IZA Institute of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 16003

Sejong's Effects on People's Health: Consequences of a Long Commute

Seulgi Lim
Soohyung Lee

MARCH 2023

IZA Institute of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 16003

Sejong's Effects on People's Health: Consequences of a Long Commute

Seulgi Lim
National Assembly Budget Office

Soohyung Lee
Seoul National University, AASLE and IZA

MARCH 2023

ISSN: 2365-9793

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

Sejong's Effects on People's Health: Consequences of a Long Commute*

We examine the health impacts of long commute time by exploiting a large-scale placed-based policy in South Korea. The policy relocated public employers in the capital area to disadvantaged cities. However, some public employees kept their residences in the capital area and spend long hours commuting. Using this change, we estimate 2SLS models whose results suggest that having a long commute substantially increases usage of medical services, particularly to treat respiratory, circulatory, and endocrine & metabolic diseases. However, we find mixed effects of long commute time on medical checkup outcomes and health-related activities such as exercise.

JEL Classification: I10, J18, H51, R11
Keywords: commute, health, place-based policy, Sejong, innovation city

Corresponding author:
Soohyung Lee
Associate Professor
Graduate School of International Studies
Seoul National University
1 Gwanak-ro
Gwanak-gu
Seoul 08826
Korea
E-mail: soohlee@snu.ac.kr

* We have benefitted from detailed comments by Prof. Wing T. Woo, Deborah Swenson, Shigeyuki Abe, Joonmo Ahn, Taehyun Ahn, Byunghil Jun, Kyung-Hwan Kim, Pramod Kumar, Minjung Park, Xuezhu Shi, Seokjin Woo, Sifan Zhou, and participants of various seminars. All errors are our own. This research was supported by the Ministry of Education of the Republic of Korea and the National Research Foundation of Korea (NRF- 2021S1A5A2A01065004).

I. Introduction

Many workers spend a significant portion of their work week commuting. For example, in 2019, the average American commuter spent 28 minutes on daily one-way commutes, a 10% increase from 2006. Travel time to work is much longer for residents in metropolitan areas (see Burd et al., 2021). The phenomenon of long commute time is not limited to developed countries. A large number of people in developing countries are concentrated in a few urban agglomerates, and they are subject to long commutes due to limited availability of transportation, poor infrastructure, and congestion (Razvadauskas, 2018).[1] The fraction of people who spend at least one hour on daily one-way commutes is reported as 42% in Bangkok, 53% in Jakarta, 47% in Istanbul, 31% in Buenos Aires, and 41% in Bogota (Moovitapp, 2021).

Despite the prevalence of long commute time around the world, there are only a limited number of studies estimating its causal impact on health. An extensive number of studies by non-economists have examined the role of commute time and the mode of commute in accounting for subjective wellbeing and objective health outcomes. A long commute has been reported to cause stress, fatigue, and an insufficient amount of sleep, worsening workers' health conditions (Gangwisch et al., 2005; Walsleben et al., 1999; Pfeifer, 2018; Stutzer and Frey, 2008; Kageyama et al., 1998; Schaefer et al., 1988; Novaco et al., 1990). However, these studies use cross-sectional analyses to report correlation and do not investigate the causal impact. Recently, a growing number of economics studies have aimed to identify a causal relationship between commute time and workers' health. Examples include Goerke and Lorenz (2017), Kunn-Nelen (2016), Martin et al., (2014), Pfeifer (2018), and Jacob et al. (2021). Those studies use individual-level panel data in the U.S., U.K., and Germany, and identify the causal effect by controlling for individual fixed effects.

[1] As of 2017, there were 33 megacities, that is, cities whose population exceeds 10 million, around the world; 26 of them are located in developing countries. See details in Razvadauskas (2018).

In sum, although long commute time is prevalent in large cities across the world and taxing many people, there are only a limited number of studies rigorously estimating its causal impact, and those studies focus on a few developed countries. This paper aims to fill this gap in the literature by exploiting a large-scale development project in South Korea that substantially increased workers' commute time. Like many workers around the world, South Koreans suffer from long commute time. In fact, it is the worst among the OECD countries in terms of average commute time (OECD, 2016).

Starting in 2012, the South Korean national government began relocating its agencies and public sector employers from the capital city area to 11 disadvantaged areas in the rest of the country. A key objective of this policy is to reduce the population concentration in Seoul and its surroundings. To meet this objective, the government intentionally chose the 11 areas and the locations of their fast-train stations to make the daily commute from the capital city area difficult. Nonetheless, a sizable share of workers whose employers were relocated decided to remain in the capital area, and thus they bear a long daily work commute.

