Rates of Numbers of Apartments Per Km^2

	Seoul Metropolitan City	Seoul to the South of the Han River	Gangnam, Seocho and Songpa
Mean	0.0259***	0.0177*	0.0119
	(0.0081)	(0.0094)	(0.0163)
P>|t|	0.001	0.061	0.465
Median	0.0178*	0.0136	-0.0010
	(0.0093)	(0.0120)	(0.0228)
P>|t|	0.057	0.259	0.967
Obs	1,263	604	199

Note: Using the data summarized in Table 2.14, each entry in this table presents the estimated coefficient when regressing the logarithm of an administrative dong's number of apartments per square kilometer on year. Mean (Median) refers to results obtained from applying the ordinary least squares (least absolute deviations) method. Standard deviations are in parentheses.

Table 2.16: Housing Price Regression Coefficients Related to Subway Measures

	Model 1	Model 2	Model 3	Model 4
Subways	9.0583***	-12.3787		
	(2.1108)	(7.9251)		
P>|t|	0.000	0.119		
Line 1			-23.4075***	-64.2167
			(6.7215)	(49.1071)
P>|t|			0.001	0.192
Line 2			52.4026***	-219.7970***
			(15.5512)	(73.3738)
P>|t|			0.001	0.003
Line 3			-7.6565	32.3376
			(18.9155)	(40.8807)
P>|t|			0.686	0.429
Line 4			-33.4110**	447.8528***
			(14.8910)	(46.5599)
P>|t|			0.025	0.000
Line 5			-17.2971	141.0726***
			(12.4899)	(48.4895)
P>|t|			0.167	0.004
Line 6			-20.1924**	-190.5405***
			(7.9890)	(29.0588)
P>|t|			0.012	0.000
Line 7			-27.7988*	-176.6421**
			(14.8771)	(83.2753)
P>|t|			0.063	0.035
Line 8			5.0253	25.5932
			(11.7946)	(68.7984)
P>|t|			0.670	0.710
Line 9			158.8678***	84.5624***
			(18.0522)	(27.7979)
P>|t|			0.000	0.003
Line 10			145.2224***	168.3939***
			(26.5385)	(32.4963)
P>|t|			0.000	0.000
Line 11			-217.5894***	64.4770***
			(45.2474)	(17.6981)
P>|t|			0.000	0.000
Line 12			-7.8913	-15.3000
			(25.3400)	(21.4677)
P>|t|			0.756	0.477
Line 13			47.2613**	43.9683*
			(19.0412)	(25.6800)
P>|t|			0.014	0.088
Line 14			-165.6307***	-177.6655***
			(22.9864)	(23.9581)
P>|t|			0.000	0.000
Line 15			-7.5093	-7.3927
			(42.2051)	(39.8133)
P>|t|			0.859	0.853
Line 16			-4.7697	-136.9383***
			(58.1087)	(52.5281)
P>|t|			0.935	0.010
Line 17			-34.7095*	-23.0795
			(20.6263)	(15.5890)
P>|t|			0.093	0.140

Note: This is the subway-related coefficients that are omitted in Table 2.5. Standard errors are in parentheses. *,**,*** denote statistical significance at the 10%, 5% and 1% levels, respectively. Line 10 denotes Bundang Line; Line 11, New Bundang Line; Line 12, Airport Line; Line 13, Gyeongui-Jungang Line; Line 14, Ui-Sinseol Line; Line 15, Sillim Line; Line 16, Gimpo-Gold Line; Line 17, Gyeongchun Line.

Table 2.17: Price Regression Results Using Distance from Gangnam District Office

	Model 1	**Model 2**	**Model 3**	**Model 4**
Seoul Metropolitan City from 2009 to 2020				
Square Meter Price	-14.0532***	-11.4802***	-13.5924***	-9.9876***
	(1.6224)	(1.7278)	(1.5105)	(1.6621)
P>|t|	0.000	0.000	0.000	0.000
Ln (Square Meter Price)	-0.0121***	-0.0070***	-0.0115***	-0.0058***
	(0.0015)	(0.0010)	(0.0013)	(0.0010)
P>|t|	0.000	0.000	0.000	0.000
Observations	1,714,881	1,714,881	1,714,881	1,714,881
Clusters (year-district)	Yes (300)	Yes (300)	Yes (300)	Yes (300)
Seoul to the South of the Han River from 2009 to 2020				
Square Meter Price	-15.1611***	-11.4588***	-16.5469***	-12.1191***
	(2.1055)	(2.1527)	(2.1810)	(2.3525)
P>|t|	0.000	0.000	0.000	0.000
Ln (Square Meter Price)	-0.0120***	-0.0064***	-0.0127***	-0.0062***
	(0.0018)	(0.0011)	(0.0017)	(0.0012)
P>|t|	0.000	0.000	0.000	0.000
Observations	864,630	864,630	864,630	864,630
Clusters (year-district)	Yes (132)	Yes (132)	Yes (132)	Yes (132)
Gangnam, Seocho and Songpa from 2009 to 2020				
Square Meter Price	-43.8499***	-34.7362***	-31.8134***	-38.7573***
	(5.3769)	(6.5837)	(3.4236)	(7.2560)
P>|t|	0.000	0.000	0.000	0.000
Ln (Square Meter Price)	-0.0410***	-0.0236***	-0.0274***	-0.0245***
	(0.0041)	(0.0041)	(0.0025)	(0.0041)
P>|t|	0.000	0.000	0.000	0.000
Observations	264,708	264,708	264,708	264,708
Clusters (year-district)	Yes (36)	Yes (36)	Yes (36)	Yes (36)
Year Fixed Effects	Yes	Yes	Yes	Yes
District Fixed Effects	Yes	Yes	Yes	Yes
Dong Fixed Effects	No	Yes	No	Yes

Note: Each entry consists of a result obtained from estimating a modification of equation (2.1) where the DID coefficient is replaced by a distance measure to Gangnam (pre-treatment distances are set at equal to zero). The other control variables are the same as in the regressions displayed in Tables 2.5 and 2.6. For simplicity and brevity, only the estimated distance coefficients are tabulated. Standard errors are in parentheses. *,**,*** denote statistical significance at the 10%, 5% and 1% levels, respectively.

Table 2.18: Transaction Regression Results Using Distance from Gangnam District Office

	All Housing Types		Only Apartments	
	Model 5	**Model 6**	**Model 7**	**Model 8**
	Seoul Metropolitan City from 2009 to 2020			
Transactions	0.3514	0.2831	0.9088	0.6830
	(0.3542)	(0.3688)	(1.0820)	(1.1193)
P>|t|	0.322	0.443	0.402	0.542
Observations	14,046	14,046	3,695	3,695
Clusters (year-district)	Yes (300)	Yes (300)	Yes (300)	Yes (300)
	Seoul to the South of the Han River from 2009 to 2020			
Transactions	0.9239**	0.7971*	4.2958***	4.1864***
	(0.3873)	(0.4156)	(1.1132)	(1.1873)
P>|t|	0.018	0.057	0.000	0.001
Observations	4,573	4,573	1,279	1,279
Clusters (year-district)	Yes (132)	Yes (132)	Yes (132)	Yes (132)
	Gangnam, Seocho and Songpa from 2009 to 2020			
Transactions	2.7227**	2.8611**	6.5865***	6.8669**
	(1.2788)	(1.3185)	(2.1937)	(2.5825)
P>|t|	0.040	0.037	0.005	0.012
Observations	1,414	1,414	398	398
Clusters (year-district)	Yes (36)	Yes (36)	Yes (36)	Yes (36)
Year Fixed Effects	Yes	Yes	Yes	Yes
District Fixed Effects	Yes	Yes	Yes	Yes
Dong Fixed Effects	Yes	Yes	Yes	Yes

Note: Each entry consists of a result obtained from estimating a modification of equation (2.2) where the DID coefficient is replaced by a distance measure to Gangnam (pre-treatment distances are set at equal to zero). The other control variables are the same as in the regressions displayed in Tables 2.8 and 2.9. For simplicity and brevity, only the estimated distance coefficients are tabulated. Standard errors are in parentheses. *,**,*** denote statistical significance at the 10%, 5% and 1% levels, respectively.

Table 2.19: Hotel Rooms Regression Results Using Distance from Gangnam District Office

	Model 9	**Model 10**	**Model 11**	**Model 12**
	Seoul Metropolitan City from 2009 to 2021			
All Hotel Rooms	-72.6921***	-100.5415***	-72.6921	-100.5415
	(27.2510)	(26.9264)	(49.5896)	(68.1929)
	0.008	0.000	0.156	0.153
Three Star Hotel Rooms	-32.2392***	-40.2894***	-32.2392	-40.2894
	(11.2006)	(11.1617)	(20.0402)	(27.1319)
	0.004	0.000	0.121	0.151
Observations	325	325	325	325
Clusters (District)	No	No	Yes (25)	Yes (25)
	Seoul to the South of the Han River from 2009 to 2021			
All Hotel Rooms	-35.8736*	-4.6661	-35.8736	-4.6661
	(18.1707)	(20.4978)	(54.3958)	(28.5148)
	0.051	0.820	0.524	0.873
Three Star Hotel Rooms	-28.8720***	-13.6767	-28.8720	-13.6767
	(8.8702)	(10.4772)	(21.3434)	(12.7517)
	0.001	0.194	0.206	0.309
Observations	143	143	143	143
Clusters (District)	No	No	Yes (11)	Yes (11)
	Gangnam, Seocho and Songpa from 2009 to 2021			
All Hotel Rooms	-356.1606***	-439.1302***	-356.1606*	-439.1302***
	(99.1751)	(117.0990)	(91.7984)	(32.0347)
	0.002	0.001	0.060	0.005
Three Star Hotel Rooms	-170.6878***	-284.7312***	-170.6878**	-284.7312***
	(58.1922)	(68.8784)	(30.1479)	(21.6939)
	0.007	0.001	0.030	0.006
Observations	39	39	39	39
Clusters (District)	No	No	Yes (3)	Yes (3)
Year Fixed Effect	Yes	Yes	Yes	Yes
District Fixed Effects	Yes	Yes	Yes	Yes

Note: Each entry consists of a result obtained from estimating a modification of equation (2.3) where the DID coefficient is replaced by a distance measure to Gangnam (pre-treatment distances are set at equal to zero). The other control variables are the same as in the regressions reported in Tables 2.10 and 2.11. For simplicity and brevity, only the estimated distance coefficients are tabulated. Standard errors are in parentheses. *,**,*** denote statistical significance at the 10%, 5% and 1% levels, respectively.

Table 2.20: Price Regression Results Using By Year – By District Fixed Effects

	Model 13	**Model 14**
Panel A: Square Meter Prices		
DID	414.9019***	446.3312***
	(4.6995)	(8.9425)
P>|t|	0.000	0.000
Distance	-22.6190***	-11.6388***
	(3.3092)	(2.3871)
P>|t|	0.000	0.000
Panel B: Logarithm of Square Meter Prices		
DID	0.1040***	0.1076***
	(0.0042)	(0.0067)
P>|t|	0.000	0.000
Distance	-0.0285***	-0.0088***
	(0.0038)	(0.0022)
P>|t|	0.000	0.000
Observations	1,714,881	1,714,881
Clusters (year-district)	Yes (300)	Yes (300)
Year Fixed Effects	Yes	Yes
District Fixed Effects	Yes	Yes
Year-District Fixed Effects	Yes	Yes
Dong Fixed Effects	No	Yes
House Characteristics	Yes	Yes
Schools / Subways	No	No

Note: In its first and third rows this table presents DID coefficients, in its second and fourth rows it displays Distance coefficients. All coefficients are obtained from estimating the specifications of equation (2.1) when using square meter prices (Panel A) or the logarithm of square meter prices as the outcome variable (Panel B). The set of controls related to housing characteristics is the same as in Tables 2.5 and 2.6. Given the multitude of fixed effects, however, I do not control for neighborhood attributes (i.e. schools and subways). All eight regressions are based on observations from the Metropolitan City of Seoul between 2009 and 2020. Standard errors are in parentheses. *,**,*** denote statistical significance at the 10%, 5% and 1% levels, respectively.

Table 2.21: Price Regression Results Using Various Clustering Levels for Standard Errors

	Model 1	Model 2	Model 3	Model 4
	Seoul Metropolitan City from 2009 to 2022			
Square Meter Price	166.6478	212.3474	184.1798	167.9614
	(36.8587)***	(32.6727)***	(46.9915)***	(30.167)***
	[35.8027]***	[63.6137]***	[40.7516]***	[60.6522]**
Ln (Square Meter Price)	0.0327	0.0746	0.0452	0.0467
	(0.0404)	(0.0378) **	(0.0534)	(0.0344)
	[0.0201]	[0.0228]***	[0.0171]**	[0.0183]**
Observations	2,069,845	2,069,845	2,069,845	2,069,845
Clusters (Dong Level)	(434)	(434)	(434)	(434)
Clusters (Two-Way)	[14 Years & 25 Dist]	[14 Years & 25 Dist]	[14 Years & 25 Dist]	[14 Years & 25 Dist]
	Seoul Metropolitan City from 2009 to 2020			
Square Meter Price	152.8086	175.2363	154.4379	139.7650
	(27.3806)***	(24.5049)***	(39.1913)***	(22.4582)***
	[41.7395]***	[70.9873]**	[42.8476]***	[65.8432]*
Ln (Square Meter Price)	0.0608	0.0855	0.0606	0.0549
	(0.0355) *	(0.0345) **	(0.0473)	(0.0293) **
	[0.0126]***	[0.0221]***	[0.0167]***	[0.0184]**
Observations	1,714,881	1,714,881	1,714,881	1,714,881
Clusters (Dong Level)	(434)	(434)	(434)	(434)
Clusters (Two-Way)	[12 Years & 25 Dist]	[12 Years & 25 Dist]	[12 Years & 25 Dist]	[12 Years & 25 Dist]
	Seoul to the South of the Han River from 2009 to 2020			
Square Meter Price	140.2927	140.6592	154.4719	150.7932
	(26.8970)***	(25.1829)***	(42.0304)***	(25.1010)***
	[39.1162]***	[57.2567]**	[43.3134]***	[67.3657]**
Ln (Square Meter Price)	0.0499	0.0695	0.0679	0.0657
	(0.0346)	(0.0323) **	(0.0518)	(0.0329) **
	[0.0158]***	[0.0241]**	[0.0225]**	[0.0222]**
Observations	864,630	864,630	864,630	864,630
Clusters (Dong Level)	(117)	(117)	(117)	(117)
Clusters (Two Way)	[12 Years & 11 Dist]	[12 Years & 11 Dist]	[12 Years & 11 Dist]	[12 Years & 11 Dist]
	Gangnam, Seocho and Songpa from 2009 to 2020			
Square Meter Price	75.0160	89.0744	79.0221	87.1756
	(35.7679)**	(36.2049)**	(46.0785)*	(39.9498)**
	[55.7351]	[62.8419]	[70.3594]	[64.2172]
Ln (Square Meter Price)	0.0455	0.0492	0.0528	0.0488
	(0.0365)	(0.0358)	(0.0534)	(0.0369)
	[0.0413]	[0.0435]	[0.0578]	[0.0375]
Observations	264,708	264,708	264,708	264,708
Clusters (Dong Level)	(37)	(37)	(37)	(37)
Clusters (Two Way)	[12 Years & 3 Dist]	[12 Years & 3 Dist]	[12 Years & 3 Dist]	[12 Years & 3 Dist]
Year Fixed Effects	Yes	Yes	Yes	Yes
District Fixed Effects	Yes	Yes	Yes	Yes
Dong Fixed Effects	No	Yes	No	Yes

Note: Since the underling models are unchanged, each entry shows the same point estimates for the DID coefficients as are displayed in Tables 2.5 to 2.7. Standard errors using dong-level clustering are in parentheses; standard errors using two-way (year and districts) clustering are in brackets. *,**,*** denote statistical significance at the 10%, 5% and 1% levels, respectively.

Table 2.22: Transaction Regression Results Using Various Clustering Levels for Standard Errors

	All Housing Types		Only Apartments	
	Model 5	**Model 6**	**Model 7**	**Model 8**
	Seoul Metropolitan City from 2009 to 2022			
Transactions	-19.6130	-18.8039	-58.6118	-51.1292
	(4.4146)***	(4.8066)***	(10.2787)***	(11.0602)***
P>|t|	0.000	0.000	0.000	0.000
	[6.9677]**	[7.0764]**	[16.0917]***	[14.4252]***
P>|t|	0.015	0.020	0.003	0.004
Observations	16,391	16,391	4,309	4,309
Clusters (Dong Level)	Yes (434)	Yes (434)	Yes (336)	Yes (336)
Clusters (Two Way)	[14 Years & 25 Dist]	[14 Years & 25 Dist]	[14 Years & 25 Dist]	[14 Years & 25 Dist]
	Seoul Metropolitan City from 2009 to 2020			
Transactions	-16.7804	-16.3469	-59.3909	-56.6538
	(4.0377)***	(3.9999)***	(10.8043)***	(10.5779)***
P>|t|	0.000	0.000	0.000	0.000
	[7.2113]**	[6.6279]**	[17.2588]***	[14.3521]***
P>|t|	0.040	0.031	0.006	0.002
Observations	14,046	14,046	3,695	3,695
Clusters (Dong Level)	Yes (434)	Yes (434)	Yes (335)	Yes (335)
Clusters (Two Way)	[12 Years & 25 Dist]	[12 Years & 25 Dist]	[12 Years & 25 Dist]	[12 Years & 25 Dist]
	Seoul to the South of the Han River from 2009 to 2020			
Transactions	-18.9771	-19.1710	-51.7026	-52.1717
	(5.2956)***	(5.4435)***	(11.9765)***	(12.6402)***
P>|t|	0.000	0.001	0.000	0.000
	[6.1698]**	[6.3379]**	[16.4607]***	[17.8810]**
P>|t|	0.012	0.013	0.010	0.015
Observations	4,573	4,573	1,279	1,279
Clusters (Dong Level)	Yes (117)	Yes (117)	Yes (110)	Yes (110)
Clusters (Two Way)	[12 Years & 11 Dist]	[12 Years & 11 Dist]	[12 Years & 11 Dist]	[12 Years & 11 Dist]
	Gangnam, Seocho and Songpa from 2009 to 2020			
Transactions	-7.7696	-8.5719	-14.2266	-13.9580
	(5.0193)	(5.2934)	(11.4151)	(12.2092)
P>|t|	0.130	0.114	0.221	0.261
	[11.4575]	[11.9342]	[18.3207]	[18.0443]
P>|t|	0.568	0.547	0.519	0.520
Observations	1,414	1,414	398	398
Clusters (Dong Level)	Yes (37)	Yes (37)	Yes (35)	Yes (35)
Clusters (Two Way)	[12 Years & 3 Dist]	[12 Years & 3 Dist]	[12 Years & 3 Dist]	[12 Years & 3 Dist]
Year Fixed Effects	Yes	Yes	Yes	Yes
District Fixed Effects	Yes	Yes	Yes	Yes
Dong Fixed Effects	No	Yes	No	Yes

Note: Standard errors using dong-level clustering are in parentheses; standard errors using two-way (year and district) clustering are in brackets. *,**,*** denote statistical significance at the 10%, 5% and 1% levels, respectively.