Using this shock to commute time, we estimate the causal effect of a long commute time on medical services usage, medical expenses, outcomes of medical checkups, and health-related activities. The Population Census (2010 and 2015) and the National Health Insurance database are our two data sources. The former provides information on commute time, while the latter includes health-related information. Unfortunately, these two datasets cannot be merged at the individual level. Thus, we use the common variables between the two datasets—namely, residence, sex and age group—to define cells and merge the two. We construct a variable measuring the share of workers in each cell who are subject to relocation due to the 2012 policy (i.e., 'at-risk' workers).

Suppose that the initial share of at-risk workers in a cell is positively correlated with the share of public workers who remained in the capital area and have a long commute time. Then, the cells with a high share of at-risk workers in 2010 should experience a greater increase in commute time between 2010 and 2015 relative to other cells with a low share of at-risk workers. Built on this conjecture, we construct a two-stage least square (2SLS) model, using the interaction between the share of at-risk workers prior to the policy intervention (namely 2010) and the post dummy indicating the year 2015 as our instrument. As the 2012 relocation policy was designed to redistribute the population in Seoul and its surroundings, we narrow our sample to those areas.

Our estimation results show that having a long commute time leads to a substantial increase in medical services usage and medical expenses. However, it has a mixed impact on outcomes of medical checkups and health-related activities. Specifically, in a cell for a given year, a 1% pt. increase in the share of long commuters, who spend at least two hours per day commuting, increases the number of workers who visit a hospital at least once by 3.5%, the number of hospital visits by 4.0%, copayments by workers by 5.3%, and medical expenses paid by the National Health Insurance System by 9.1%. Long commute time significantly increases the medical services usage for circulatory, respiratory, endocrine & metabolic, and pregnancy & childbirth-related diseases. For example, for circulatory diseases, a 1% pt. increase in the share of long commuters leads to a 17.1% increase in the number of workers who visit a hospital at least once and a 17.2% increase in the number of hospital visits. As for outcomes of medical checkups, the increase in commute time reduces good cholesterol (HDL), increasing fasting blood sugar levels (FBS) as well as creatinine, which may increase health risks. In contrast, the increase in commute time leads to a reduction in BMI, waist circumference, total cholesterol level, LDL and ALT, which are considered to reduce health risks. These mixed effects can be accounted for by our finding that people in our sample respond to a long commute by decreasing both bad and good health-related activities (e.g., reducing smoking as well as exercise).

In addition to the health economics studies citied above, this paper contributes to the rich economics literature on place-based policies targeting disadvantaged areas, by examining an understudied program, namely Korea's relocation policy. The literature provides economic theories justifying place-based policies and empirical examinations of specific programs as well as their causal effects. See Neumark and Simpson (2014) for an overview of the literature. Although Korea's relocation policy reflects Korea-specific environments, the policy was motivated by equity, which is common to other place-based policies in the literature. This paper demonstrates that Korea's relocation policy has had only limited success in redistributing its population despite massive spending on development projects. For example, building new offices for the public employers cost 10 trillion won, equivalent to 3.4% of the government budget and 0.8% of the GDP in 2010 (NABO, 2016). The total cost is even greater if we include the subsequent costs for infrastructure and administration services. However, this relocation policy has not only failed to reduce the population concentration in the SMA but also poses health risks to workers. This finding is worth sharing with other researchers because it shows that massive spending on infrastructure and relocation of employers may not be sufficient to guarantee a policy's success.

The reminder of our paper proceeds as follows: Section II provides the institutional background, while Section II describes the data and the sample. Sections IV and V present the empirical framework and results, respectively. Section VI concludes.

II. Institutional Background

II.1 Redistribution Policy of Public Employers

Most developing countries show a concentration of population and economic resources in their capital areas (Henderson, 2002). South Korea shares this pattern. Its capital and its surrounding areas, namely Seoul, Incheon and Gyeonggi, are home to 25.9 million people (50% of Korea's population in 2019), and they accounted for 52% of Korea's gross domestic product as of 2019, although they make up only 11.8% of the South Korean territory (Statistics Korea, 2009, 2020). For conciseness, we refer to Seoul, Incheon and Gyeonggi, as the Seoul Metropolitan area (SMA), hereafter. Figure 1 indicates the SMA in gray.

This immense concentration of people and economic activities in a small geographical area has generated various challenges for Koreans, such as traffic jams, soaring house prices, and environmental degradation in the SMA. At the same time, the rest of the country has been losing residents and economic vitality, which has raised political concerns in South Korea for the past several decades.