Table 2.23: Hotel Room Regression Results Using Two-Way Clustering for Standard Errors

	Model 11	Model 12
Seoul Metropolitan City from 2009 to 2021		
All Hotel Rooms	1537.1280***	1567.3140**
	[369.2888]	[517.9114]
P>|t|	0.001	0.011
Three Star Hotel Rooms	821.2875***	782.8063***
	[197.6210]	[205.7560]
P>|t|	0.001	0.003
Observations	325	325
Clusters (Two Way)	[13 Years & 25 Dist]	[13 Years & 25 Dist]
Seoul to the South of the Han River from 2009 to 2021		
All Hotel Rooms	1787.7870***	1022.3050*
	[317.5755]	[510.1121]
P>|t|	0.000	0.073
Three Star Hotel Rooms	898.9833***	572.1942*
	[165.2074]	[302.8051]
P>|t|	0.000	0.088
Observations	143	143
Clusters (Two Way)	[13 Years & 11 Dist]	[13 Years & 11 Dist]
Gangnam, Seocho and Songpa from 2009 to 2021		
All Hotel Rooms	1756.6000**	1938.8510**
	[395.4862]	[362.7973]
P>|t|	0.047	0.033
Three Star Hotel Rooms	821.3167*	1240.8360**
	[215.9193]	[226.1107]
P>|t|	0.063	0.032
Observations	39	39
Clusters (Two Way)	[13 Years & 3 Dist]	[13 Years & 3 Dist]
Year Fixed Effects	Yes	Yes
District Fixed Effects	Yes	Yes

Note: Standard errors using two-way (year and district) clustering are in brackets. *,**,*** denote statistical significance at the 10%, 5% and 1% levels, respectively.

Table 2.24: Price and Transactions Regressions Using Housing Types

Seoul Metropolitan City from 2009 to 2022					
Square Meter Price	**Interactions**	**Model 1**	**Model 2**	**Model 3**	**Model 4**
	DID x Apartments	323.8568***	329.9733***	319.3838***	283.6869***
		(80.4440)	(82.3424)	(80.5400)	(80.6842)
	P>|t|	0.000	0.000	0.000	0.000
	DID x Multi-Hh House	17.9489	52.4694	32.3515	-4.4403
		(44.0805)	(43.8390)	(53.4448)	(40.3730)
	P>|t|	0.684	0.232	0.545	0.912
	DID x Regular House	61.6525	160.2728**	76.6437	101.5437
		(78.6095)	(70.2087)	(81.9872)	(65.2163)
	P>|t|	0.433	0.023	0.351	0.120
	DID x Dual Purpose Bld	-178.7029***	-50.1698***	-108.1807***	-103.6703***
		(26.0346)	(19.2085)	(31.7775)	(20.0175)
	P>|t|	0.000	0.009	0.001	0.000
	Observations	2,069,845	2,069,845	2,069,845	2,069,845
	Cluster (year-district)	Yes (350)	Yes (350)	Yes (350)	Yes (350)
Ln (Square Meter Price)	**Interactions**	**Model 1**	**Model 2**	**Model 3**	**Model 4**
	DID x Apartments	0.1336***	0.1391***	0.1295***	0.1106***
		(0.0211)	(0.0219)	(0.0222)	(0.0211)
	P>|t|	0.000	0.000	0.000	0.000
	DID x Multi-Hh House	-0.0013	0.0259	0.0031	-0.0101
		(0.0197)	(0.0207)	(0.0289)	(0.0190)
	P>|t|	0.948	0.211	0.915	0.596
	DID x Regular House	-0.0293	0.0583*	-0.0245	0.0212
		(0.0386)	(0.0350)	(0.0436)	(0.0329)
	P>|t|	0.448	0.097	0.575	0.520
	DID x Dual Purpose Bld	-0.2490***	-0.1154***	-0.1860***	-0.1477***
		(0.0392)	(0.0342)	(0.0397)	(0.0335)
	P>|t|	0.000	0.001	0.000	0.000
	Observations	2,069,845	2,069,845	2,069,845	2,069,845
	Cluster (year-district)	Yes (350)	Yes (350)	Yes (350)	Yes (350)
Transactions	**Interactions**	**Model 5**	**Model 6**	**Model 7**	**Model 8**
	DID x Apartments	-18.6961	-17.6289		
		(16.7533)	(17.2263)		
	P>|t|	0.265	0.307		
	DID x Multi-Hh House	-26.0500***	-25.3762***		
		(6.4874)	(6.8838)		
	P>|t|	0.000	0.000		
	DID x Regular House	-1.6893	-1.0049		
		(8.0790)	(8.5636)		
	P>|t|	0.834	0.907		
	DID x Dual Purpose Bld	-35.6269***	-34.9672***		
		(10.7779)	(10.7777)		
	P>|t|	0.001	0.001		
	Observations	16,391	16,391		
	Cluster (year-district)	Yes (350)	Yes (350)		

Note: Each entry consists of a result obtained from estimating modifications of equations (2.1) for prices and (2.2) for transactions using the (Year×Dong×HousingType) panel. The control variables are the same as in the regressions shown in Tables 2.5, 2.6 and 2.8. For simplicity and brevity, however, only the estimated DID coefficients are tabulated. Standard errors are in parentheses. *,**,*** denote statistical significance at the 10%, 5% and 1% levels, respectively.

Table 2.25: Restaurant Regression Results Using Various Control Regions

	Model 9	**Model 10**	**Model 11**	**Model 12**
	Seoul Metropolitan City from 2009 to 2022			
Restaurants	87.0453***	87.0453***	74.2391***	69.8531***
	(13.0674)	(13.5358)	(13.4138)	(12.9980)
P>|t|	0.000	0.000	0.000	0.000
Observations	6,538	6,538	6,538	6,538
Clusters (year-district)	Yes (350)	Yes (350)	Yes (350)	Yes (350)
	Seoul Metropolitan City from 2009 to 2020			
Restaurants	77.3338***	77.3338***	63.1235***	60.8489***
	(12.2383)	(12.7552)	(12.5950)	(12.2371)
P>|t|	0.000	0.000	0.000	0.000
Observations	5,604	5,604	5,604	5,604
Clusters (year-district)	Yes (300)	Yes (300)	Yes (300)	Yes (300)
	Seoul to the South of the Han River from 2009 to 2020			
Restaurants	84.7011***	84.7011***	65.1256***	74.0970***
	(12.4653)	(12.9747)	(13.3251)	(12.0086)
P>|t|	0.000	0.000	0.000	0.000
Observations	1,452	1,452	1,452	1,452
Clusters (year-district)	Yes (132)	Yes (132)	Yes (132)	Yes (132)
	Gangnam, Seocho and Songpa from 2009 to 2020			
Restaurants	71.1505***	71.1505***	67.1267***	70.4352***
	(12.4368)	(12.9611)	(14.5727)	(11.5681)
P>|t|	0.000	0.000	0.000	0.000
Observations	444	444	444	444
Clusters (year-district)	Yes (36)	Yes (36)	Yes (36)	Yes (36)
Year Fixed Effects	Yes	Yes	Yes	Yes
District Fixed Effects	Yes	Yes	Yes	Yes
Dong Fixed Effects	No	Yes	Yes	Yes

Note: This table presents DID coefficients that are obtained from estimating the specifications of equation (2.2) with two modifications. First, there are no controls related to house characteristics. Second, the outcome variable is the number of restaurants per legal-status dong. The various columns report regression results after adding more and more controls. In Model 14 I include legal-status dong fixed effects. Model 15 includes school information and the Subways count variable. In Model 16 I replace the Subways count variable with the 17 subway line dummies. Standard errors are in parentheses. *,**,*** denote statistical significance at the 10%, 5% and 1% levels, respectively.

Table 2.26: Price Regression Results after Controlling for Transactions (Seoul Metropolitan City)

Variables	Treatment	Model 1	Model 2	Model 3	Model 4
Square Meter Price	DID	163.5844***	207.1534***	181.6083***	163.3801***
		(40.6836)	(48.5724)	(44.3286)	(48.0375)
	P>|t|	0.000	0.000	0.000	0.001
Observations		2,069,845	2,069,845	2,069,845	2,069,845
Clusters (year-district)		Yes (350)	Yes (350)	Yes (350)	Yes (350)
Ln (Square Meter Price)	DID	0.0315	0.0727***	0.0446**	0.0451***
		(0.0236)	(0.0159)	(0.0207)	(0.0151)
	P>|t|	0.183	0.000	0.032	0.003
Observations		2,069,845	2,069,845	2,069,845	2,069,845
Clusters (year-district)		Yes (350)	Yes (350)	Yes (350)	Yes (350)
Square Meter Price	DID x Apartment	321.6950***	325.5377***	317.1180***	279.8162***
		(79.7394)	(81.1569)	(79.9280)	(79.4968)
	P>|t|	0.000	0.000	0.000	0.000
	DID x Multi-Hh House	10.9179	39.2741	27.5914	-16.9447
		(43.1525)	(41.8835)	(52.4670)	(38.2647)
	P>|t|	0.800	0.349	0.599	0.658
	DID x Regular House	58.8397	156.2612**	74.9628	98.0272
		(76.9415)	(66.9868)	(80.4401)	(61.8744)
	P>|t|	0.445	0.020	0.352	0.114
	DID x Dual Purpose Bld	-181.5421***	-49.6239***	-110.1224***	-102.6774***
		(25.3137)	(18.2301)	(31.4322)	(18.8282)
	P>|t|	0.000	0.007	0.001	0.000
Observations		2,069,845	2,069,845	2,069,845	2,069,845
Clusters (year-district)		Yes (350)	Yes (350)	Yes (350)	Yes (350)
Ln (Square Meter Price)	DID x Apartment	0.1327***	0.1375***	0.1289***	0.1092***
		(0.0210)	(0.0218)	(0.0222)	(0.0210)
	P>|t|	0.000	0.000	0.000	0.000
	DID x Multi-Hh House	-0.0041	0.0212	0.0019	-0.0143
		(0.0194)	(0.0202)	(0.0288)	(0.0185)
	P>|t|	0.831	0.296	0.947	0.439
	DID x Regular House	-0.0305	0.0569*	-0.0249	0.0200
		(0.0381)	(0.0342)	(0.0433)	(0.0320)
	P>|t|	0.424	0.097	0.566	0.533
	DID x Dual Purpose Bld	-0.2502***	-0.1152***	-0.1865***	-0.1474***
		(0.0388)	(0.0336)	(0.0395)	(0.0330)
	P>|t|	0.000	0.001	0.000	0.000
Observations		2,069,845	2,069,845	2,069,845	2,069,845
Clusters (year-district)		Yes (350)	Yes (350)	Yes (350)	Yes (350)

Note: Each entry consists in a result obtained from estimating modifications of equation (2.1). The control variables are the same as in the regressions reported in Tables 2.5, 2.6 and 2.24 with the addition of the numbers of transactions (by housing type). For simplicity and brevity, however, only the estimated DID coefficients are tabulated. Standard errors are in parentheses. *,**,*** denote statistical significance at the 10%, 5% and 1% levels, respectively.

CHAPTER 3

How Are They Doing?

The Academic Performance and Mental Wellbeing of World Cup Babies

This essay was previously published by *SSM-Population Health* (Bethmann and Cho, 2024).

3.1 Introduction

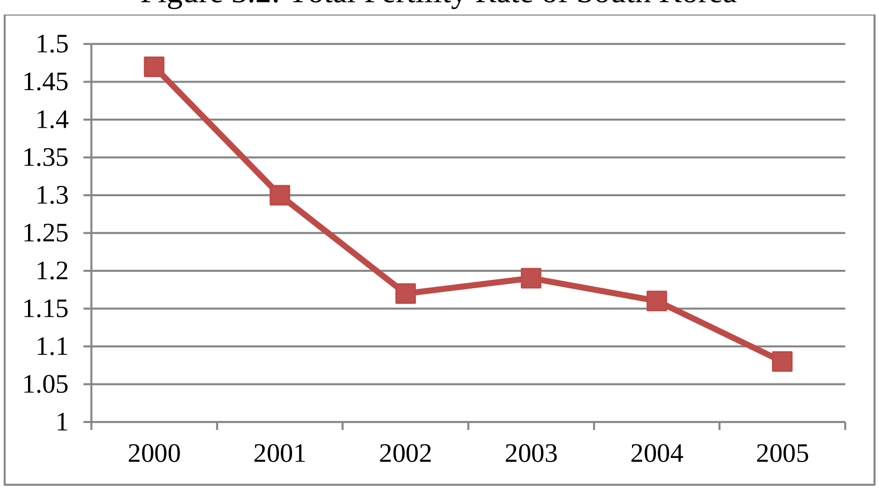
In 2002 between the 31st of May and the 30th of June, South Korea and Japan jointly hosted the 17th FIFA World Cup. Although FIFA ranked the Korean National Football Team only 43rd in its 2001 edition of the men's world ranking, the team accelerated to the semi-finals.[37] The unexpected match results made Korean people gather in large numbers on the streets and in stadiums to cheer for their national team as shown in Figure 3.1. Conservative media viewed the excessive joy and general euphoria with skepticism. The main fear was that there could be adverse effects on labor productivity, public security, and mental health issues caused by the prolonged cheering events and mass gatherings (Dong-A Ilbo, 2002a; Dong-A Ilbo, 2002b; Yonhap News Agency, 2002). In the aftermath, some newspapers showed additional concerns due to the unexpected increase of pregnancies (Seoul Broadcasting Service, 2003; Chosun Ilbo, 2010; Seoul Broadcasting Service, 2010; The Korea Economic Daily, 2014).

[37] The South Korean national team played their first game on June 4th and won against Poland. On June 29th, the South Korean team was defeated by Turkey in the third-place play-off.

Figure 3.1: Mass Gathering in Seoul 2002[38]

In fact, South Korea experienced a temporary increase in the total fertility rate[39] the year following the 2002 FIFA World Cup. Although the total fertility rate had continuously decreased since the 1990s, the rate exceptionally increased in 2003 as shown in Figure 3.2 (Statistics Korea, 2006).

Figure 3.2: Total Fertility Rate of South Korea



As can be seen, the total fertility rate was 1.17 in 2002; then rose to 1.19 in 2003 before it fell back to 1.16 in 2004. In particular, the ratio of babies born in spring (roughly ten months after

[38] Credit: Ministry of Culture, Sports, and Tourism (2017)

[39] The total fertility rate (TFR) in South Korea is calculated as follows: TFR = ∑(ASFR)/1000 where ASFR refers to the age-specific fertility rate (Statistics Korea, 2023) and the summation extends over all age groups.

the June 2002 World Cup) relative to the January borns increased sharply in 2003.[40] Table 3.1 shows the ratios of March, April, and May borns compared to the January borns.

Table 3.1: Ratio of Spring Borns Compared to the January Borns by Years

Year	2000	2001	2002	2003	2004	2005
$\frac{\text{March borns}}{\text{January borns}}$	0.97	0.96	0.98	1.01	0.99	0.99
$\frac{\text{April borns}}{\text{January borns}}$	0.86	0.85	0.89	0.95	0.89	0.89
$\frac{\text{May borns}}{\text{January borns}}$	0.85	0.82	0.86	0.91	0.85	0.87

Note: The number of monthly new borns by year is provided by Statistics Korea (2021).

As the above table shows, the March/January ratio increased by 2~3 percentage points compared to 2002 and 2004. Similarly, the April/January ratio increased by 6 percentage points and the May/January ratio increased by 5~6 percentage points compared to 2002 and 2004. Several statistical tests indicate that the temporary increase in the spring of 2003 was indeed significant (see Table 3.16 in the appendix).

The primary goal of this paper is to use the 2002 World Cup induced upward blip in the Korean fertility rate as an experiment to check whether a quantity-quality trade-off in reproduction exists and how it affects the wellbeing of children born during this episode. From the standpoint of economic theory, the event with its overwhelming excitement and joy temporarily lowered the costs of pursuing a quantity-oriented reproductive strategy. For about one month rollicking parties distracted the Korean population from the worries of everyday life and anxieties about the future, which ultimately affected the fertility rate. Important for statistical identification, the way the event was received and celebrated by the Korean public was unforeseen. The Korean government, in particular, did not intend to affect fertility rates when it

[40] Because Korean schools officially start in early March, Korean parents wish to give birth in January or February to ensure that their children are relatively old at school based on the (false) belief in a relative age effect (Bethmann and Cho, 2021). As a result, January usually has the largest number of newly born babies compared to other months. For this reason, we use the number of January born children as a reference when reporting the births increases in the March, April and May of 2003 (cf. Table 3.1).

decided to host the 2002 World Cup. Even in hindsight, the whole episode seems unlikely considering the history of poor performances of the Korean National Football Team at tournaments preceding the 2002 World Cup.