Since 1964, the Korean government has been trying to deter further concentration in the SMA through various restrictions such as limiting the number of colleges located in the capital area and land use (KERI, 2006). A recurring policy tool the Korean government has been using is relocating public sector employers further away from Seoul and its surroundings. For example, in 1975, a national development plan was established to develop a rural area south of Seoul, called Gwacheon (indicated by the triangle in the upper-left corner of Figure 1), and to relocate 13 ministries of the central government from central Seoul to it.[2] The relocation spanned from 1982 to 1994. In 1990, the government decided to relocated another 11 of its ministries near Daejeon, located in the middle of South Korea (indicated by the triangle in Figure 1), and the relocation started in 1998.[3]

[2] The 13 ministries are Ministry of Health and Welfare, Emergency Planning Commission, Ministry of Science and ICT, Ministry of Transportation, Ministry of Agriculture, Food and Rural Affairs, Ministry of Justice, Ministry of Trade, Industry and Energy, Ministry of Employment and Labor, Ministry of Finance, Ministry of Energy and Resources, Ministry of Economy and Finance, Ministry of Transportation, Ministry of Environment (Ministry of the Interior and Safety, 2019).

[3] The 11 government ministries are the Korea Customs Service, Public Procurement Service, Statistics Korea, Military Manpower Administration, Korean Intellectual Property Office, Korea Forest Service, Korean national Railroad,

In 2004, the government announced another relocation plan on an unprecedented scale. The plan was introduced because President Rho Moo-hyun had made an election promise to establish a new city specializing in public administration. The plan was enacted as a Special Act in 2005, titled "The Special Act on Balanced National Development" (*Gukga KyunHyung Baljeon Bub* in Korean). Unlike its predecessors, this plan applied not only to ministries in the central government but also public agencies and for-profit firms whose majority shareholder is the Korean government. The relocation plan consists of two parts. The first is to establish a new city in the middle of South Korea, called Sejong Special Self-Governing City (Sejong, hereafter), and relocate almost all central government ministries there.[4] The second is to develop rural areas adjacent to existing cities outside of the SMA. There are 10 such areas, which are referred to as "Innovation Cities."[5] These Innovation Cities are designed to host public agencies (e.g., research institutions and the National Pension Service) and for-profit firms whose main shareholder is the Korean government (e.g., Korea Electric Power Corporation and Korea Gas Corporation). Figure 1 shows each relocation area.

At the time of policy announcement in 2005, it was not clear whether the relocation policy would be implemented and, if so, its scope and timeline. Specifically, there was immense political opposition to the plan as well as several lawsuits starting from 2002. Furthermore, the opposition party won the presidential election in December 2007, just 5 months after the construction for Sejong started. The successor, President Lee Myung-bak, was publicly against the relocation policy, especially establishing Sejong, and several attempts to diminish the scope of the relocation plan

---

Ministry of SMEs and Startups, Cultural Heritage Administration, National Archives of Korea, Government Buildings Management Office (Daejeon) (Ministry of the Interior and Safety, 2019).

[4] See the Special Act on the Construction of Administrative City [Act No.7391, Mar.18, 2005].

[5] See the Special Act on the Construction and Support of Innovation Cities Following Relocation of Public Agencies [Act No.8238, Jan.11, 2007].

were made throughout his presidency. In 2012, the final year of Lee's presidency, Sejong was established and recognized as an administrative unit in July. In September it received the first government agency, with only 150 regular employees. As Ms. Park Geun-hyu, the president inaugurated in February 2013, supported the relocation policy, the rest of the government agencies and other public employers continued to relocate to Sejong and other areas. As for the policy's scope, the Ministry of the Interior and Safety initially chose 194 public employers to be relocated in 2005, but the list of employers subject to relocation was revised several times, eventually leading to 170 employers as of 2015.[6] Hereafter, we refer to this relocation policy as the 2012 policy, because the list of public employers was changed multiple times since 2005 and the actual relocation of government agencies started in 2012.

As of 2015, 95% of the targeted employers had been relocated, and the number of their regular employees amounted to 53,652 (0.41% relative to the number of employees in the SMA). By regular employees, we mean full-time employees with unlimited terms, often guaranteed with life-time employment till age 60. Note that the targeted employers also hired individuals as fixed-term workers and part-time workers (often referred to as non-regular workers in Korea). A study estimates the number of non-regular workers to be 30% that of regular workers in the Korean public sector (KIPF, 2013). Applying this estimate, we expect approximately 70,000 workers (and their family members) to be subject to relocation.