The exogenous fertility shock may have affected Korean couples in diverse ways. Couples with pre-born children, for example, had a higher chance of conceiving an additional child thereby lowering human capital investments per child. Similarly, childless couples had higher chances of having an unexpected pregnancy that typically results in subpar parental investments (Gipson et al., 2008; Marston and Cleland, 2010; Cavalcanti et al., 2020). Although the two groups most likely differ with respect to average ages, marital statuses, stability and durations of the relationships, they may both share comparatively low parental expectations with respect to the academic performance of their newly conceived offspring. We investigate whether this hypothesis is indeed true by using academic performance (school test scores) as a measure of child quality and examine whether children born in the spring of 2003 ("World Cup children") do under-perform. Our regression results show that World Cup children tend to perform worse at school (using test scores in five major subject areas) but they also show higher degrees of mental wellbeing (showing less aggressive or depressive symptoms) than children born in different years and months.

3.2 Background

A growing body of literature shows that seemingly irrelevant events affect the decisions of agents. The weather and seasonal changes, for example, are shown to influence people's emotions and/or moods (Sanders and Brizzolara, 1982; Denissen et al., 2008). The induced changes of emotions and moods, in turn, affect individual decision making (Schwarz and Clore, 1983; Alengoz et al., 2017).

Several studies have a more specific focus on sports activities affecting people's decision process, especially in light of crime (Kalist and Lee, 2016; Munyo and Rossi, 2013; Rees and Schnepel, 2009) and domestic violence (Card and Dahl, 2011). Munyo and Rossi (2013) show that the frustration after a surprising loss in a soccer game leads to an increase in criminal activities (and the opposite effect after a surprising win). In general, growing evidence suggests that crime rates increase during and after professional sports events (Rees and Schnepel, 2009, for college football; Kalist and Lee, 2016, for the National Football League). Similarly, Card and Dahl (2011) show that domestic violence increased on Sundays during the professional football season in the US due to frustration after disappointing match results.

Some papers have checked whether major (professional) sports events have an effect on fertility. Montesinos et al. (2013) attribute the spike in the Catalan fertility rate approximately nine to ten months after the 2009 season of the UEFA Champions League to the spectacular win of FC Barcelona against Chelsea FC at the semifinal. Similarly, Bernardi and Cozzani (2021) find that unexpected losses of local teams lead to a small decrease in the number of births approximately nine to ten months later. They claim that unexpected losses have a greater effect on fertility than expected losses. Hayward and Rybińska (2017), in contrast, show that the United States has not experienced an increase in the fertility rate after the Super Bowl.

Despite the possible connections between major sports events and changes in fertility, no study is using the exogenous fertility shock originating from a major sports event to investigate the possible quantity-quality trade-off of children suggested by Becker (1960). The theory postulates a negative relationship between a family's number of children (quantity) and their outcomes (quality) (Becker and Lewis, 1973; Becker and Tomes, 1976). The latter dimension is typically proxied by educational achievements (Conley and Glauber, 2006; Glick et al., 2007; Lee, 2007; Rosenzweig and Zhang, 2009) or health outcomes of children (Glick et al., 2007;

Angrist et al., 2010; Millimet and Wang, 2011). Most of the empirical work relies on siblings and twins as the main source of variations to measure the quantity-quality effects on children (see, for example, Black et al., 2005).

The statistical identification of the quantity-quality trade-off may well be at risk if the blip in Korean fertility was mainly driven by families that tend to produce less educated and/or less healthy offspring. Young, poor, and/or unmarried mothers, in particular, could distort results (Shields and Hanneke, 2008; Maani and Kalb 2007; Blau, 1999; Ermisch and Francesconi, 2001; Case and Katz, 1991; Feinstein and Symons, 1999; Finer and Zolna, 2014; Font-Ribera et al., 2007; Henshaw, 1998). Such a selection bias could also stem from different maternal attitudes towards risky behavior. A greater tendency to consume alcohol during pregnancy, for example, is shown to harm children’s performance in school and also to adversely affect (mental) health measures (Nilsson, 2017; O'Connor et al., 2002; O'Connor and Paley, 2009). Smoking is another example of such risky parental behavior (Ekblad et al., 2010; Nigg and Breslau, 2007; Rahu et al 2010). Finally, children from less supportive, cold, and neglectful parents are more likely to exhibit mental health disorders (Repetti et al., 2002).

Our paper contributes to the existing literature by examining not only the fertility increase caused by the 2002 World Cup but also analyzing the effect this event had on the quality of children born approximately ten months after the World Cup season. Unlike the UEFA Champions League Final or the NFL Super Bowl, World Cup matches are played within the boundary of the hosting country for about one month: compared to the one-evening Champions League Final and Super Bowl events, the World Cup season is hence much longer.[41] Moreover, the boisterous sentiment during the 2002 World Cup gave ample opportunities for South Koreans

[41] The Korean National team played their first game on June 4th and managed to remain in the tournament until June 29th - the day the third place play-off was held.

to indulge in the joys of the moment. Not surprisingly, the effect of the 2002 World Cup on the South Korean fertility rate was stronger than comparable effects of the UEFA Champions League Final in Europe or the Super Bowl in the US. Figure 3.2 and Table 3.1 both show the resulting blip in Korean fertility.

Although several papers focus on the fertility increase after major sports events (Montesinos et al., 2013; Hayward and Rybińska, 2017; Bernardi and Cozzani, 2021), there have been no studies examining the possible quantity-quality trade-off of children using the increased fertility rate caused by major sports events. Our study therefore fills a gap in the existing literature by investigating the effect the exogenous fertility shock caused by the 2002 World Cup had on child quality outcomes in Korea.

In addition, we propose to expand the view beyond the actual trade-off between the quantity and quality of children and to add an analysis of the children's mental wellbeing. The existing literature mainly focuses on the human capital formation of children from the parents' perspective using measures of academic achievements (Conley and Glauber, 2006; Glick et al., 2007; Lee, 2007; Rosenzweig and Zhang, 2009) or physical health (Glick et al., 2007; Angrist et al., 2010; Millimet and Wang, 2011). By using student mental wellbeing as the dependent variable, our paper addresses the quality dimension also from the children's perspective. Several indicators of aggressive and depressive symptoms are used to complete the picture of how World Cup children fared.

3.3 Data and Methodology

Our study uses the first grade cohort (children born in 2003) and fourth grade cohort (children born in 2000) from the Korean Children and Youth Panel Survey (KCYPS)[42] which is conducted

[42] The dataset is publicly available, and its detailed description can be found at https://www.nypi.re.kr/archive/board?menuId=MENU00329 [last accessed December 10, 2022].

by the National Youth Policy Institute (NYPI) and administered by the Prime Minister's Office. The NYPI chose schools based on the size and population of South Korea's seventeen primary administrative districts. Using a proportional stratified sampling method the NYPI then randomly selected individual students. The dataset traces both the first and the fourth grade cohort from 2010 to 2016. In our analysis, we use the seventh grade for the academic and the sixth grade for the mental wellbeing outcomes as shown in Table 3.2.

Table 3.2: Data Structure

Dependent Variables	Grade	First Grade Cohort (Born in 2003)	Fourth Grade Cohort (Born in 2000)
Academic Outcomes	7th Grade	7th Wave (surveyed in 2016)	4th Wave (surveyed in 2013)
Mental Wellbeing	6th Grade	6th Wave (surveyed in 2015)	3rd Wave (surveyed in 2012)

Note: The two cohorts that make up the first and fourth graders progressed to the sixth grade in 2015 and 2012 (advancing to the seventh grade in 2016 and 2013). Hence, we employ KCYPS data collected during the years 2012, 2013, 2015, and 2016, which corresponds to the 3rd, 4th, 6th, and 7th waves.

While the survey contains detailed information about actual school test scores of seventh-graders, its fourth wave in 2013 does not provide wellbeing information of seventh grade students. Thus, we use the information from sixth-graders in our mental wellbeing analysis.

The following Table 3.3 summarizes our differences-in-differences (DID) research design. As can be seen, students born in the spring of 2003 constitute the treatment group in our analysis.

Table 3.3: Research Design

Control / Treatment	Born in 2000	Born in 2003	Note
Mar, Apr, May	Control	Treated	The dummy variable 'treated months' denotes March, April, and May borns.
Other Birth Months	Control	Control	

In our main analysis, we used the pooled ordinary least squares (OLS) regression model to estimate the following equation (3.1):

$$\mathrm{Y_{ist}} = \beta_0 + \beta_1(\mathrm{year_t}) + \beta_2(\text{treated months}_\mathrm{i}) + \theta(\mathrm{worldcup_{it}}) + \beta_3(\mathrm{covariates_{ist}}) + \gamma(\text{school location}_\mathrm{s}) + \varepsilon_{\mathrm{ist}} \qquad (3.1)$$

Note that equation (3.1) embodies the difference-in-differences (DID) design. The variable *year* indicates whether a student was born after the 2002 World Cup. To be precise, students born in the year 2000 (2003) are assigned zero (one).[43] The *treated months* variable identifies students born in March, April, or May. Our DID variable *worldcup* results from the interaction between *year* and *treated months*. Through this design, we can capture whether World Cup children perform worse in school tests than the control group. *Y* and *covariates* denote our dependent variable(s) and a vector of individual characteristics respectively. Fixed *school location* effects control for time-invariant observable and unobservable characteristics of school provinces (and alternatively school districts) that might influence the outcome variable. The error term is assumed to have the usual ideal properties. Finally, in addition to applying the pooled OLS regression model, we also used the fact that most outcome variables are categorical in nature and estimated equation (3.1) using the ordered probit regression model. Because this model is non-linear, the size of the estimated coefficient of the interaction term ($\mathrm{worldcup}_{it}$) does not directly depict the magnitude of the treatment effect; its sign, however, does coincide with the sign of the treatment effect (Puhani, 2012). Thus, the ordered probit model allows us to double-check the signs of the pooled OLS regression coefficients.

Since our analysis uses two different cohorts, we also conducted several robustness checks using placebo treated months, and clustered standard errors (with clustering on the school level and on the school district level). Using alternative treated months, our placebo tests should reveal whether observations in the control group exhibit statistically significant differences in

[43] We restrict our analysis to students who entered elementary school according to their legal school age. In other words, students who were enrolled either later or earlier (i.e. not following the legal school age possibly due to mental or physical disabilities or talents) are not included. We dropped 80 students among 4,221 survey participants (about 1.89 percent).

school test scores.[44] The clustered standard errors may add to the precision of the regression results. Because school characteristics (difficulty of exams, quality of teachers, and location of the school) are not changing drastically over three years, possible inconsistency issues caused by using two different cohorts maybe partially alleviated by clustering standard errors on the school level or school district level.

In our academic outcome (i.e. child quality) analysis, we use school test scores in the following five subjects as our dependent variables: mathematics (*math*), social science (*sosci*), natural science (*nasci*), *korean*, and *english*. Table 3.4 summarizes the learning results after pooling the two cohorts.

Table 3.4: Descriptive Statistics for Academic Outcome Variables

Variables	Obs	Mean (sd)	Min	Max	Description
math	3,997	4.347511 (2.541759)	1	8	mathematics test score
sosci	3,864	4.461957 (2.334915)	1	8	social science test score
nasci	3,995	4.254568 (2.369694)	1	8	natural science test score
korean	3,993	4.722014 (2.218604)	1	8	Korean language test score
english	3,997	4.817113 (2.529570)	1	8	English language test score

Note: Test score is categorized as follows: 1 = 64 or less; 2 = 65 to 69; 3 = 70 to 74; 4 = 75 to 79; 5 = 80 to 84; 6 = 85 to 89; 7 = 90 to 95; 8 = 96 or above. Standard deviations are in parentheses.

For our mental wellbeing analyses, we focus on the following five dependent variables: *depressed* denotes "I feel miserable and depressed"; *suicidal*, "I want to die"; *self-reproach*, "Bad things happened by me"; *bullying*, "How many times I have bullied other students"; *violent*, "How many times I have hit others (very hard)". The first three variables *depressed*, *suicidal*, *and self-reproach* are categorized from 1 (strong yes) to 4 (strong no). Variables *bullying* and

[44] The placebo tests check whether the students in the control group do or do not exhibit statistically significant differences in their academic or mental wellbeing outcomes. When conducting these tests, we drop the treated observations from the main regressions.

violent, in contrast, are count variables. As mentioned before, we use the responses from sixth-graders as we do not have mental wellbeing information about seventh grade students. See Table 3.5 for a detailed definition and descriptive statistics.

Table 3.5: Descriptive Statistics for Mental Wellbeing Variables

Variables	Obs	Mean (sd)	Min	Max	Description
depressed	4,205	3.342212 (0.769383)	1	4	I feel miserable and depressed 1 = strong yes 2 = yes 3 = no 4 = strong no
suicidal	4,205	3.558859 (0.696375)	1	4	I want to die 1 = strong yes 2 = yes 3 = no 4 = strong no
self-reproach	4,205	3.132461 (0.840124)	1	4	Bad things happened by me 1 = strong yes 2 = yes 3 = no 4 = strong no
bullying	4,205	0.047562 (0.378567)	0	10	How many times I have bullied other students (did not report bullying = 0)
violence	4,205	0.035434 (0.472286)	0	15	How many times I have hit others very hard (did not report violence = 0)

Note: Standard deviations are in parentheses.

The individual covariates controlled in the regressions are mother's education, log of annual household income, log of monthly allowance, students' reported physical health, having elder siblings, gender, and school districts. It is well established that family background characteristics such as parental education levels, monthly allowance, and annual household income strongly affect students' test scores (Davis-Kean, 2005; Duncan et al., 2011; Dahl and Lochner, 2012). As we argue below, it is advisable to control whether students are having elder siblings as their presence might interfere with our identification of any quantity-quality trade-off between children (Becker and Lewis, 1973; Hanushek, 1992; Booth and Kee, 2009; Nitsch et al., 2013). Moreover, gender and physical health conditions of students are controlled as they are commonly known to influence academic outcomes. Because student physical health can be endogenous to our dependent variables (especially mental wellbeing variables), we take an agnostic stand and will present regression results with and without controlling for physical health. Table 3.6 summarizes the descriptive statistics for the control variables.

Table 3.6: Descriptive Statistics for Control Variables

Variables	**7th Grade**		**6th Grade**		**Description**
	obs	mean(sd)	obs	mean(sd)	
year	4,141	0.467761 (0.49902)	4,205	0.475862 (0.499476)	1st grade cohort (2003 borns) = 1 4th grade cohort (2000 borns) = 0
treated months	4,141	0.263946 (0.440824)	4,205	0.263496 (0.440581)	March, April, and May borns = 1 others months = 0
worldcup	4,141	0.123159 (0.328659)	4,205	0.124851 (0.330590)	year ×treated months
income	3,952	4963.059 (2670.548)	4,142	4835.662 (2639.491)	annual income in 10,000 KRW
moeduc	3,881	2.89075 (0.964269)	4,070	2.889189 (0.971077)	mother’s education
allowance	3,977	3.694569 (3.87776)	4,028	2.256356 (1.624053)	monthly allowance in 10,000 KRW
health	4,015	1.708842 (0.562629)	4,198	1.62101 (0.581673)	students’ reported physical health 1 = very healthy 2 = healthy 3 = unhealthy 4 = very unhealthy
eldersibling	4,141	0.505675 (0.500028)	4,205	0.518668 (0.499711)	did not report elder siblings = 0 have elder siblings = 1
gender	4,015	1.47995 (0.499660)	4,205	1.480618 (0.499684)	male = 1 female = 2

Note: moeduc is categorized as follows: 1 = Middle School or Less; 2 = High School; 3 = Community College; 4 = University; 5 = Graduate School. Moreover, students are categorized into a total of 163 school districts in the dataset. Standard deviations are in parentheses.

As discussed in the introduction, previously married couples but also older couples more generally may have had a higher probability of conceiving an additional child during the World Cup. If this is indeed the case, World Cup children were more likely to have an elder sibling than students born in different years or months. As Table 3.7 illustrates, World Cup children (our treatment group) have a 5 percentage points higher probability of reporting an elder sibling with a one-tail p-value of 0.014.

Table 3.7: Probability of Having Elder Siblings (Unconditional Mean)

	Treated Group		Control Group				
	Obs	Mean	Obs	Mean	Mean diff	t-value	P>t
Pr (eldersibling = 1)	510	0.550980	3,631	0.499312	0.051669	2.1861	0.0144
		(0.022047)		(0.008299)	(0.023635)		

Note: Standard deviations are in parentheses.

Assuming that older mothers were already in a stable relationship, we conducted mean tests where we restricted our sample to students whose mothers were more than 30 years old at the time of their births. Table 3.8 presents the respective probabilities and the one-sided t-test statistic.

Table 3.8: Probability of Having Elder Siblings (Mother's Age at Birth > 30)

	Treated Group		Control Group				
	Obs	Mean	Obs	Mean	Mean diff	t-value	P>t
Pr (eldersibling = 1)	222	0.711712	1,374	0.660844	0.050868	1.4933	0.0678
		(0.03047)		(0.012777)	(0.034064)		

Note: Standard deviations are in parentheses.

As can be seen, compared to Table 3.7 the probabilities to have an elder sibling increase by about 16 percentage points for both the treated and the control group. Consequently, the conditional t-test result when mothers' age at birth is greater than 30 resembles the corresponding result in Table 3.7. The treated group has a 5 percentage points higher probability of having an elder sibling with a one-tail p-value of 0.068. The mean tests in Tables 3.7 and 3.8 suggest that the World Cup fertility shock is mostly due to couples in long-term relationships including married couples.[45] To proxy for potentially confounding family background factors, we also control for the presence of an elder sibling in our main regressions.

[45] Unfortunately, the dataset only provides information about the existence of elder (and younger) siblings, but not about their birth years and months. Therefore, we cannot examine whether the quantity-quality trade-off also affects the siblings of World Cup babies.

As discussed in the background section, weak academic performance could also result from a negative selection of parents. If such a distortion was indeed at work, the three family characteristics of World Cup children, mother's age at birth, mother's education level, and household income, should all be significantly lower compared to the control group (Finer and Zolna, 2014; Font-Ribera et al., 2007; Henshaw, 1998). However, if anything these three characteristics point into the opposite direction (see Table 3.9).

Table 3.9: Mean Difference on Family Traits between Treated and Control Groups

	Treated Group		Control Group				
Family Trait	Obs	Mean	Obs	Mean	Mean Diff	t-value	P>t
mother's age	482	30.0104 (0.1602)	3,394	29.3176 (0.0653)	0.6928 (0.1834)	3.7776	0.000
moeduc	491	2.9165 (0.0427)	3,281	2.8817 (0.0168)	0.0348 (0.0466)	0.7465	0.228
income	495	5407.374 (115.369)	3,353	4894.769 (46.305)	512.605 (128.411)	3.9919	0.000
tutortime	510	114.2255 (3.5193)	3,456	118.1623 (1.4823)	-3.9368 (4.0887)	-0.9629	0.168

Note: Standard deviations are in parentheses. Family trait moeduc refers to the mother's education level, income refers to annual household income, and tutortime refers to the average tutoring time per day.

According to Table 3.9, mothers of World Cup children are marginally older at the time of their births, more educated, and wealthier than mothers of control group children.[46] Despite these favorable family characteristics, World Cup children had less tutoring time than the children from the control group. Consequently, it is highly unlikely that a negative selection of families could cause an underperformance of World Cup Children at school.

[46] Our simple DID estimates using family traits as outcome variables also point out that World Cup children may have marginally more favorable (at least no statistically significant differences) family characteristics (see Appendix Table 3.17) than children born in different years or months.

3.4 Results

Table 3.10 displays our main DID regression results: *worldcup* is our DID coefficient showing the academic gap between the World Cup children and controlled students.

Table 3.10: (OLS) DID Coefficients on Academic Outcomes

Variables	Statistics	Model 1	Model 2	Model 3	Model 4
math	worldcup	-0.3372* (0.1871)	-0.3436* (0.1871)	-0.3129* (0.1887)	-0.3202* (0.1887)
	P>|t|	0.072	0.066	0.097	0.090
	observations	3,516	3,516	3,516	3,516
sosci	worldcup	-0.4762*** (0.1736)	-0.4887*** (0.1736)	-0.5091*** (0.1758)	-0.5178*** (0.1759)
	P>|t|	0.006	0.005	0.004	0.003
	observations	3,408	3,408	3,408	3,408
nasci	worldcup	-0.5371*** (0.1732)	-0.5516*** (0.1732)	-0.4716*** (0.1749)	-0.4829*** (0.1749)
	P>|t|	0.002	0.001	0.007	0.006
	observations	3,515	3,515	3,515	3,515
korean	worldcup	-0.3578** (0.1612)	-0.3627** (0.1613)	-0.3491** (0.1627)	-0.3516** (0.1628)
	P>|t|	0.027	0.025	0.032	0.031
	observations	3,514	3,514	3,514	3,514
english	worldcup	-0.6126*** (0.1829)	-0.6253*** (0.1828)	-0.6101*** (0.1850)	-0.6197*** (0.1851)
	P>|t|	0.001	0.001	0.001	0.001
	observations	3,516	3,516	3,516	3,516
health controlled		no	yes	no	yes
school province f.e		yes	yes	-	-
school district f.e		no	no	yes	yes

Note: The table shows the estimated DID coefficients from equation (3.1) for the five academic outcome variables. All regressions control for year, treated months, and individual covariates as they were listed in Section III. A total of 163 school districts or 17 school provinces is also controlled in the regressions. For income and allowance, natural logarithm values are used. Standard errors are in parentheses. *,**,*** denote statistical significance at the 10%, 5% and 1% level.

Our results confirm that students who were born approximately ten months after the 2002 World Cup perform worse than students from the controlled year and months with high statistical significance in all five major subjects. The academic performance gap between World Cup children and other students is most pronounced in English and least in Korean and Mathematics. Equation (3.1) can also be estimated using the ordered probit regression model. As can be seen in Appendix Table 3.18, the corresponding results imply that World Cup children have higher probability of performing worse on school exams than students in the control group.

We also checked the robustness of our main findings for different combinations of treated months and for two specifications with clustered standard errors. Table 3.11 and Table 3.12 report the DID coefficients of interest when using narrower measures of treated months: March and April respectively April and May.

Table 3.11: (OLS) Robustness Checks with
Different Combinations of Treated Months (March & April)

Variables	Statistics	Model 1	Model 2	Model 3	Model 4
math	worldcup	-0.3674*	-0.3845*	-0.3798*	-0.3972*
	(Mar & Apr)	(0.2124)	(0.2126)	(0.2146)	(0.2147)
	P>|t|	0.084	0.071	0.077	0.064
	observations	3,516	3,516	3,516	3,516
sosci	worldcup	-0.2835	-0.2971	-0.3517*	-0.3601*
	(Mar & Apr)	(0.1977)	(0.1977)	(0.2006)	(0.2008)
	P>|t|	0.152	0.133	0.080	0.073
	observations	3,408	3,408	3,408	3,408
nasci	worldcup	-0.6558***	-0.6798***	-0.6099***	-0.6283***
	(Mar & Apr)	(0.1966)	(0.1967)	(0.1989)	(0.1990)
	P>|t|	0.001	0.001	0.002	0.002
	observations	3,515	3,515	3,515	3,515
korean	worldcup	-0.2482	-0.2575	-0.2948	-0.3011
	(Mar & Apr)	(0.1832)	(0.1833)	(0.1851)	(0.1853)
	P>|t|	0.175	0.160	0.111	0.104
	observations	3,514	3,514	3,514	3,514
english	worldcup	-0.5524***	-0.5744***	-0.5747***	-0.5915***
	(Mar & Apr)	(0.2078)	(0.2079)	(0.2106)	(0.2107)
	P>|t|	0.008	0.006	0.006	0.005
	observations	3,516	3,516	3,516	3,516
health controlled		no	yes	no	yes
school province f.e.		yes	yes	-	-
school district f.e		no	no	yes	yes

Note: The same control variables as in the main regressions are used. Standard errors are in parentheses. *,**,*** denote statistical significance at the 10%, 5% and 1% level.

Table 3.12: (OLS) Robustness Checks with
Different Combinations of Treated Months (April & May)

Variables	Statistics	Model 1	Model 2	Model 3	Model 4
math	worldcup	-0.2941	-0.2955	-0.2241	-0.2252
	(Apr & May)	(0.2162)	(0.2162)	(0.2179)	(0.2179)
	P>|t|	0.174	0.172	0.304	0.301
	observations	3,516	3,516	3,516	3,516
sosci	worldcup	-0.5526***	-0.5648***	-0.5249***	-0.5338***
	(Apr & May)	(0.2005)	(0.2004)	(0.2028)	(0.2029)
	P>|t|	0.006	0.005	0.010	0.009
	observations	3,408	3,408	3,408	3,408
nasci	worldcup	-0.4180**	-0.4294**	-0.2981	-0.3069
	(Apr & May)	(0.2004)	(0.2003)	(0.2022)	(0.2021)
	P>|t|	0.037	0.032	0.140	0.129
	observations	3,515	3,515	3,515	3,515
korean	worldcup	-0.3555*	-0.3590*	-0.2814	-0.2826
	(Apr & May)	(0.1864)	(0.1865)	(0.1879)	(0.1880)
	P>|t|	0.057	0.054	0.134	0.133
	observations	3,514	3,514	3,514	3,514
english	worldcup	-0.5942***	-0.6037***	-0.5351**	-0.5421**
	(Apr & May)	(0.2114)	(0.2114)	(0.2138)	(0.2138)
	P>|t|	0.005	0.004	0.012	0.011
	observations	3,516	3,516	3,516	3,516
health controlled		no	yes	no	yes
school province f.e		yes	yes	-	-
school district f.e		no	no	yes	yes

Note: The same control variables as in the main regressions are used. Standard errors are in parentheses. *,**,*** denote statistical significance at the 10%, 5% and 1% level.

By and large, Table 3.11 and Table 3.12 show the robustness of our main regressions. Students born during the treated months in 2003 tend to perform weaker than students born during the controlled year and months. Again, World Cup children had particularly low school test scores in English but suffer less in Korean and Mathematics. Moreover, the ordered probit results corresponding to Tables 11 and 12 (cf. Appendix Tables 3.19 and 3.20) also reconfirm our view that World Cup children underperform at school.

Because we use two different cohorts for the DID design, the reported standard errors may lead to inconsistency issues due to the (possibly) heterogeneous tests that the two cohorts took. In this light, we clustered the observations on the school level and on the school district level. The school level clustering might resolve some of the heterogeneous test issues as school teachers and curricula do not change a lot in three years. Appendix Table 3.21 shows the robustness check results after clustering observations on the school and school district levels. As the table shows, the coefficients are still statistically significant although the standard errors are adjusted after clustering (the corresponding ordered probit regression results are presented in Appendix Table 3.22). The similarity of the estimated coefficients and standard errors lend some additional credibility to the main regression results. Again the academic performance of World Cup children is worse than that of students in the control group.

To show that the control group does not exhibit statistical differences in academic test scores, we ran several tests using placebo months. Table 3.13 summarizes the results from testing the null hypotheses that the interaction terms equal zero when using irrelevant birth months.

Table 3.13: Placebo Test Results on Academic Outcomes

Placebo Months			math	sosci	nasci	korean	english
			school province fixed effects				
Jan	Feb	Jun	-0.1956	0.2045	0.3385*	0.3549*	0.1565
			(0.2110)	(0.1955)	(0.1959)	(0.1821)	(0.2076)
reject the null			No	No	Yes	Yes	No
Feb	Jun	Jul	-0.1567	-0.0113	0.1510	0.0874	0.1261
			(0.2044)	(0.1897)	(0.1898)	(0.1764)	(0.2011)
reject the null			No	No	No	No	No
Jun	Jul	Aug	-0.0204	0.0369	0.0061	-0.0470	0.0422
			(0.2036)	(0.1894)	(0.1892)	(0.1758)	(0.2004)
reject the null			No	No	No	No	No
Jul	Aug	Sep	0.3590*	0.1685	0.3375*	0.1974	0.4295**
			(0.1991)	(0.1852)	(0.1850)	(0.1719)	(0.1958)
reject the null			Yes	No	Yes	No	Yes
Aug	Sep	Oct	0.3211	0.1683	0.2901	0.0201	0.2806
			(0.1973)	(0.1837)	(0.1834)	(0.1704)	(0.1942)
reject the null			No	No	No	No	No
Sep	Oct	Nov	0.2018	0.0123	0.1446	0.0233	0.2711
			(0.1972)	(0.1830)	(0.1832)	(0.1701)	(0.1940)
reject the null			No	No	No	No	No
Oct	Nov	Dec	0.1638	0.2083	-0.0603	-0.0637	0.0390
			(0.2004)	(0.1864)	(0.1858)	(0.1729)	(0.1971)
reject the null			No	No	No	No	No
			school district fixed effects				
Jan	Feb	Jun	-0.2505	0.2034	0.3075	0.3104*	0.1128
			(0.2138)	(0.1992)	(0.1981)	(0.1844)	(0.2111)
reject the null			No	No	No	Yes	No
Feb	Jun	Jul	-0.2798	-0.0720	0.0274	-0.0319	-0.0273
			(0.2066)	(0.1927)	(0.1914)	(0.1781)	(0.2040)
reject the null			No	No	No	No	No
Jun	Jul	Aug	-0.0813	0.0162	-0.0241	-0.1363	0.0436
			(0.2068)	(0.1934)	(0.1916)	(0.1783)	(0.2041)
reject the null			No	No	No	No	No
Jul	Aug	Sep	0.3442*	0.1811	0.3005	0.1524	0.4374**
			(0.2022)	(0.1893)	(0.1875)	(0.1744)	(0.1995)
reject the null			Yes	No	No	No	Yes
Aug	Sep	Oct	0.3744*	0.1820	0.3245*	0.0668	0.3383*
			(0.2001)	(0.1874)	(0.1855)	(0.1727)	(0.1975)
reject the null			Yes	No	Yes	No	Yes
Sep	Oct	Nov	0.2026	-0.0109	0.0771	0.0470	0.2250
			(0.2000)	(0.1864)	(0.1854)	(0.1722)	(0.1974)
			No	No	No	No	No
Oct	Nov	Dec	0.1435	0.1539	-0.0878	-0.0259	0.0110
			(0.2034)	(0.1900)	(0.1881)	(0.1752)	(0.2006)
reject the null			No	No	No	No	No
Observations			3,060	2,959	3,059	3,058	3,060

Note: The same control variables as in the main regression are used except for the treated months. Reject the null-hypothesis (θ=0) if p<0.10. Observations from students born in the treated months and treated year are not used in these regressions. School province (district) fixed effect results are based on estimating Model 2 (Model 4).

As can be seen in Table 3.13, using an interaction term resulting from irrelevant birth months by and large exerts no statistically significant influence on academic outcome measures. This reassures us of the validity of the significant negative coefficients reported in Table 3.10. World Cup children indeed underperform in school tests.

Table 3.14 displays our main mental wellbeing DID regression results. As before, *worldcup* is our DID coefficient showing the mental wellbeing gap between the World Cup children and students in the control group.

Table 3.14: (OLS) DID Coefficients on Mental Wellbeing Outcomes

Variables	Statistics	Model 1	Model 2	Model 3	Model 4
depressed	worldcup	0.1256**	0.1271**	0.1417**	0.1417**
		(0.0579)	(0.0558)	(0.0587)	(0.0565)
	P>|t|	0.030	0.023	0.016	0.012
	observations	3,566	3,559	3,566	3,559
suicidal	worldcup	0.0967*	0.0978*	0.1098**	0.1100**
		(0.0520)	(0.0511)	(0.0526)	(0.0516)
	P>|t|	0.063	0.056	0.037	0.033
	observations	3,566	3,559	3,566	3,559
self-reproach	worldcup	0.0663	0.0643	0.0787	0.0763
		(0.0632)	(0.0618)	(0.0640)	(0.0626)
	P>|t|	0.294	0.298	0.219	0.223
	observations	3,566	3,559	3,566	3,559
bullying	worldcup	-0.0766**	-0.0770**	-0.0870***	-0.0872***
		(0.0302)	(0.0303)	(0.0308)	(0.0308)
	P>|t|	0.011	0.011	0.005	0.005
	observations	3,566	3,559	3,566	3,559
violence	worldcup	-0.0612	-0.0615	-0.0621	-0.0623
		(0.0381)	(0.0381)	(0.0390)	(0.0391)
	P>|t|	0.108	0.107	0.111	0.111
	observations	3,566	3,559	3,566	3,559
health controlled		no	yes	no	yes
school province f.e		yes	yes	-	-
school district f.e		no	no	yes	yes

Note: The same control variables as in the main regressions are used. Standard errors are in parentheses. *,**,*** denote statistical significance at the 10%, 5% and 1% level. The variables **depressed**, **suicidal**, and **self-reproach** are categorical where ”1“ refers to strong yes, ”2“ refers to yes, ”3“ refers to no, and ”4“ refers to strong no. Therefore, students with better mental health report higher numerical values for these variables. The variables **bullying** and **violence**, in contrast, measure the frequency of instances in which students have exhibited bullying behavior or physical aggression toward others. If students do not engage in such actions, they are assigned 0. Therefore, students who behaved well in a school context tend to have lower numerical values for these variables.

As is evident, World Cup children fare better than students in the control group. They generally feel less depressed and have fewer suicidal impulses than students born in different years or months. Although it is not statistically significant at conventional levels, World Cup children tend to self-blame less than students in the control group. Moreover, they tend to cause less trouble among their peers. They generally exercise less bullying and also direct less violence against other students or classmates than students in the control group. The ordered probit regression results for the categorical outcome variables (i.e., depressed, suicidal, and self-reproach) also reconfirm our view that World Cup children generally exhibit better mental wellbeing than students born in different year and months (see Appendix Table 3.23).

As in the academic outcome regressions, we also clustered observations on the school and school district level in our mental wellbeing analysis (see Appendix Table 3.24). Again the statistical significance and magnitude of our clustered regression coefficients show the robustness of our main empirical results that World Cup children generally have higher degrees of mental wellbeing than the students in the control group. Ordered probit regressions on categorical variables corroborate our OLS findings (see Appendix Table 3.25).

We also conducted the placebo tests for our mental wellbeing outcome regressions to show there is no statistical difference in mental wellbeing among students in the control group (see Table 3.15).