II.2 Intended Policy Goal and Impacts on Commute Time

[6] It is worth noting that eight public employers relocated to innovation cities prior to 2012: Customs Human Resources Development Institute, Police human Resources Development Institute, National Institute of Food and Drug Safety Evaluation, Ministry of Food and Drug Safety, Korea Disease Control and Prevention Agency, Korea Health Industry Development Institute, National Institute of Special Education, National Education Training Institute. This relocation decision was made by each employer, not by the central government. For this reason, we do not consider them part of the 2012 relocation policy.

A key policy goal is to relocate the population residing in the SMA to the rest of country, inducing employees of the public sector employers to move to the newly built cities. To ensure their relocation, the government deliberately chose the locations of train and bus stations in the newly built cities such that a trip between the SMA and a newly built city would take a long time. For example, there is no train station in Sejong, and the nearest one is located 16.7 km (10.4 miles) away, taking approximately 15 minutes by car and 30 minutes on public transportation. To gauge the travel time between the SMA and a newly built city, we select two locations—the center of Sejong and the National Assembly—and calculate estimates of travel time between the two. Without any traffic, a one-way trip takes over two and half hours on public transportation and approximately two hours by car.

Despite this policy objective, a significant fraction of the employees have not relocated to the new cities with their families. For example, 27.9% of government officials have remained in the SMA and commute to Sejong (KIPA, 2014). As for the rest of the public sector workers whose employers moved far away from the SMA, 6% commute daily from the SMA, and another 35% stay in the newly built cities only during weekdays but go back to the SMA to join their families on weekends (NABO, 2016). Moreover, even the employees who relocated their entire families to the new cities are not free from long work-related travel times. This is because the key stakeholders of the public employers, such as the National Assembly, the Presidential Office, and the Financial Supervisory Agencies, remain in Seoul. To attend meetings with those stakeholders, some employees travel to Seoul regularly, which equally increases travel time for work. This pattern suggests that the policy may substantially increase the commute time of at least some public sector employees.

Not surprisingly, there have been anecdotal news reports showing that employees in public sector are distressed due to the long commute time and frequent long-distance trips for work. For example, over 86% of government officials in Sejong are concerned about excessive travel and the resulting time costs to the National Assembly (KIPA, 2013). The associated travel costs are estimated to range from 3.57 to 6.72 billion won per year — about USD 2.9 to 5.9 million (KIPA, 2017).[7] Health challenges associated with the long travel times have been also reported in local media. A common complaint among workers is herniated disks in the back and neck due to 3+ hours of commuting time, for which they frequently need to visit medical clinics specializing in rehabilitation and acupuncture (MT, 2015). Some government officials are concerned about the lack of time spent with their family members in the SMA, which generates conflicts at home and taxes their mental health (MT, 2015).

III. Data and Sample

III.1 Data Sources

We rely on two sources to construct a dataset including occupation, commute time, and health outcomes. The first is the supplemental survey on "Commute to Work or School," which is part of the Population Census (2010 and 2015). We obtain the restricted-use version that includes individual level information on demographic variables such as sex, age, residence, educational attainment, and (if the person works) industry.

The second source is the National Health Insurance Service (NHIS). South Korea has a national health insurance system covering all South Koreans, and the NHIS is in charge of it. Hospitals and pharmacies are required to report to the NHIS all medical services eligible for the national insurance to get reimbursement. The NHIS maintains its database based on those reports

[7] The payroll expenses for the employees of the central government were on average 32 trillion won between 2015 and 2017. Relative to the payroll expenses, the travel costs were 0.11 to 0.21 percent (Hankook, 2019).

and constructs a dataset for researchers. The dataset, called the National Sample Cohort (NSC), includes a 2% random sample of the total Korean population in 2006, including their demographic information, medical services usage, and health outcomes between 2002 and 2015. We were granted to access the cross-sectional data including individual-level information for 2010 and 2015. This information includes sex, age, residence, insurance type, use of medical services, and other health measures.

Insurance is classified into five types: employee, dependent of the employee, self-employed, dependent of the self-employed, and medical aid beneficiary. For those who have a job, their insurance types are either employee or self-employed. The dependent status is granted only if the person does not work for pay, regardless of whether he/she has a spouse with health insurance. Medical aid is for those who are under the poverty line set by the NHIS. The NSC records the frequency of hospital visits, and associated costs paid by individuals and by the NHIS, as well as the types of illnesses. In addition, the dataset includes other health measures such as body mass index (BMI), waist circumference, and cholesterol level as well as the surveys on the frequency of health-related activities, such as drinking, smoking, and exercising. These health measures are recorded as a result of health checkups that the NHIS grants to all South Koreans every year for blue-collar workers and every two years for the rest, for early detection and treatment of diseases.