Table 3.15: Placebo Test Results on Mental Wellbeing Outcomes

Placebo Months	depressed	suicidal	self-reproach	bullying	violence
	school province fixed effects				
Jan Feb Jun	0.0243	0.0639	0.0880	0.0215	0.0253
	(0.0638)	(0.0582)	(0.0695)	(0.0362)	(0.0454)
reject the null	No	No	No	No	No
Feb Jun Jul	0.0110	0.0182	0.0064	0.0173	-0.0200
	(0.0621)	(0.0566)	(0.0676)	(0.0353)	(0.0441)
reject the null	No	No	No	No	No
Jun Jul Aug	0.0171	-0.0141	-0.0329	-0.0018	-0.0195
	(0.0614)	(0.0560)	(0.0669)	(0.0349)	(0.0436)
reject the null	No	No	No	No	No
Jul Aug Sep	-0.0256	-0.0605	-0.0647	0.0193	0.0120
	(0.0601)	(0.0548)	(0.0654)	(0.0341)	(0.0427)
reject the null	No	No	No	No	No
Aug Sep Oct	0.0230	-0.0099	-0.0195	0.0573*	0.0503
	(0.0596)	(0.0543)	(0.0649)	(0.0338)	(0.0423)
reject the null	No	No	No	Yes	No
Sep Oct Nov	-0.0569	-0.0139	0.0037	0.0662	0.0489
	(0.0596)	(0.0544)	(0.0649)	(0.0338)	(0.0424)
reject the null	No	No	No	Yes	No
Oct Nov Dec	-0.0788	-0.0791	-0.0349	0.0290	0.0254
	(0.0604)	(0.0552)	(0.0659)	(0.0344)	(0.0430)
reject the null	No	No	No	No	No
	school district fixed effects				
Jan Feb Jun	0.0244	0.0496	0.0766	0.0319	0.0177
	(0.0651)	(0.0591)	(0.0706)	(0.0371)	(0.0467)
reject the null	No	No	No	No	No
Feb Jun Jul	0.0013	-0.0078	-0.0124	0.0169	-0.0255
	(0.0630)	(0.0572)	(0.0684)	(0.0360)	(0.0453)
reject the null	No	No	No	No	No
Jun Jul Aug	0.0124	-0.0195	-0.0320	-0.0053	-0.0207
	(0.0626)	(0.0568)	(0.0679)	(0.0357)	(0.0450)
reject the null	No	No	No	No	No
Jul Aug Sep	-0.0259	-0.0516	-0.0678	0.0109	0.0063
	(0.0611)	(0.0555)	(0.0664)	(0.0349)	(0.0439)
reject the null	No	No	No	No	No
Aug Sep Oct	0.0167	-0.0042	-0.0246	0.0560	0.0513
	(0.0607)	(0.0551)	(0.0659)	(0.0346)	(0.0436)
reject the null	No	No	No	No	No
Sep Oct Nov	-0.0564	-0.0132	0.0007	0.0780**	0.0416
	(0.0610)	(0.0553)	(0.0662)	(0.0348)	(0.0438)
reject the null	No	No	No	Yes	No
Oct Nov Dec	-0.0878	-0.0856	-0.0198	0.0427	0.0337
	(0.0617)	(0.0560)	(0.0671)	(0.0352)	(0.0444)
reject the null	No	No	No	No	No
Observations	3,100	3,100	3,100	3,100	3,100

Note: The same control variables as in the main regression are used except for the treated months. Reject the null-hypothesis (θ=0) if p<0.10. Observations from students born in the treated months and treated year are not used in these regressions. School province (district) fixed effect results are based on estimating Model 2 (Model 4).

As Table 3.15 shows, in almost all specifications we cannot reject the null hypothesis that students born in hypothetically treated months enjoy the same level of mental wellbeing than students in the control group.

3.5 Discussion

Our analysis reveals strong empirical evidence that the positive fertility shock caused by the 2002 World Cup also had a significant adverse effect on students' human capital formation. Our findings, therefore, produce evidence for the existence of a trade-off between child quantity and quality in South Korea. As Tables 3.10, 3.11, and 3.12 show, World Cup children performed worse than students in the control group in all subject areas and the effects are especially pronounced in English and less pronounced in Korean and Mathematics. At the same time we find evidence that students fare better in terms of mental wellbeing, which might be a reflection of less pressure and lower expectations from the parents of World Cup children.

Given the linguistic difference between English and Korean, the acquisition of English as a second language is particularly difficult for Korean students. Its different structure, pronunciation, and phrasing make learning the English language very time intensive for Korean students. Not surprisingly, private tutoring expenditures on English education are the highest among all academic subjects (Statistics Korea, 2019). The fact that World Cup children produce significantly lower test scores in English may therefore result from lower parental investments relative to the control group (see Table 3.9 and Appendix Table 3.17). The above line of argument does not apply - or only to a much lesser degree - to the acquisition of the Korean language and mathematics skills. Korean language skills in particular can be gained more naturally until puberty around the age of 12 or 13 (Fromkin et al., 2014). We suspect that differences in the skill acquisition processes most likely explain why the test score gaps between

World Cup children and students in the control group are less pronounced in Korean and Mathematics than they are in English.

The results of our analysis are also consistent with the notion that the parents of World Cup children have lower expectations with respect to the academic performance of their offspring. Compared to students born in different years and months, World Cup children might therefore feel less pressure from their parents. Empirical evidence suggests that students facing high parental expectations get more stress from low test scores (Lee and Kang, 2018; Shin et al., 2018). This stress factor then lowers the mental wellbeing measure or deteriorates the mental health statuses of adolescents (Ma et al., 2018; Almroth et al., 2019). World Cup children, in contrast, exhibit generally higher degrees of mental wellbeing than students born in different years or months (see Table 3.14 and Appendix Table 3.24). They feel less depressed and have fewer suicidal impulses than students in the control group. In this light, our results indeed insinuate that World Cup children experienced less academic pressure from their parents than students in the control group. The empirical results corroborate our view that both parental expectations and investment in human capital formation are significantly lower for World Cup children which then leads to low test scores but content students.

The alternative mechanism via a negative selection of parents, in contrast, is very unlikely. First and foremost, mothers of World Cup children are marginally older at the time of birth, more educated, and have higher incomes than mothers of children born in different years or months. Previous empirical studies suggest that such a selection of mothers would indeed lead to favorable academic and mental health child outcome. Therefore this selection of mothers might have even caused an attenuation of our results. Similarly, if a selection of mothers with a lower aversion against risky behavior was driving our results (unfortunately we can neither confirm nor reject this hypothesis), we would expect their children to underperform at school and to exhibit severe mental health problems which is of course not the case. We therefore argue that our regression results are a reflection of the child quantity-quality trade-off in South Korea.

Last but not least, we would like to note that the empirical findings presented in this paper also hint at the adverse consequences that are associated with a competitive educational environment. South Korea is known for its pervasive "education fever" and relentless educational system aimed at qualifying students for admissions to prestigious high schools and universities (Anderson & Kohler, 2012; Lee, 2005). Given the substantial educational expenses and the high expectations imposed by parents, students in South Korea experience greater stress and poorer mental health compared to their peers in other middle and high income countries (Rudolf and Bethmann, 2022), which is reflected, among other things, in a high suicide rate among young Koreans.[47] In such a competitive environment, reducing educational expenditures and parental expectations may actually increase the mental wellbeing and contentment of students.

3.6 Conclusion

The Korean National Football team experienced miraculous match results during the home World Cup in the June of 2002. The events caused a euphoria among Koreans that led to a temporary and significant increase in the country's fertility rate in the subsequent spring. Given its long duration and unforeseen nature, the football tournament hence provides us with the quasi-experimental event needed for statistical identification. In a first step, we showed that the World Cup indeed had a significant positive impact on South Korean fertility. Second, we used the episode to study the Beckerian trade-off between child quantity and quality. Being more numerous, we hypothesized that the "World Cup children" were likely to show a lower academic

[47] The average suicide rate among 15 to 19 years old teenagers in the 38 OECD countries is 6.26 (per 100,000 population) with a standard deviation of 0.59. In South Korea this rate is 9.90. Source: World Health Organization (Suicide Rate Estimates, Crude, 15-19) at https://apps.who.int/gho/data/view.main.MHSUICIDE15TO19v [last accessed on October 9th, 2023].

performance in major school subjects. Last, we changed our perspective and checked whether the event also affected the children's mental wellbeing.

Our empirical results show that World Cup children – born approximately ten months after the tournament – tend to underperform in all five academic subjects. The results are more pronounced in English and less pronounced in Korean and Mathematics. Since these findings most likely result from inferior parental investments, they are in line with the Beckerian notion of a trade-off between child quantity and quality. Our mental wellbeing regression results, in turn, indicate that World Cup children may experience less pressure from their parents as they generally feel less depressed, have fewer suicidal impulses, exert less self-blame, exercise less bullying, and direct less violence against classmates than students in the control group.

Our research adds to the existing literature by producing additional empirical evidence in favor of the existence of a trade-off between the quantity and quality of children. In our empirical strategy we used an unusually long-lasting exogenous shock to South Korean fertility caused by the 2002 World Cup. Two contributions of our work are worth mentioning. First, we complemented the existing studies with their focus on Western countries by providing corroborative evidence from an East Asian country. Second, we also added the children's perspective. Interestingly, we found evidence that children fare better in terms of mental wellbeing despite underperforming at school. It goes without saying that both of these contributions should not be viewed as final or definite answers but rather as inspirations for further work in that same direction.

3.7 Appendix

3.7.1 Additional Results

We conducted a series of statistical tests to check whether the increases of monthly new births in March, April, and (or) May of 2003 are in fact significant. For this purpose, we used the Monthly New Borns dataset from Statistics Korea (2021) from 2000 to 2020 and analyzed the data using heteroskedasticity robust OLS and Tobit regressions after controlling for both years and months. Table 3.16 shows the results from these tests. Note that the null-hypotheses assume no change in new births.

Table 3.16: Hypotheses Testing on Monthly New Birth with World Cup Dummies

treated month(s)	test type	regression type	test statistic	reject the null
2003 Mar	t-test	ols	1.96	Yes
		tobit	2.11	Yes
2003 Apr	t-test	ols	2.55	Yes
		tobit	2.75	Yes
2003 May	t-test	ols	0.96	No
		tobit	1.03	No
2003 Mar & Apr	f-test	ols	3.26	Yes
		tobit	3.77	Yes
2003 Mar & May	f-test	ols	8.49	Yes
		tobit	9.82	Yes
2003 Mar & Apr & May	f-test	ols	5.95	Yes
		tobit	6.88	Yes

Note: Reject the null-hypothesis if [P>|t-statistic|] < 0.10 or [P>f-statistic] < 0.10. The number of observations is 252.

As Table 3.16 shows, the number of new births increased significantly in the spring of 2003. The rejection of joint hypothesis tests using March, April, and (or) May of 2003 reconfirms our view that Korea indeed experienced a temporary increase in fertility roughly ten months after the World Cup.

Mean difference tests in Table 9 analyze the unconditional means of control and treatment groups, showing that the mothers of World Cup children generally show marginally better (family) characteristics (i.e., older at the time of their birth, more educated, and have higher household income). These family traits of World Cup children can still be observed even if we control for gender and regional fixed effects. The results in Table 3.17 are the estimated β_3 coefficients of the following regression model:

$$Y = \beta_0 + \beta_1 Year + \beta_2 Treated\ Months + \beta_3 World\ Cup + \beta_4\ Gender + \delta_{region\ FE} + \varepsilon$$

Y denotes outcome variables such as mother's age, mother's education level, household income, and tutoring time. $\delta_{region\ FE}$ denotes school province or school district fixed effects.

Table 3.17: Family Traits of World Cup Children

Variables	**Statistics**	**Model A**	**Model B**
eldersibling	worldcup	0.0393	0.0428
		(0.0359)	(0.0363)
	P>|t|	0.273	0.238
	observations	3,972	3,972
mother's age	worldcup	0.2726	0.2574
		(0.2754)	(0.2808)
	P>|t|	0.322	0.359
	observations	3,729	3,729
moeduc	worldcup	0.0788	0.0690
		(0.0701)	(0.0682)
	P>|t|	0.261	0.312
	observations	3,711	3,711
income	worldcup	155.5359	147.8808
		(192.5402)	(188.8262)
	P>|t|	0.419	0.434
	observations	3,794	3,794
tutortime	worldcup	-16.0236***	-15.0060**
		(6.1668)	(6.1684)
	P>|t|	0.009	0.015
	observations	3,965	3,965
school province f.e		yes	no
school district f.e		no	yes

Note: Standard errors are in parentheses. *,**,*** denote statistical significance at the 10%, 5% and 1% level.

Table 3.18: (Ordered-Probit) DID Coefficients on Academic Outcomes

Variables	Statistics	Model1	Model2	Model3	Model4
math	**worldcup**	-0.1408*	-0.1442*	-0.1416*	-0.1456*
		(0.0800)	(0.0800)	(0.0816)	(0.0817)
	P>|t|	0.078	0.072	0.083	0.075
	observations	3,516	3,516	3,516	3,516
sosci	**worldcup**	-0.2366***	-0.2428***	-0.2604***	-0.2649***
		(0.0804)	(0.0804)	(0.0820)	(0.0820)
	P>|t|	0.003	0.003	0.001	0.001
	observations	3,408	3,408	3,408	3,408
nasci	**worldcup**	-0.2383***	-0.2460***	-0.2149***	-0.2213***
		(0.0793)	(0.0793)	(0.0809)	(0.0809)
	P>|t|	0.003	0.002	0.008	0.006
	observations	3,515	3,515	3,515	3,515
korean	**worldcup**	-0.1515*	-0.1547**	-0.1526*	-0.1547*
		(0.0786)	(0.0787)	(0.0802)	(0.0803)
	P>|t|	0.054	0.049	0.057	0.054
	observations	3,514	3,514	3,514	3,514
english	**worldcup**	-0.2803***	-0.2874***	-0.2937***	-0.2997***
		(0.0801)	(0.0802)	(0.0818)	(0.0819)
	P>|t|	0.000	0.000	0.000	0.000
	observations	3,516	3,516	3,516	3,516
health controlled		no	yes	no	yes
school province f.e		yes	yes	-	-
school district f.e		no	no	yes	yes

Note: The table shows the estimated DID coefficients from equation (3.1) for the five academic outcome variables. All regressions control for year, treated months, and individual covariates as they were listed in Section III. A total of 163 school districts or 17 school provinces is also controlled in the regressions. For income and allowance, natural logarithm values are used. Standard errors are in parentheses. *,**,*** denote statistical significance at the 10%, 5% and 1% level.

Table 3.19: (Ordered-Probit) Robustness Checks with
Different Combinations of Treated Months (March & April)

Variables	Statistics	Model1	Model2	Model3	Model4
math	worldcup	-0.1253	-0.1331	-0.1409	-0.1492
	(Mar & Apr)	(0.0908)	(0.0909)	(0.0929)	(0.0930)
	P>|t|	0.168	0.143	0.129	0.109
	observations	3,516	3,516	3,516	3,516
sosci	worldcup	-0.1475	-0.1541*	-0.1811*	-0.1856**
	(Mar & Apr)	(0.0914)	(0.0915)	(0.0934)	(0.0935)
	P>|t|	0.107	0.092	0.053	0.047
	observations	3,408	3,408	3,408	3,408
nasci	worldcup	-0.3026***	-0.3150***	-0.2901***	-0.3001***
	(Mar & Apr)	(0.0900)	(0.0901)	(0.0920)	(0.0921)
	P>|t|	0.001	0.000	0.002	0.001
	observations	3,515	3,515	3,515	3,515
korean	worldcup	-0.1015	-0.1073	-0.1296	-0.1343
	(Mar & Apr)	(0.0893)	(0.0894)	(0.0913)	(0.0914)
	P>|t|	0.256	0.230	0.156	0.142
	observations	3,514	3,514	3,514	3,514
english	worldcup	-0.2407***	-0.2518***	-0.2627***	-0.2721***
	(Mar & Apr)	(0.0908)	(0.0909)	(0.0929)	(0.0930)
	P>|t|	0.008	0.006	0.005	0.003
	observations	3,516	3,516	3,516	3,516
health controlled		no	yes	no	yes
school province f.e		yes	yes	-	-
school district f.e		no	no	yes	yes

Note: The same control variables as in the main regressions are used. Standard errors are in parentheses. *,**,*** denote statistical significance at the 10%, 5% and 1% level.

Table 3.20: (Ordered-Probit) Robustness Checks with
Different Combinations of Treated Months (April & May)

Variables	Statistics	Model1	Model2	Model3	Model4
math	worldcup	-0.1409	-0.1423	-0.1218	-0.1232
	(Apr & May)	(0.0923)	(0.0923)	(0.0941)	(0.0941)
	P>|t|	0.127	0.123	0.196	0.190
	observations	3,516	3,516	3,516	3,516
sosci	worldcup	-0.2713***	-0.2774***	-0.2692***	-0.2740***
	(Apr & May)	(0.0927)	(0.0927)	(0.0944)	(0.0944)
	P>|t|	0.003	0.003	0.004	0.004
	observations	3,408	3,408	3,408	3,408
nasci	worldcup	-0.1882**	-0.1942**	-0.1390	-0.1438
	(Apr & May)	(0.0916)	(0.0916)	(0.0934)	(0.09340)
	P>|t|	0.040	0.034	0.137	0.124
	observations	3,515	3,515	3,515	3,515
korean	worldcup	-0.1550*	-0.1575*	-0.1226	-0.1240
	(Apr & May)	(0.0908)	(0.0908)	(0.0926)	(0.0926)
	P>|t|	0.088	0.083	0.185	0.181
	observations	3,514	3,514	3,514	3,514
english	worldcup	-0.2836***	-0.2888***	-0.2748***	-0.2790***
	(Apr & May)	(0.0925)	(0.0926)	(0.0944)	(0.0944)
	P>|t|	0.002	0.002	0.004	0.003
	observations	3,516	3,516	3,516	3,516
health controlled		no	yes	no	yes
school province f.e		yes	yes	-	-
school district f.e		no	no	yes	yes

Note: The same control variables as in the main regressions are used. Standard errors are in parentheses. *,**,*** denote statistical significance at the 10%, 5% and 1% level.

Table 3.21: (OLS) Academic Robustness Checks Using Clustered Standard Errors

Variables	Statistics	Model 5	Model 6	Model 7	Model 8
math	worldcup	-0.3372*	-0.3436*	-0.3372*	-0.3436*
		(0.1921)	(0.1925)	(0.1845)	(0.1837)
	P>|t|	0.08	0.075	0.069	0.063
	clusters	876	876	163	163
sosci	worldcup	-0.4762***	-0.4887***	-0.4762**	-0.4887**
		(0.1844)	(0.1838)	(0.2116)	(0.2098)
	P>|t|	0.01	0.008	0.026	0.021
	clusters	858	858	163	163
nasci	worldcup	-0.5371***	-0.5516***	-0.5371***	-0.5516***
		(0.1809)	(0.1804)	(0.2006)	(0.1997)
	P>|t|	0.003	0.002	0.008	0.006
	clusters	876	876	163	163
korean	worldcup	-0.3578**	-0.3627**	-0.3578**	-0.3627**
		(0.1700)	(0.1699)	(0.1845)	(0.1839)
	P>|t|	0.036	0.033	0.054	0.05
	clusters	874	874	163	163
english	worldcup	-0.6126***	-0.6253***	-0.6126***	-0.6253***
		(0.1850)	(0.1846)	(0.1852)	(0.1834)
	P>|t|	0.001	0.001	0.001	0.001
	clusters	876	876	163	163
health controlled		no	yes	no	yes
school province f.e		yes	yes	yes	yes
school level clustering		yes	yes	no	no
school district level clustering		no	no	yes	yes

Note: The same control variables as in the main regressions are used in the school level or school district level standard error clustering regression. Standard errors are in parentheses. *,**,*** denote statistical significance at the 10%, 5% and 1% level.