III.2 Sample Construction

Using the Census and NSC data to answer our research question poses two key challenges. One is that there is no individual-level identifier allowing us to link the two data sources. For this reason, we aggregate each data source up to the cell level and merge the two sources for a given year. We define cells based on observables commonly reported in the two data sources, namely sex, age group, and residential location. As our research goal is to measure the impact of the relocation policy on workers in terms of commute time and health outcomes, we restrict our sample to relevant demographic groups, namely those who reside in the SMA and who are full-time employees, excluding self-employees, between 25 and 59 years old.[8] We further narrow our sample only to those who commute to work, by excluding the employees who stay at their workplace (e.g., in a work-dormitory, at a construction site, truck drivers). In total, there are 1,050 cells in our sample.[9]

The other challenge is that neither of the two data sources has information on whether a person works for a public sector employer subject to the 2004 relocation policy. As a result, we cannot directly calculate the fraction of workers who are employed by the targeted public sector employers in each cell. As an alternative, we use the following procedure to create a proxy variable correlated to the risk of relocation for each cell. For each of the 170 public employers subject to the 2012 policy, we obtain the number of its employees as of 2010.[10] We then aggregate the number up to the 3-digit industry level (total of 228 categories) and calculate its share relative to the number of employees in the corresponding industry.[11] Next, for each cell in our sample, we calculate the share of at-risk workers, that is, the average of these shares weighted by the industry composition among the employees in the corresponding cell. To gauge whether this variable is a

[8] We impose this age restriction because it is the prime working age and public sector employers' set retirement age as 60 (Lee, 2014).

[9] There are 2 sexes (men and women), 7 age groups (25-29, 30-34, 35-39, 40-44, 45-49, 50-54, 55-59) and 75 residential locations in the SMA. The product of 2, 7, and 75 yields 1,050.

[10] For the government agencies, we collect the list of employers from the notices posted by the Ministry of the Interior and Safety (October 2005, August 2010, October 2015) and identify the number of employees each agency hired in 2010 (Ministry of the Interior and Safety Notices No. 2005-9, 2010-52, 2015-37). We identify the other employers based on the list from the Ministry of Interior and Safety, and we identify their number of employees from the Ministry of Land, Infrastructure, and Transport. We classify the 170 targeted public employers into the 3-digit industry classification by adopting the definition of industries from NICE biz info. See https://www.nicebizinfo.com/cm/CM0100M001GE.nice.

[11] That share ranges from 0.23 (e.g., activities of head offices & management consulting) to 1 (e.g., mining of hard coal, extraction of crude petroleum, extraction of natural gas).

reasonable proxy, we compare two cities — Gwacheon and Incheon. Gwacheon was home to most central government agencies before 2012, housing many government officials as its residents. In contrast, Incheon is the second largest city in the SMA, but it does not host major central government agencies or public employers. Therefore, if our variable, the share of at-risk workers, captures the risk of relocation, it should be higher in Gwacheon than in Incheon, which in fact is the case. The average share of at-risk workers in 2010 is 3.092 in the cells belonging to Gwacheon, but only 0.931 in the cells belonging to Incheon.

III.3 Summary Statistics

Table 1 shows the summary statistics, weighted by the number of workers in each observation. Panel A reports the statistics regarding commute time and the share of at-risk employees working for the public agencies subject to the relocation. The Korean Census classifies a worker's commute time, each way to/from work, into 7 categories (0-14 minutes, 15-29 minutes, 30-44 minutes, 45-59 minutes, 60-89 minutes, 90-119 minutes, 120 or more minutes). We regard workers as being long commuters if they spend at least one hour each way to go to/from work. The share of long commuters in the SMA was 29.4% in 2010, when the relocation had not yet started, and 31.7% in 2015, when the relocation was almost completed. The share of at-risk employees was 1.1% in 2010 and 1.0% in 2015. This pattern — that the share of at-risk workers is smaller but still greater than zero — is consistent with our report in Section II that some public sector workers relocated out of the SMA while others remained in the SMA.

Panel B of Table 1 reports the demographic characteristics of the workers in our sample. Their average age is 40, and 42% are female. Approximately half of the workers are graduates from 4-year colleges and own houses (47% and 52%, respectively). Sixty-two percent of the workers are currently married, and 60% hold white-collar jobs (e.g., managers, experts, and office workers). We also calculate the share of employees (i.e., workers in our sample), relative to all workers in the corresponding cell, denoted by "Employee(%)." This variable may reflect time-varying economic conditions in each cell because people may switch their work status depending on the economic situation. For example, if a middle-aged person gets laid off, he/she may become a self-employed person, thus decreasing "Employee(%)," because of difficulty in finding a job. These characteristics could affect commute time and health outcomes – such as educational attainment, marital status, homeownership, and occupation (see Kunn-Nelen, 2016; Goerke and Lorenz, 2017). Thus, we use them as control variables in our regression analyses explained in Sections IV and V.

Health outcomes are reported in Panel C. The top rows report the usage of health care services and medical expenses. In a given cell and year, on average, 149 workers (88%) visited a hospital at least once, while 22 workers (12%) never visited a hospital in a given year. There were 1,423 visits in total — costing the NHIS 48.1 million won and patients 17.9 million won.[12]

The middle rows of the table present outcomes from medical checkups. For our assessments, we adopt the criteria provided by the US medical institutes. The National Heart, Lung, and Blood Institute (NHLBI) regards a person as being at low risk for heart diseases and Type 2 diabetes if the person's BMI is the range of 18.5 to 24.9, and his/her waist circumference is less than 35 inches for women and 40 inches for men (NHLBI, 2021). The average BMI in our sample is 23.6, and the average waist circumference is 73cm (28 inches) for women and 84cm (33 inches) for men. Thus, on average, the workers in our sample are healthy. However, we see some warning

[12] That is, in each year, 88% of workers visited a hospital at least once, and the average worker visited a hospital 8 times, costing him/her 104,961 won (approximately US$88) and the NHIS 281,710 won (approximately US$235).

signs regarding cholesterol. Cholesterol levels – total, LDL, HDL – are commonly used to predict the risk of coronary artery diseases. For adults, healthy ranges are 125 to 200mg/dL for total cholesterol, less than 100mg/dL for LDL (bad cholesterol), and 50mg/dL or higher for HDL(good cholesterol) (NHLBI, 2005). In our sample, the average total cholesterol level is 194mg/dL, close to the upper limit of the healthy level (200mg/dL), while the LDL is on average 112mg/dL, exceeding the upper limit (100mg/dL). The average HDL in our sample is rather on the low side (57mg/dL), although it falls within the healthy range. A high level of fasting blood sugar (FBS) indicates diabetes. FBS ranging between 70 and 99 ml/dL is considered normal, whereas FBS exceeding 126 mg/dL indicates diabetes (CDC, 2021). The sample average is 95 ml/dL, which is close to the upper limit of the normal range. A large amount of creatinine in the blood is associated with poor kidney health. The normal range is 0.9 to 1.3 mg/dL for men and 0.6 to 1.1 mg/dL for women (URMC, 2022). As our sample contains 42% women, the weighted normal range for our data is between 0.78 and 1.2 mg/dL. The average in our sample is 0.964, within the normal range. A high value on the Alanine Aminotransferase Test (ALT) indicates liver damage. The normal range of ALT is between 4 and 36 mg/dL (USCF Health, 2022). The sample average is 25.6, considered normal.

The rest of Table 1 reports the statistics regarding people's activities affecting health outcomes. In our sample, 31% of employees drink alcohol at least two times per week, and 27% of employees are smokers. The average number of cigarettes consumed (including non-smokers) is 3.9 cigarettes per day. Regarding exercise that lasts at least 30 minutes with medium intensity, 34% of workers report exercising at least two times per week.

IV. Empirical Framework

IV.1 Model, Causes of Endogeneity, and Identification Strategy

We first present an individual-level regression model in equation (1) to highlight the nature of endogeneity. We then explain the cell-level model that we employ for our estimation. Equation (1) presents a linear model mapping a person's commute time and other explanatory variables to health outcomes ($Health_{iclt}$). Subscript $iclt$ refers to a person i belonging to a demographic category c (i.e., cell) residing in locality l in year t.

$$\begin{aligned} Health_{iclt} &= \alpha + \beta LongCommute_{iclt} + \gamma HR_{icl0} + \delta Post_t \\ &\quad + \boldsymbol{X}'_{iclt}\Lambda + \mu_{lt} + \varepsilon_{iclt} \end{aligned} \tag{1}$$