Table 3.22: (Ordered-Probit) Academic Robustness Checks Using Clustered Standard Errors

Variables	Statistics	Model5	Model6	Model7	Model8
math	worldcup	-0.1408*	-0.1442*	-0.1408*	-0.1442*
		(0.0823)	(0.0826)	(0.0780)	(0.0778)
	P>|t|	0.087	0.081	0.071	0.064
	clusters	876	876	163	163
sosci	worldcup	-0.2366***	-0.2428***	-0.2366**	-0.2428**
		(0.0857)	(0.0854)	(0.0988)	(0.0981)
	P>|t|	0.006	0.004	0.017	0.013
	clusters	858	858	163	163
nasci	worldcup	-0.2383***	-0.2460***	-0.2383***	-0.2460***
		(0.0825)	(0.0825)	(0.0930)	(0.0925)
	P>|t|	0.004	0.003	0.010	0.008
	clusters	876	876	163	163
korean	worldcup	-0.1515*	-0.1547*	-0.1515*	-0.1547*
		(0.0834)	(0.0832)	(0.0913)	(0.0910)
	P>|t|	0.069	0.063	0.097	0.089
	clusters	874	874	163	163
english	worldcup	-0.2803***	-0.2874***	-0.2803***	-0.2874***
		(0.0826)	(0.0824)	(0.0813)	(0.0806)
	P>|t|	0.001	0.000	0.001	0.000
	clusters	876	876	163	163
health controlled		no	yes	no	yes
school province f.e		yes	yes	yes	yes
school level clustering		yes	yes	no	no
school district level clustering		no	no	yes	yes

Note: The same control variables as in the main regressions are used in the school level or school district level standard error clustering regression. Standard errors are in parentheses. *,**,*** denote statistical significance at the 10%, 5% and 1% level.

Table 3.23: (Ordered-Probit) DID Coefficients on Mental Wellbeing Outcomes

Variables	**Statistics**	**Model1**	**Model2**	**Model3**	**Model4**
depressed	worldcup	0.1720**	0.1864**	0.2060**	0.2196**
		(0.0868)	(0.0879)	(0.0890)	(0.0901)
	P>|t|	0.048	0.034	0.021	0.015
	observations	3,566	3,559	3,566	3,559
suicidal	worldcup	0.1670*	0.1738*	0.1955**	0.1992**
		(0.0939)	(0.0947)	(0.0968)	(0.0976)
	P>|t|	0.075	0.066	0.043	0.041
	observations	3,566	3,559	3,566	3,559
self-reproach	worldcup	0.1006	0.1032	0.1214	0.1240
		(0.0837)	(0.0843)	(0.0857)	(0.0862)
	P>|t|	0.229	0.221	0.156	0.151
	observations	3,566	3,559	3,566	3,559
health controlled		no	yes	no	yes
school province f.e		yes	yes	-	-
school district f.e		no	no	yes	yes

Note: The same control variables as in the main regressions are used. Standard errors are in parentheses. *,**,*** denote statistical significance at the 10%, 5% and 1% level.

Table 3.24: (OLS) Diff-in-Diff Results with Clustered Standard Errors

Variables	**Statistics**	**Model 5**	**Model 6**	**Model 7**	**Model 8**
depressed	worldcup	0.1256**	0.1271**	0.1256**	0.1271**
		(0.0556)	(0.0536)	(0.0562)	(0.0552)
	P>|t|	0.024	0.018	0.027	0.023
	clusters	564	563	155	155
suicidal	worldcup	0.0967*	0.0978*	0.0967*	0.0978*
		(0.0527)	(0.0510)	(0.0524)	(0.0508)
	P>|t|	0.067	0.055	0.067	0.056
	clusters	564	563	155	155
self-reproach	worldcup	0.0663	0.0643	0.0663	0.0643
		(0.0657)	(0.0641)	(0.0691)	(0.0668)
	P>|t|	0.314	0.316	0.339	0.337
	clusters	564	563	155	155
bullying	worldcup	-0.0766**	-0.0770**	-0.0766**	-0.0770**
		(0.0359)	(0.0360)	(0.0375)	(0.0375)
	P>|t|	0.033	0.033	0.043	0.042
	clusters	564	563	155	155
violence	worldcup	-0.0612	-0.0615	-0.0612	-0.0615
		(0.0443)	(0.0445)	(0.0411)	(0.0411)
	P>|t|	0.168	0.167	0.138	0.137
	clusters	564	563	155	155
health controlled		no	yes	no	yes
school province f.e		yes	yes	yes	yes
school level clustering		yes	yes	no	no
school district level clustering		no	no	yes	yes

Note: The same control variables as in the main regressions are used. Standard errors are in parentheses. *,**,*** denote statistical significance at the 10%, 5% and 1% level. The variables **depressed**, **suicidal**, and **self-reproach** are categorical where ”1“ refers to strong yes, ”2“ refers to yes, ”3“ refers to no, and ”4“ refers to strong no. Therefore, students with better mental health report higher numerical values for these variables. The variables **bullying** and **violence**, in contrast, measure the frequency of instances in which students have exhibited bullying behavior or physical aggression toward others. If students do not engage in such incidents, they are assigned 0. Therefore, students who behaved well in a school context tend to have lower numerical values for these variables.

Table 3.25: (Ordered-Probit) Diff-in-Diff Results with Clustered Standard Errors

Variables	**Statistics**	**Model5**	**Model6**	**Model7**	**Model8**
depressed	worldcup	0.1720**	0.1864**	0.1720**	0.1864**
		(0.0830)	(0.0839)	(0.0826)	(0.0860)
	P>|t|	0.038	0.026	0.037	0.030
	clusters	564	563	155	155
suicidal	clusters	0.1670*	0.1738*	0.1670*	0.1738*
		(0.0949)	(0.0940)	(0.0941)	(0.0934)
	P>|t|	0.078	0.064	0.076	0.063
	clusters	564	563	155	155
self-reproach	worldcup	0.1006	0.1032	0.1006	0.1032
		(0.0876)	(0.0880)	(0.0918)	(0.0921)
	P>|t|	0.250	0.241	0.273	0.263
	clusters	564	563	155	155
health controlled		no	yes	no	yes
school province fixed effect		yes	yes	yes	yes
school level clustering		yes	yes	no	no
school district level clustering		no	no	yes	yes

Note: The same control variables as in the main regressions are used in the school level or school district level standard error clustering regression. Standard errors are in parentheses. *,**,*** denote statistical significance at the 10%, 5% and 1% level.

References

Alengoz, L., Castellani, M., and Squazzoni, F. (2017). Mood Implications on Social Behaviour in Complex Societies. A Literature Review. *Sociologica, 12*(3).

Almroth, M., László, K. D., Kosidou, K., and Galanti, M. R. (2019). Academic Expectations and Mental Health in Adolescence: A Longitudinal Study Involving Parents' and Their Children's Perspectives. *Journal of Adolescent Health, 64*(6), 783–789.

Anderson, M. L. (2008). Multiple Inference and Gender Differences in the Effects of Early Intervention: A Reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects. *Journal of the American Statistical Association, 103*(484), 1481–1495.

Anderson, T., & Kohler, H.-P. (2013). Education Fever and the East Asian Fertility Puzzle. *Asian Population Studies, 9*(2), 196–215.

Angrist, J., Lavy, V., and Schlosser, A. (2010). Multiple Experiments for the Causal Link between the Quantity and Quality of Children. *Journal of Labor Economics, 28*(4), 773–824.

Arceo, E., Hanna, R., & Oliva, P. (2016). Does the Effect of Pollution on Infant Mortality Differ between Developing and Developed countries? Evidence from Mexico City. *The Economic Journal, 126*(591), 257–280.

Armstrong, C., Hill, M., & Secker, J. (2000). Young People's Perceptions of Mental Health. *Children & Society, 14*(1), 60–72.

Astbury, Jill (2001). Gender Disparities in Mental health. In: Mental health. Ministerial Round Tables 2001, 54th World Health Assemble, WHO, Geneva, Switzerland.

Ayuso-Álvarez, A., Simón, L., Nuñez, O., Rodríguez-Blázquez, C., Martín-Méndez, I., Bel-lán, A., López-Abente, G., Merlo, J., Fernandez-Navarro, P., & Galán, I. (2019). Association between Heavy Metals and Metalloids in Topsoil and Mental Health in the Adult Population of Spain. *Environmental Research, 179*, 108784.

Bakis, O., Inci, E., & Senturk, R. O. (2019). Unbundling Curbside Parking Costs from Housing Prices. *Journal of Economic Geography*, 19(1), 89–119.

Bakolis, I., Hammoud, R., Stewart, R., Beevers, S., Dajnak, D., MacCrimmon, S., Broadbent, M., Pritchard, M., Shiode, N., Fecht, D., Gulliver, J., Hotopf, M., Hatch, S. L., & Mudway, I. S. (2020). Mental Health Consequences of Urban Air Pollution: Prospective Population-based Longitudinal Survey. *Social Psychiatry and Psychiatric Epidemiology*, *56*(9), 1587–1599.

Ball, M. J. (1973). Recent Empirical Work on the Determinants of Relative House Prices, *Urban Studies*, 10(2), 213–233.

Balakrishnan, U., & Tsaneva, M. (2021). Air Pollution and Academic Performance: Evidence from India. *World Development, 146*, 105553.

Banzhaf, H. S. (2021). Difference-in-differences Hedonics, *Journal of Political Economy, 129*(8), 2385–2414.

Bartik, T. J. (1987). The Estimation of Demand Parameters in Hedonic Price Models. *Journal of Political Economy*, *95*(1), 81–88.

Bauer, T.K., Braun, S.T. & Kvasnicka, M. (2017). Nuclear Power Plant Closures and Local Housing Values: Evidence from Fukushima and the German Housing Market. *Journal of Urban Economics*, *99*, 94–106.

Beck, J., Lin, M. 2020. The Impact of Sea Level Rise on Real Estate Prices in Coastal Georgia. *The Review of Regional Studies*, 50(1), 43-52.

Becker, G. S. (1960). An Economic Analysis of Fertility. *NBER*, 209–240.

Becker, G. S., and Lewis, H. G. (1973). On the Interaction between the Quantity and Quality of Children. *Journal of Political Economy, 81*(2), S279–288.

Becker, G. S., and Tomes, N. (1976). Child Endowments and the Quantity and Quality of Children. *Journal of Political Economy 84*(4), S143–S162.

Becker, N. B., Jesus, S. N., João, K. A., Viseu, J. N., & Martins, R. I. (2017). Depression and Sleep Quality in Older Adults: A Meta-analysis. *Psychology, Health & Medicine, 22(*8), 889–895.

Bernardi, F., and Cozzani, M. (2021). Soccer Scores, Short-Term Mood and Fertility. *European Journal of Population 37*, 625–641.

Bernstein, A., Gustafson, M.T., Lewis, R. 2019. Disaster on the Horizon: The Price Effect of Sea Level Rise, *Journal of Financial Economics*, 134(2), 253-272.

Bethmann, D., and Cho, J.I (2021). How Long Does It Last? The Relative Age Effect in Korean Elementary Education. *Korea University Institute of Economic Research Discussion Paper Series*, 2106.

Bhagure, G. R., & Mirgane, S. R. (2011). Heavy Metal Concentrations in Groundwaters and Soils of Thane Region of Maharashtra, India. *Environmental Monitoring and Assessment*, *173*, 643-652.

Binter, A.-C., Kusters, M. S. W., van den Dries, M. A., Alonso, L., Lubczyńska, M. J., Hoek, G., White, T., Iñiguez, C., Tiemeier, H., & Guxens, M. (2022). Air Pollution, White Matter Microstructure, and Brain Volumes: Periods of Susceptibility from Pregnancy to Preadolescence. *Environmental Pollution, 313*, 120109.

Bishop, K., Ketcham, J., & Kuminoff, N. (2017). Hazed and Confused: The Effect of Air Pollution on Dementia. *Working Paper.*

Black, S.E. 1999. Do Better Schools Matter? Parental Valuation of Elementary Education, *Quarterly Journal of Economics*, 114(2), 577–599.

Black, S. E., Devereux, P. J., Salvanes, K. J. (2005). The More the Merrier? The Effect of Family Size and Birth Order on Children's Education. *Quarterly Journal of Economics, 120*(2), 669-700.

Blau, D. M. (1999). The Effect of Income on Child Development. *Review of Economics and Statistics, 81*(2), 261–276.

Block, M. L., Wu, X., Pei, Z., Li, G., Wang, T., Qin, L., Wilson, B., Yang, J., Hong, J. S., & Veronesi, B. (2004). Nanometer Size Diesel Exhaust Particles are Selectively Toxic to Dopaminergic Neurons: The Role of Microglia, Phagocytosis, and NADPH Oxidase. *The FASEB Journal, 18*(13), 1618–1620.

Bondy, M., Roth, S., & Sager, L. (2020). Crime is in the Air: The Contemporaneous Relationship between Air Pollution and Crime. *Journal of the Association of Environmental and Resource Economists, 7*(3), 555–585.

Booth, A. L., and Kee, H. J. (2009). Birth Order Matters: The Effect of Family Size and Birth Order on Educational Attainment. *Journal of Population Economics, 22*(2), 367–397.

Borgoni, R., Michelangeli, A., & Pontarollo, N. (2018). The Value of Culture to Urban Housing Markets. *Regional Studies, 52*(12), 1672–1683.

Borowiecki, K. J. (2009). The Determinants of House Prices and Construction: An Empirical Investigation of the Swiss Housing Economy. *International Real Estate Review, 12*(3), 193-220.

Bound, J., Jaeger, D. A., & Baker, R. M. (1995). Problems with Instrumental Variables Estimation when the Correlation between the Instruments and the Endogeneous Explanatory Variable is Weak. *Journal of the American Statistical Association, 90*(430), 443-450.

Bresson, G., & Hsiao, C. (2011). A Functional Connectivity Approach for Modeling Cross-sectional Dependence with an Application to the Estimation of Hedonic Housing Prices in Paris. *aStA Advances in Statistical Analysis, 95*(4), 501–529.

Buoli, M., Grassi, S., Caldiroli, A., Carnevali, G. S., Mucci, F., Iodice, S., Cantone, L., Pergoli, L., & Bollati, V. (2018). Is There a Link between Air Pollution and Mental Disorders? *Environment International, 118*, 154–168.

Burkhardt, J., Bayham, J., Wilson, A., Carter, E., Berman, J. D., O'Dell, K., Ford, B., Fischer, E. V., & Pierce, J. R. (2019). The Effect of Pollution on Crime: Evidence from Data on Particulate Matter and Ozone. *Journal of Environmental Economics and Management, 98*, 102267.

Campbell, J. Y., Giglio, S., & Pathak, P. (2011). Forced Sales and House Prices. *American Economic Review, 101*(5), 2108–2131.

Cao, Z., Zhou, J., Li, M., Huang, J., & Dou, D. (2023). Urbanites' Mental Health Undermined by Air Pollution. *Nature Sustainability, 6*, 470-478.

Card, D., and Dahl, G. B. (2011). Family Violence and Football: The Effect of Unexpected Emotional Cues on Violent Behavior. *The Quarterly Journal of Economics, 126*(1), 103–143.

Casas, J. B., & Benuto, L. T. (2022). Work-related Traumatic Stress Spillover in First Responder Families: A Systematic Review of the Literature. *Psychological Trauma: Theory, Research, Practice, and Policy, 14*(2), 209–217.

Case, A., and Katz, L. (1991). The Company You Keep: The Effects of Family and Neighborhood on Disadvantaged Youths. *NBER Working Paper*, 3705.

Cavalcanti, T., Kocharkov, G., and Santos, C. (2020). Family Planning and Development: Aggregate Effects of Contraceptive Use. *The Economic Journal, 131*(634), 624–657.

Cebula, R.J. (2009). The Hedonic Pricing Model Applied to the Housing Market of the City of Savannah and its Savannah Historic Landmark District. *The Review of Regional Studies,* 39(1), 9-22.

Chay, K., & Greenstone, M. (2003). Air Quality, Infant Mortality, and the Clean Air Act of 1970. *NBER Working Paper* 10053.

Chay, K.Y. & Greenstone, M. (2005). Does Air Quality Matter? Evidence from the Housing Market. *Journal of Political Economy*, 113(2), 376-424.

Chay, K., Dobkin, C., & Greenstone, M. (2003). The Clean Air Act of 1970 and Adult Mortality. *Journal of Risk and Uncertainty, 27*(3), 279–300.

Chen, Y., Ebenstein, A., Greenstone, M., & Li, H. (2013). Evidence on the Impact of Sustained Exposure to Air Pollution on Life Expectancy from China's Huai River Policy. *Proceedings of the National Academy of Sciences, 110*(32), 12936–12941.

Chen, S., Oliva, P., & Zhang, P. (2018). Air Pollution and Mental Health: Evidence from China. *NBER Working Paper* 24686.

Cheng, C., Mayer, W. J., & Mayer, Y. (2018). The Effect of Legalizing Retail Marijuana on Housing Values: Evidence from Colorado, *Economic Inquiry*, 56(3), 1585-1601.

Chun, Y. S. (2004). Asian Dust Evens in Korea. *Journal of the Korean Professional Engineers Association, 37*(2), 58–61.

Clark, D. E., Michelbrink, L., Allison, T., & Metz W. C. (1997). Nuclear Power Plants and Residential Housing Prices. *Growth and Change*, 28(4), 496-519.

Cohen, E., & Avieli, N. (2004). Food in Tourism. *Annals of Tourism Research, 31*(4), 755–778.

Conklin, J., Diop, M., & Li, H. (2020). Contact High: The External Effects of Retail Marijuana Establishments on House Prices, *Real Estate Economics*, 48(1), 135-173.

Conley, D., and Glauber, R. (2006). Parental Educational Investment and Children's Academic Risk: Estimates of the Impact of Sibship Size and Birth Order from Exogenous Variation in Fertility. *Journal of Human Resources, 41*(4), 722–737.

Coulson, N. E., Morris, A. C., & Neill, H. R. (2019). Are New Homes Special? *Real Estate Economics*, 784-806.

Currie, J., Davis, L., Greenstone, M., & Walker, R. (2015). Environmental Health Risks and Housing Values: Evidence from 1,600 Toxic Plant Openings and Closings. *American Economic Review*, *105*(2), 678–709.

Currie, J., & Neidell, M. (2005). Air Pollution and Infant Health: What Can We Learn from California's Recent Experience? *Quarterly Journal of Economics, 120*(3), 1003–1030.

Currie, J., Neidell, M., & Schmieder, J. F. (2009). Air Pollution and Infant Health: Lessons from New Jersey. *Journal of Health Economics, 28*(3), 688–703.

Currie, J., & Walker, R. (2011). Traffic Congestion and Infant Health: Evidence from E-zpass. *American Economic Journal: Applied Economics*, *3*(1), 65–90.

Cutler, D., & Miller, G. (2005). The Role of Public Health Improvements in Health Advances: The Twentieth-century United States. *Demography*, *42*(1), 1-22.

Dahl, G. B., and Lochner, L. (2012). The Impact of Family Income on Child Achievement: Evidence from the Earned Income Tax Credit. *American Economic Review*, *102*(5), 1927–1956.

Davis-Kean, P. E. (2005). The Influence of Parent Education and Family Income on Child Achievement: The Indirect Role of Parental Expectations and the Home Environment. *Journal of Family Psychology*, *19*(2), 294–304.

de Oliveira, C., Saka, M., Bone, L., & Jacobs, R. (2023). The Role of Mental Health on Workplace Productivity: A Critical Review of the Literature. *Applied Health Economics and Health Policy, 21*(2), 167-193.

Dearing, E. (2008). Psychological Costs of Growing up Poor. *Annals of the New York Academy of Sciences, 1136*(1), 324–332.

Demichelis, O. P., Grainger, S. A., Burr, L., & Henry, J. D. (2023). Emotion Regulation Mediates the Effects of Sleep on Stress and Aggression. *Journal of Sleep Research, 32*(3), e13787.

Denissen, J. J. A., Butalid, L., Penke, L., and van Aken, M. A. G. (2008). The Effects of Weather on Daily Mood: A Multilevel Approach. *Emotion, 8*(5), 662–667.

Deschênes, O., Greenstone, M., & Shapiro, J. S. (2017). Defensive Investments and the Demand for Air Quality: Evidence from the Nox Budget Program. *American Economic Review, 107*(10), 2958–2989.

Deryugina, T., Heutel, G., Miller, N. H., Molitor, D., & Reif, J. (2019). The Mortality and Medical Costs of Air Pollution: Evidence from Changes in Wind Direction. *American Economic Review*, *109*(12), 4178–4219.

Downes, T. A., & Zabel, J. E. (2002). The Impact of School Characteristics on House Prices: Chicago 1987–1991. *Journal of Urban Economics, 52*(1), 1–25.

Duncan, G. J., Morris, P. A., and Rodrigues, C. (2011). Does Money Really Matter? Estimating Impacts of Family Income on Young Children's Achievement with Data from Random-assignment Experiments. *Developmental Psychology*, *47*(5), 1263–1279.

Dwyer, P., Scullion, L., Jones, K., McNeill, J., & Stewart, A. B. (2020). Work, Welfare, and Wellbeing: The Impacts of Welfare Conditionality on People with Mental Health Impairments in the UK. *Social Policy & Administration*, *54*(2), 311-326.

Ebenstein, A. (2012). The Consequences of Industrialization: Evidence from Water Pollution and Digestive Cancers in China. *Review of Economics and Statistics, 94*(1), 186-201.

Ebru, Ç., & Eban, A. (2011). Determinants of House Prices in Istanbul: A Quantile Regression Approach. *Quality & Quantity, 45*(2), 305–317.

Ekblad, M., Gissler, M., Lehtonen, L., and Korkeila, J. (2010). Prenatal Smoking Exposure and the Risk of Psychiatric Morbidity into Young Adulthood. *Archives of General Psychiatry, 67*(8), 841.

EPI. (2016).Yale Environmental Performance Index – 2016. *Yale Center for Environmental Law & Policy.*

Erkuş-Öztürk, H., & Terhorst, P. (2016). Innovative Restaurants in a Mass-tourism City: Evidence from Antalya. *Tourism Management*, *54*, 477–489.

Ermisch, J., and Francesconi, M. (2001). Family Matters: Impacts of Family Background on Educational Attainments. *Economica, 68*(270), 137–156.

Feinstein, L., and Symons, J. (1999). Attainment in Secondary School. *Oxford Economic Papers, 51*(2), 300–321.

Fesselmeyer, E., & Seah, K. Y. (2018). The Effect of Localized Density on Housing Prices in Singapore. *Regional Science and Urban Economics, 68,* 304–315.

Finer, L. B., and Zolna, M. R. (2014). Shifts in Intended and Unintended Pregnancies in the United States, 2001–2008. *American Journal of Public Health, 104*(S1), S43–S48.

Fletcher, M., Gallimore, P., & Mangan, J. (2000). Heteroscedasticity in Hedonic House Price Models. *Journal of Property Research, 17*(2), 93–108.

Font-Ribera, L., Pérez, G., Salvador, J., and Borrell, C. (2007). Socioeconomic Inequalities in Unintended Pregnancy and Abortion Decision. *Journal of Urban Health, 85*(1), 125–135.

Forrest, D. (1991). An Analysis of House Price Differentials between English Regions. *Regional Studies*, *25*(3), 231–238.

Fromkin, V., Rodman, R., and Hyams, N. (2014). *An Introduction to Language, (10th ed)*. Wadsworth Cengage Learning (ISBN-10: 1133310680; ISBN-13: 978-1133310686).

Gibbons, S. (2004). The Costs of Urban Property Crime. *Economic Journal*, 114(499), F441–F463.

Gipson, J. D., Koenig, M. A., and Hindin, M. J. (2008). The Effects of Unintended Pregnancy on Infant, Child, and Parental Health: A Review of the Literature. *Studies in Family Planning*, *39*(1), 18–38.

Glick, P. J., Marini, A., and Sahn, D. E. (2007). Estimating the Consequences of Unintended Fertility for Child Health and Education in Romania: An Analysis Using Twins Data. *Oxford Bulletin of Economics and Statistics, 69*(5), 667–691.

Goetzel, R. Z., Hawkins, K., Ozminkowski, R. J., & Wang, S. (2003). The Health and Productivity Cost Burden of the "Top 10" Physical and Mental Health Conditions Affecting Six Large US Employers in 1999. *Journal of Occupational and Environmental Medicine, 45*(1) 5-14.

Goodman, A. C., & Thibodeau, T. G. (1995). Age-related Heteroskedasticity in Hedonic House Price Equations. *Journal of Housing Research*, *6*(1), 25–42.

Goodman, A. C. & Thibodeau, T. G. (2003). Housing Market Segmentation and Hedonic Prediction Accuracy, *Journal of Housing Economics*, 12(3), 181-201.

Greenstone, M. (2017). The Continuing Impact of Shervin Rosen's "Hedonic Prices and Implicit Markets". *Journal of Political Economy*, *125*(6), 1891–1902.

Greenstone, M., & Gallagher, J. (2008). Does Hazardous Waste Matter? Evidence from the Housing Market and the Superfund Program. *Quarterly Journal of Economics, 123*(3), 951–1003.

Gronau, W., & Kagermeier, A. (2007). Key Factors for Successful Leisure and Tourism Public Transport Provision. *Journal of Transport Geography, 15*(2), 127–135.

Gu, H., Yan, W., Elahi, E., & Cao, Y. (2020). Air Pollution Risks Human Mental Health: An Implication of Two-stages Least Squares Estimation of Interaction Effects. *Environmental Science and Pollution Research, 27*(2), 2036–2043.

Gunnell, D., Kidger, J., & Elvidge, H. (2018). Adolescent Mental health in Crisis. *BMJ,* 361: k2608.

Hamidi, S., Bonakdar, A., Keshavarzi, G., & Ewing, R. (2020). Do Urban Design Qualities Add to Property Values? An Empirical Analysis of the Relationship between Urban Design Qualities and Property Values. *Cities*, *98*, 102564.

Hanushek, E. A. (1992). The Trade-off between Child Quantity and Quality. *Journal of Political Economy, 100*(1), 84–117.

Harrison, D., & Rubinfeld, D. L. (1978). Hedonic Housing Prices and the Demand for Clean Air. *Journal of Environmental Economics and Management, 5*(1), 81–102.

Hayashino, Y., Yamazaki, S., Takegami, M., Nakayama, T., Sokejima, S., & Fukuhara, S. (2010). Association between Number of Comorbid Conditions, Depression, and Sleep Quality using the Pittsburgh Sleep Quality Index: Results from a Population-based Survey. *Sleep Medicine, 11*(4), 366–371.

Hayward, G. M., and Rybińska, A. (2017). “Super Bowl Babies”: Do Counties with Super Bowl Winning Teams Experience Increases in Births Nine Months Later? *Socius, 3,* 1–14.

Hendershott, P.H., Kim, K.H., Lee, J.M. & Shilling, J.D. (2021) Announcement Effects: Taxation of Housing Capital Gains in Seoul. *Journal of Real Estate Finance and Economics*, 62(3), 319–341.

Henshaw, S. K. (1998). Unintended Pregnancy in the United States. *Family Planning Perspectives*, *30*(1), 24–29.

Herrnstadt, E., Heyes, A., Muehlegger, E., & Saberian, S. (2021). Air Pollution and Criminal Activity: Microgeographic Evidence from Chicago. *American Economic Journal: Applied Economics*, *13*(4), 70–100.

Hill, J. (2002). Biological, Psychological and Social Processes in the Conduct Disorders. *Journal of Child Psychology and Psychiatry*, *43*(1), 133–164.

Hollingsworth, A., & Rudik, I. (2021). The Effect of Leaded Gasoline on Elderly Mortality: Evidence from Regulatory Exemptions. *American Economic Journal: Economic Policy*, *13*(3), 345–373.

Hong, S. I., & Lim, S. Y. (2018). A Study on the Change of the Housing Culture Koreans through the First-generation Apartment in Korea. *Asia-pacific Journal of Multimedia S ervices Convergent with Art, Humanities, and Sociology,* 8(4), 985-998.

Horwitz, S. M., Chamberlain, P., Landsverk, J., & Mullican, C. (2010). Improving the Mental Health of Children in Child Welfare through the Implementation of Evidence-based Parenting Interventions. *Administration and Policy in Mental Health and Mental Health Services Research*, *37*, 27-39.

Hossain, B., & Latif, E. (2009). Determinants of Housing Price Volatility in Canada: A Dynamic Analysis. *Applied Economics*, *41*(27), 3521–3531.

Jalan, J., & Ravallion, M. (2003). Does Piped Water Reduce Diarrhea for Children in Rural India? *Journal of Econometrics*, *112*(1), 153-173.

Johns Hopkins Medicine. (2023). Conduct Disorder. https://www.hopkinsmedicine.org/health/conditions-and-diseases/conduct-disorder

Johnson, D., Dupuis, G., Piche, J., Clayborne, Z., & Colman, I. (2018). Adult Mental Health Outcomes of Adolescent Depression: A Systematic Review. *Depression and Anxiety, 35*(8), 700-716.

Joo, Y. S., Kim, J., Lee, J., & Chung, I. J. (2022). Fine Particulate Matter and Depressive Symptoms in Children: A Mediation Model of Physical Activity and a Moderation Model of Family Poverty. *SSM - Population Health*, *17*, 101015.

Jung, H., & Lee, J. (2017). The Effects of Macroprudential Policies on House Prices: Evidence from an Event Study using Korean Real Transaction Data. *Journal of Financial Stability*, *31*, 167–185.

Kain, J.F., & Quigley, J.M. 1970. Measuring the Value of Housing Quality. *Journal of the American Statistical Association*, 65(330), 532-548.

Kalist, D. E., and Lee, D. Y. (2016). The National Football League. *Journal of Sports Economics*, *17*(8), 863–882.

Kelly, F. J. (2003). Oxidative Stress: Its Role in Air Pollution and Adverse Health Effects. *Occupational and Environmental Medicine*, *60*(8), 612–616.

Kiely, K. M., Brady, B., & Byles, J. (2019). Gender, Mental Health and Ageing. *Maturitas*, *129*, 76–84.

Kim, C. W., Phipps, T. T., & Anselin, L. (2003). Measuring the Benefits of Air Quality Improvement: A Spatial Hedonic Approach. *Journal of Environmental Economics and Management*, *45*(1), 24–39.

Kim, D., Kim, J., Jeong, J., & Choi, M. (2019). Estimation of Health Benefits from Air Quality Improvement using the MODIS AOD Dataset in Seoul, Korea. *Environmental Research*, *173*, 452–461.

Kim, S. E., Xie, Y., Dai, H., Fujimori, S., Hijioka, Y., Honda, Y., Hashizume, M., Masui, T., Hasegawa, T., Xu, X., Yi, K., & Kim, H. (2020). Air Quality Co-benefits from Climate Mitigation for Human Health in South Korea. *Environment International, 136*, 105507.

King, J. D., Zhang, S., & Cohen, A. (2022). Air Pollution and Mental Health: Associations, Mechanisms and Methods. *Current Opinion in Psychiatry, 35*(3), 192–199.

Kuroda, S., & Yamamoto, I. (2018). Good Boss, Bad Boss, Workers' Mental Health and Productivity: Evidence from Japan. *Japan and the World Economy, 48*, 106-118.

Kwak, S. J., Cho, J. I., & Choi, E. C. (2022). Fine Dust is Coming Again: The Effect of Air Pollution on Health Using Seasonal Weather Patterns. *The Journal of Korean Data Analysis Society, 24*(5), 1625–1637.

Lamb, C., & Murphy, M. (2013). The Divide between Child and Adult Mental Health Services: Points for Debate. *British Journal of Psychiatry, 202*(s54), s41-s44.

Lee, C. J. (2005). Korean Education Fever and Private Tutoring, *KEDI Journal of Education Policy, 2*(1), 99-107.

Lee, J. (2007). Sibling Size and Investment in Children's Education: An Asian Instrument. *Journal of Population Economics*, *21*(4), 855–875.

Lee, J., and Kang, S. (2018). Perceived Helicopter Parenting and Korean Emerging Adults' Psychological Adjustment: The Mediational Role of Parent–Child Affection and Pressure from Parental Career Expectations. *Journal of Child and Family Studies, 27*(11), 3672–3686.

Lee, K., & Choi, G. H. (2016). Determinants of House Prices using Hedonic Price Model: By Analyzing Ichon, Seoul (in Korean). *2016 Korean Urban Management Association Annual Meeting.*

Leiter, M. P., & Durup, M. J. (1996). Work, Home, and In-between: A Longitudinal Study of Spillover. *The Journal of Applied Behavioral Science, 32*(1), 29-47.

Lieske, S. N., van den Nouwelant, R., Han, J. H., & Pettit, C. (2021). A Novel Hedonic Price Modelling Approach for Estimating the Impact of Transportation Infrastructure on Property Prices. *Urban Studies*, *58*(1), 182–202.

Lin, N. F., Tang, J., & Ismael, H. S. M. (2000). Study on Environmental Etiology of High Incidence Areas of Liver Cancer in China. *World Journal of Gastroenterology*, *6*(4), 572-576.

Lin, W. H., Pan, W. C., & Yi, C. C. (2019). "Happiness in the Air?" The Effects of Air Pollution on Adolescent Happiness. *BMC Public Health*, *19*, 795.

Linden, L. & Rockoff, J. E. (2008). Estimates of the Impact of Crime Risk on Property Values from Megan's Laws. *American Economic Review*, 98(3), 1103–1127.

Lopuszanska, U., & Samardakiewicz, M. (2020). The Relationship between Air Pollution and Cognitive Functions in Children and Adolescents: A Systematic Review. *Cognitive and Behavioral Neurology*, *33*(3), 157–178.

Lu, W., Hackman, D. A., & Schwartz, J. (2021). Ambient Air Pollution Associated with Lower Academic Achievement among US Children. *Environmental Epidemiology*, *5*(6) e174.

Lu, Y., Song, S., Wang, R., Liu, Z., Meng, J., Sweetman, A. J., Jenkins, A., Ferrier, R. C., Li, H., Luo, W., & Wang, T. (2015). Impacts of Soil and Water Pollution on Food Safety and Health Risks in China. *Environment International*, *77*, 5-15.

Luttik, J. (2000). The Value of Trees, Water and Open Space as Reflected by House Prices in the Netherlands. *Landscape and Urban Planning*, *48*(3-4), 161–167.

Lynch, A. K., & Rasmussen, D. W. (2001). Measuring the Impact of Crime on House Prices. *Applied Economics*, *33*(15), 1981–1989.

Ma, Y., Siu, A., and Tse, W. S. (2018). The Role of High Parental Expectations in Adolescents' Academic Performance and Depression in Hong Kong. *Journal of Family Issues*, *39*(9), 2505–2522.

Maani, S. A., and Kalb, G. (2007). Academic Performance, Childhood Economic Resources, and the Choice to Leave School at Age 16. *Economics of Education Review, 26*(3), 361–374.

Marston, C., and Cleland, J. (2003). Do Unintended Pregnancies Carried to Term Lead to Adverse Outcomes for Mother and Child? An Assessment in Five Developing Countries. *Population Studies, 57*(1), 77–93.