That person's health outcome is accounted for by five types of explanatory variables: whether the person spends at least 2 hours per day commuting ($LongCommute_{iclt} \in \{0,1\}$), whether, in 2010, the person works for a public employer subject to the 2012 relocation ($HR_{icl0} \in \{0,1\}$), a year dummy indicating 2015 ($Post_t \in \{0,1\}$), individual characteristics ($\boldsymbol{X}'_{iclt}$), and location-by-year fixed effect (μ_{lt}). Vector $\boldsymbol{X}_{clt}$ includes all characteristics of the cell that could affect commutes and health outcomes, reported in Panel B of Table 1. Note that for age, we include dummies for 7 age groups (25-29, 30-34, 35-39, 40-44, 45-49, 50-54, 55-59). Variable ε_{iclt} captures a random shock affecting the person's health. We allow for the random shocks to be correlated across cells that belong to the same location, by clustering at the location level (75 locations in total).

The parameter of interest is β to gauge the causal effect of long commute time on health. However, estimating equation (1) with an OLS may provide a biased estimate if the random shock ε_{iclt} is correlated with $LongCommute_{iclt}$ (i.e., endogeneity). To examine the possible causes of correlation, we specify the error term as the sum of three components: unobserved individual specific shock (ν_i), cell-by-year specific shock (ξ_{clt}), and pure random shock (τ_{iclt}):

$$\varepsilon_{iclt} = \nu_i + \xi_{clt} + \tau_{iclt} \; . \tag{2}$$

The first possible scenario that may generate endogeneity is the case in which people make their locational decision, ultimately affecting their commute time ($LongCommute_{iclt}$), based on their unobserved characteristics (ν_i) that could affect health outcomes. For example, people in good health may be more willing to bear a long commute time. If this is the case, the OLS estimate will be biased downward when a researcher examines outcomes indicating poor health conditions. Alternatively, people who pay close attention to staying in good health may be less willing to bear a long commute time, to reduce the stress on their body.

$$LongCommute_{iclt} = \pi_0 + \pi_1 HR_{icl0} \times Post_t + \pi_2 HR_{icl0} + \pi_3 Post_t + \boldsymbol{X'}_{iclt}\Theta + \rho_{lt} + \omega_{iclt} \quad (3)$$

To address this possible endogeneity, we exploit the 2012 relocation policy and establish the first stage model as equation (3). As the 2012 policy exogenously relocated the public employers far away from the SMA, the public sector employees in the SMA (i.e., $HR_{icl0} = 1$) would be more likely to experience a long commute time in 2015 (i.e., $LongCommute_{iclt} = 1$ with $t = 2015$). For that, we expect π_1 to be positive. If π_1 is statistically significant (relevance) and $HR_{icl0} \times Post_t$ is not correlated with ε_{iclt} (exclusion restriction), we can use equation (3) as our first stage regression. The combination of equations (1) and (3) yields our regression framework based on 2SLS. Note that the remaining variables in equation (3) are the same as in equation (1), while ρ_{lt} and ω_{iclt} capture location by year fixed effects and random shocks to commute time, respectively.

Another possible scenario can generate endogeneity in our setting. Specifically, in each year, people belonging to the same cell may experience a common economic shock that is correlated with their commute time. Specifically, consider the area called Gwacheon, a small-scale city that hosted most central government agencies until 2012. Its residents and retailers expressed concern over the possibility that the relocation policy would hit hard the local demand for goods and services as well as for housing, which generated multiple political demonstrations against the policy. If the economic effects triggered by the 2012 policy are location-specific, then the effects should be captured by the location-by-time fixed effects (μ_{lt}) in our model, not occurring endogeneity. However, the effects can vary by cells in a location. Thus, a person's health outcomes can be indirectly affected by the extent to which his/her peers belonging to the same cell were subject to the 2012 policy. That is, the cell-by-year specific shock ξ_{clt} can be correlated with our instrument ($HR_{icl0} \times Post_t$), which violates the exclusion restriction assumption. We find evidence that this scenario may not be a concern in our setting, which we will report in detail in the subsequent section.

For the estimation, we aggregate equations (1) and (3) up to the cell-by-year level because we do not observe individual-level information, as explained in Section III.2. The regression models remain equivalent to equations (1) and (3) except for the fact that an individual-level variable is now defined as the average across individuals in the corresponding cell. For example, the new dependent variable $Health_{clt}$ is the average of $Health_{iclt}$ across i in the same cell c in a given year t, $LongCommute_{clt}$ is the share of workers commuting at least two hours per day in the cell, and HR_{cl0} is the share of workers in cell c who are subject to the 2012 relocation policy in year 2010.