McHale, S. M., Crouter, A. C., & Tucker, C. J. (2001). Free-time Activities in Middle Childhood: Links with Adjustment in Early Adolescence. *Child Development, 72*(6), 1764–1778.

Mckercher, B., & Lau, G. (2008). Movement Patterns of Tourists within a Destination. *Tourism Geographies, 10*(3), 355–374.

McLeod, J. D., & Shanahan, M. J. (1996). Trajectories of Poverty and Children's Mental Health. *Journal of Health and Social Behavior, 37*(3), 207–220.

McLeod, G. F., Horwood, L. J., & Fergusson, D. M. (2016). Adolescent Depression, Adult Mental Health and Psychosocial Outcomes at 30 and 35 years. *Psychological medicine, 46*(7), 1401-1412.

Miles, S. B., & Stipek, D. (2006). Contemporaneous and Longitudinal Associations between Social Behavior and Literacy Achievement in a Sample of Low-income Elementary School Children. *Child Development, 77*(1), 103–117.

Millimet, D. L., and Wang, L. (2011). Is the Quantity-Quality Trade-Off a Trade-Off for All, None, or Some? *Economic Development and Cultural Change, 60*(1), 155–195.

Min, K.-H., & Lee, T. J. (2014). Customer Satisfaction with Korean Restaurants in Australia and Their role as Ambassadors for Tourism Marketing. *Journal of Travel & Tourism Marketing, 31*(4), 493–506.

Ministry of Public Administration and Security (2019), February 29th Government Press Release.

Montero, J.-M., & Fernández-Avilés, G. (2014). Hedonic Price Model. *Encyclopedia of Quality of Life and Well-Being Research*, 2834–2837.

Montesinos, J., Cortes, J., Arnau, A., Sanchez, J. A., Elmore, M., Macia, N., and Bosch, J. (2013). Barcelona Baby Boom: Does Sporting Success Affect Birth Rate? *BMJ, 347*, 7938.

Morales-Suarez-Varela, M. M., Llopis-Gonzalez, A., & Tejerizo-Perez, M. L. (1995). Impact of Nitrates in Drinking Water on Cancer Mortality in Valencia, Spain. *European Journal of Epidemiology, 11*(1), 15-21.

Munyo, I., and Rossi, M. A. (2013). Frustration, Euphoria, and Violent Crime. *Journal of Economic Behavior and Organization, 89*, 136–142.

Murphy, A. (2018). A Dynamic Model of Housing Supply, *American Economic Journal: Economic Policy*, 10(4), 243-267.

Nanda, A., & Ross, S. L. (2012). The Impact of Property Condition Disclosure Laws on Housing Prices: Evidence from an Event Study using Propensity Scores. *The Journal of Real Estate Finance and Economics, 45*(1), 88–109.

Ng, F., Berk, M., Dean, O., & Bush, A. I. (2008). Oxidative Stress in Psychiatric Disorders: Evidence Base and Therapeutic Implications. *The International Journal of Neuropsychopharmacology, 11*(6), 851–876.

Nigg, J. T., and Breslau, N. (2007). Prenatal Smoking Exposure, Low Birth Weight, and Disruptive Behavior Disorders. *Journal of the American Academy of Child & Adolescent Psychiatry, 46*(3), 362–369.

Nilsson, J. P. (2017). Alcohol Availability, Prenatal Conditions, and Long-term Economic Outcomes. *Journal of Political Economy, 125*(4), 1149–1207.

Nitsch, A., Faurie, C., and Lummaa, V. (2013). Are Elder Siblings Helpers or Competitors? Antagonistic Fitness Effects of Sibling Interactions in Humans. *Proceedings of the Royal Society B: Biological Sciences*, *280*(1750), 20122313.

OECD. (2016). Economic Consequences of Outdoor Air Pollution. Organization for Economic Co-operation and Development.

O'Connor, M. J., and Paley, B. (2009). Psychiatric Conditions Associated with Prenatal Alcohol Exposure. *Developmental Disabilities Research Reviews*, *15*(3), 225–234.

O'Connor, M. J., Shah, B., Whaley, S., Cronin, P., Gunderson, B., and Graham, J. (2002). Psychiatric Illness in a Clinical Sample of Children with Prenatal Alcohol Exposure. *The American Journal of Drug and Alcohol Abuse, 28*(4), 743–754.

O'Leary, K., Bylsma, L. M., & Rottenberg, J. (2017). Why Might Poor Sleep Quality Lead to Depression? A Role for Emotion Regulation. *Cognition and Emotion, 31*(8), 1698-1706.

Osland, L. (2010). An Application of Spatial Econometrics in Relation to Hedonic House Price Modeling. *Journal of Real Estate Research, 32*(3), 289–320.

Paris, S. G., Lawton, T. A., Turner, J. C., & Roth, J. L. (1991). A Developmental Perspective on Standardized Achievement Testing. *Educational Researcher, 20*(5), 12–20.

Park, J., Lee, D., Park, C., Kim, H., Jung, T., & Kim, S. (2017). Park Accessibility Impacts Housing Prices in Seoul. *Sustainability, 9*(2), 185.

Park, J. H. (2022). Crisis in the Tourism Industry due to COVID-19 and Future Countermeasures. *International Journal of Tourism and Hospitality Research, 36*(1), 19-29.

Park, J. W., & Hwang, B. D. (2017). Analysis of Effects of Regional Air Pollutants (PM10, O3, CO) on Respiratory Disease. *Korean Public Health Research, 43*(3), 53–66.

Park, K. A. (2019). An Event Study in Relative Prices and Choice of Loan Term. *Journal of Housing Economics, 46*, 101637.

Park, S., & Shin, H. J. (2017). Analysis of the Factors Influencing PM2.5 in Korea: Focusing on Seasonal Factors. *Journal of Environmental Policy and Administration*, *25*(1), 227–248.

Paul, M., Ford, T., Kramer, T., Islam, Z., Harley, K., & Singh, S. (2013). Transfers and Transitions between Child and Adult Mental Health Services. *British Journal of Psychiatry, 202*(S54), S36-S40.

Persico C., & Marcotte D. V. (2022). Air Quality and Suicide. NBER Working Paper #30626

Peterson, B. S., Rauh, V. A., Bansal, R., Hao, X., Toth, Z., Nati, G., Walsh, K., Miller, R. L., Arias, F., Semanek, D., & Perera, F. (2015). Effects of Prenatal Exposure to Air Pollutants (Polycyclic Aromatic Hydrocarbons) on the Development of Brain White Matter, Cognition, and Behavior in Later Childhood. *JAMA Psychiatry, 72*(6), 531 – 540.

Polo-Peña, A. I., Frías-Jamilena, D. M., & Rodríguez-Molina, M. Á. (2012). Marketing Practices in the Spanish Rural Tourism Sector and Their Contribution to Business Outcomes. *Entrepreneurship & Regional Development*, *24*(7-8), 503–521.

Pope, J. C. (2008). Fear of Crime and Housing Prices: Household Reactions to Sex Offender Registries. *Journal of Urban Economics*, 64(3), 601–614.

Power, E., Hughes, S., Cotter, D., & Cannon, M. (2020). Youth Mental Health in the Time of Covid-19. *Irish Journal of Psychological Medicine, 37*(4), 301–305.

Powers, S. I., Hauser, S. T., & Kilner, L. A. (1989). Adolescent Mental Health. *American Psychologist, 44*(2), 200–208.

Puhani, P. A. (2012). The Treatment Effect, the Cross Difference, and the Interaction Term in Nonlinear “Difference-in-Differences” Models. *Economics Letters, 115*(1), 85–87.

Quinton, D., Pickles, A., Maughan, B., & Rutter, M. (1993). Partners, Peers, and Pathways: Assortative Pairing and Continuities in Conduct Disorder. *Development and Psychopathology, 5*(4), 763–783.

Rahu, K., Rahu, M., Pullmann, H., and Allik, J. (2010). Effect of Birth Weight, Maternal Education and Prenatal Smoking on Offspring Intelligence at School Age. *Early Human Development, 86*(8), 493–497.

Rees, D. I., and Schnepel, K. T. (2009). College Football Games and Crime. *Journal of Sports Economics, 10*(1), 68–87.

Repetti, R. L., Taylor, S. E., and Seeman, T. E. (2002). Risky Families: Family Social Environments and the Mental and Physical Health of Offspring. *Psychological Bulletin, 128*(2), 330–366.

Richardson, H. W., Vipond, J., & Furbey, R. A. (1974). Determinants of Urban House prices. *Urban Studies, 11*(2), 189–199.

Ridker, R. G. & Henning, J. A. (1967). The Determinants of Residential Property Values with Special Reference to Air Pollution. *Review of Economics and Statistics*, 49(2), 246-257.

Risom, L., Møller, P., & Loft, S. (2005). Oxidative Stress-induced DNA Damage by Particulate Air Pollution. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis, 592*(1-2), 119–137.

Roberts, S., Arseneault, L., Barratt, B., Beevers, S., Danese, A., Odgers, C. L., Moffitt, T. E., Reuben, A., Kelly, F. J., & Fisher, H. L. (2019). Exploration of NO2 and PM2.5 Air Pollution and Mental Health Problems using High-resolution Data in London-based Children from a UK Longitudinal Cohort Study. *Psychiatry Research, 272,* 8–17.

Rosenfield, S., & Mouzon, D. (2012). Gender and Mental Health. *In Handbooks of Sociology and Social Research* (pp. 277–296). Springer.

Rosenzweig, M., and Zhang, J. (2009). Do Population Control Policies Induce more Human Capital Investment? Twins, Birth Weight and China's “One-child” Policy. *Review of Economic Studies*, *76*(3), 1149–1174.

Roushdy, R., Sieverding, M., & Radwan, H. (2012). The Impact of Water Supply and Sanitation on Child Health: Evidence from Egypt. New York Population Council – Poverty, Gender, and Youth Working Paper #24.

Rudolf, R., & Bethmann, D. (2022). The Paradox of Wealthy Nations’ Low Adolescent Life Satisfaction. *Journal of Happiness Studies*, *24*(1), 79–105.

Sanders, J. L., and Brizzolara, M. S. (1982). Relationships between Weather and Mood. *Journal of General Psychology, 107*(1), 155–156.

Sanders, N. J., & Stoecker, C. (2015). Where Have all the Young Men Gone? Using Sex Ratios to Measure Fetal Death Rates. *Journal of Health Economics, 41,* 30–45.

Sareen, J., Afifi, T. O., McMillan, K. A., & Asmundson, G. J. (2011). Relationship between Household Income and Mental Disorders. *Archives of General Psychiatry*, *68*(4), 419–427.

Schwarz, N., and Clore, G. (1983). Mood, Misattribution, and Judgments of Well-being: Informative and Directive Functions of Affective States. *Journal of Personality and Social Psychology*, *45*(3), 513–523.

Shen, Y., & Karimi, K. (2017). The Economic Value of Streets: Mix-scale Spatio-functional Interaction and Housing Price Patterns. *Applied Geography*, *79*, 187–202.

Shields, N., and Hanneke, C. (2008). The Effects of Parental Age and Sibling Configurations on Family Environment and Academic Achievement of Children. *Journal of Applied Social Science, 2*(1), 13–35.

Shier, V., Nicosia, N., Shih, R., & Datar, A. (2019). Ambient Air Pollution and Children’s Cognitive Outcomes. *Population and Environment*, *40*(3), 347–367.

Shin, D. J. D., Lee, H. J., Ha, J. E., Park, J. H., Son, E., and Bong, M. (2018). Why aren't Korean Students Happy? Tracing Back to the Sources of Their Academic Distress 1. *In Asian Education Miracles* (pp. 124-138). Routledge.

Simmons, J. (1984). Railways, Hotels, and Tourism in Great Britain 1839-1914. *Journal of Contemporary History*, *19*(2), 201–222.

Skantz, T., & Strickland, T. (1987). House Prices and a Flood Event: An Empirical Investigation of Market Efficiency. *Journal of Real Estate Research, 2*(2), 75–83.

Slack, K., & Webber, M. (2007). Do We care? Adult Mental Health Professionals' Attitudes towards Supporting Service Users' Children. *Child & Family Social Work, 13*(1), 72–79.

Smith, S. L. J. (1983). Restaurants and Dining Out. *Annals of Tourism Research, 10*(4), 515–549.

Smith, V. K. & Huang, J. C. (1993). Hedonic Models and Air Pollution: Twenty-five Years and Counting. *Environmental and Resource Economics*, 3(4), 381–394

South Coast AQMD. (2016). 2016 Air Quality Data Tables. Available at https://www.aqmd.gov/home/air-quality/historical-air-quality-data/historical-data-by-year

Staiger, D., & Stock, J. H. (1997). Instrumental Variables Regression with Weak Instruments. *Econometrica, 65*(3), 557-586.

Statistics Korea (2021), Monthly, Quarterly, and Yearly Population Status.

Statistics Korea. (2019). 2019 Primary and Secondary School Private Tutoring Expenditure Survey. https://kostat.go.kr/board.es?mid=a10301070100&bid=245&act=view&list_no=381064

Statistics Korea (2006), 2006 Birth Statistics Results.

Statistics Korea (2023), Definition of Total Fertility Rate. https://kostat.go.kr/statTerm.es?act=view&mid=a10503000000&nPage=1&stat_term_cd=SL4289

Strohschein, L. (2005). Household Income Histories and Child Mental Health Trajectories. *Journal of Health and Social Behavior*, *46*(4), 359–375.

Sui, G., Liu, G., Jia, L., Wang, L., & Yang, G. (2018). The Association between Ambient Air Pollution Exposure and Mental Health Status in Chinese Female College Students: A Cross-sectional Study. *Environmental Science and Pollution Research*, *25*(28), 28517–28524.

Szyszkowicz, M., Zemek, R., Colman, I., Gardner, W., Kousha, T., & Smith-Doiron, M. (2020). Air Pollution and Emergency Department Visits for Mental Disorders among Youth. *International Journal of Environmental Research and Public Health*, *17*(12), 4190.

Tremblay, P. (1998). The Economic Organization of Tourism. *Annals of Tourism Research, 25*(4), 837–859.

Wang, Q., & Yang, Z. (2016). Industrial Water Pollution, Water Environment Treatment, and Health Risks in China. *Environmental Pollution, 218*, 358-365.

Wenzel, H. G., Bakken, I. J., Johansson, A., Götestam, K. G., & Øren, A. (2009). Excessive Computer Game Playing among Norwegian Adults: Self-reported Consequences of Playing and Association with Mental Health Problems. *Psychological Reports*, *105*(3), 1237–1247.

Wickham, S., Bentley, L., Rose, T., Whitehead, M., Taylor-Robinson, D., & Barr, B. (2020). Effects on Mental Health of a UK Welfare Reform, Universal Credit: A Longitudinal Controlled Study. *The Lancet Public Health, 5*(3), e157-e164.

World Bank. (2006). Pakistan Strategic Country Environmental Assessment. World Bank Report #36946

Wooldridge, J. M. (2011). *Econometric Analysis of Cross Section and Panel Data.* MIT Press.

Yanik, M., Erel, O., & Kati, M. (2004). The Relationship between Potency of Oxidative Stress and Severity of Depression. *Acta Neuropsychiatrica, 16*(4), 200–203.

Yoo, H. G., Hong, J. W., Hong, J., Sung, S., Yoon, E. J., Park, J.-H., & Lee, J.-H. (2020). Impact of Meteorological Conditions on the PM2.5 and PM10 Concentrations in Seoul. *Journal of Climate Change Research, 11*(5-2), 521–528.

Zietz, J., Zietz, E. N., & Sirmans, G. S. (2008). Determinants of House Prices: A Quantile Regression Approach. *The Journal of Real Estate Finance and Economics, 37*(4), 317–333.

Korean Newspapers and Media

"'Again 2002'… Street Gatherings for U-20 World Cup." *The Ministry of Culture, Sports, and Tourism*, May 8, 2017. https://www.korea.kr/news/policyNewsView.do?newsId=148833298.

"Be Aware of World Cup Babies, Korean Men and Women Could not Hide Their Passion…." *The Korea Economic Daily*, June 24, 2014. https://www.hankyung.com/society/article/2014062317767.

"[Business] The 'World Cup Side Effects' at the Production Sites." *Dong-A-Ilbo*, July 4, 2002a. https://n.news.naver.com/mnews/article/020/0000138986?sid=101.

"Four Dangerous Outcomes of the 'World Cup Baby Boom'." *Chosun Ilbo*, June 25, 2010. https://m.health.chosun.com/svc/news_view.html?contid=2010062501478&.

"'Gangnam Style' Boosts S. Korean Tourism." *VOA (Voice of America)*, November 16, 2012. https://www.voanews.com/a/gangman-style-boosting-toruisn-in-south-korea/1547441.html

"Gangnam Brings Fans—and Tourism Revenue—to Korea." *CNBC (Consumer News and Business Channel)*, January 23, 2013. https://www.cnbc.com/id/100399553

"[Health] Many People Suffer World Cup Side Effects." *Dong-A-Ilbo,* June 16, 2002b, https://n.news.naver.com/mnews/article/020/0000137572?sid=102.

Korea Tourism Organization, (2012). 'Gangnam Style' is Indeed Powerful. October 10, 2012. https://kto.visitkorea.or.kr/kor/notice/news/press/board/view.kto?id=416775&rnum=1641

"Miraculous Win and Miraculous Night: Again 2002 World Cup Baby Boom?" *Seoul Broadcasting Service*, June 25, 2010. https://news.sbs.co.kr/news/endPage.do?news_id=N1000762503.

"Unexpected, More Pregnancy and 'World Cup Babies'." *Seoul Broadcasting Service*, April 7, 2003. https://news.sbs.co.kr/news/endPage.do?news_id=N0311404679.

"Where is the Public Order?" Yonhap News Agency, July 31, 2002, https://n.news.naver.com/mnews/article/001/0000214221?sid=102.