IV.2 Validity of the Identification Strategy

Our identification strategy relies on two conditions. One is that the variable $HR_{cl0} \times Post_t$ should have a statistically significant correlation with $LongCommute_{clt}$ (i.e., $\pi_1 \neq 0$). The other is that it should not affect health outcomes directly but only through $LongCommute_{clt}$ (i.e., exclusion restriction). The regression results from the first-stage equation can provide information on whether the first condition holds.

Relevance: First-Stage Regression

Column (1) of Table 2 presents the results. The estimate of π_1 is 1.877, statistically different from zero at the 1 percent level. Not only is the estimate different from zero at conventional levels, but it is positive, as we hypothesized. Consider two cells that share the same characteristics except for one dimension. That is, one cell has a one standard deviation (0.507) higher share of high-risk workers than the other in 2010. Then, our estimate implies that the share of long commuters in 2015 will be larger by 1% pts. (i.e., 0.952 = 1.877 x 0.507) in the former cell compared to the latter.

Exclusion Restriction: Falsification Test

To assess the plausibility of our exclusion restriction, we conduct the following empirical test. We are concerned about the possibility that a person's health outcomes may be indirectly affected by the extent to which his/her peers belonging to the same cell were subject to the 2012 policy (i.e., ξ_{clt} is correlated with $HR_{icl0} \times Post_t$). If this concern is relevant to our setting, then our IV will be correlated to the health outcomes of people whose commute time should not be affected by the policy, namely self-employed and retirees. Thus, we access the NHIS data on the health outcomes of self-employed and retirees, which we exclude from our main sample. We then construct their average health outcomes for each cell and use them as dependent variables in our 2SLS regression model (equations (1) and (3)), instead of the health outcomes of those who are employed. Note that the number of cells covered by the self-employed and retirees is 2,090, 10 cells fewer than our baseline sample. For that reason, we re-estimate the first-stage using those 2,090 cells and report the results, comparable to our baseline results, in column (2) of Table 2. Columns (3) and (4) report the second-stage results. The estimated effect of the long commute on the number of workers who visited a hospital at least once per year (i.e., the number of patients, hereafter) is -1.171, statistically insignificant at conventional levels. Likewise, we find no statistically significant impact on number of hospital visits (column (4)). This insignificant effect is also found when we examine other outcomes such as amount of copayments and NHIS expenses. Consequently, we conclude that concern over exclusion restriction may not be relevant in our setting.

V. Results

V.1 Medical Services Usage and Medical Expenses

Estimation Results

Panel A of Table 3 presents the main results, while Panel B reports the OLS estimates for a comparison. Our baseline results show that a long commute increases medical services usage and associated costs. For example, the estimated coefficients of "LongCommute" imply that a 1% pt. increase in the share of long commuters in a cell increases the number of workers who visit a hospital at least once by 5.2 persons (a 3.5% increase), the number of hospital visits by 57 (a 4.0% increase), spending by the workers by 947,000 won (a 5.3% increase) and spending by the NHIS by 4,357,000 won (a 9.1% increase). All estimates are statistically significant at either the 1 percent or 5 percent level.

In contrast, the OLS results in Panel B report statistically insignificant impacts of long commutes on medical services usage except for the expenses paid by the NHIS. Even for these expenses, the estimated coefficient is less than one-tenth of the estimate from the 2SLS. This pattern suggests that workers who need more medical services are likely to reside near their employers and avoid a long commute, which is consistent with the findings from existing studies (e.g., Goerke and Lorenz, 2017). These differences between the 2SLS and OLS estimates highlight the importance of controlling for endogeneity to identify the true health impacts of long commute time.

Implications of the Estimates: LATE and ATE

Assuming that the health effect of a long commute is heterogeneous, our 2SLS estimate measures the average of the treatment effect across compliers, employees who worked for the 170 public employers targeted in the 2012 relocation policy and who decided to stay in the SMA, spending a substantial amount of time commuting. This local average treatment effect (LATE) can be larger or smaller than the average treatment effect (ATE), and we have no additional information to assess the comparison between the two. However, we suspect that LATE may be smaller than ATE in our setting for the following two reasons. The OLS estimate is downward biased compared to the 2SLS, suggesting that employees in good health tend to have a long commute time. Then, by the same logic, we expect that employees in good health may be more likely to stay in the SMA, bearing the burden of long commute time, while their peers in less good health may be more likely to move out of the SMA. If this is the case, our LATE will be an underestimate of the average treatment effect among the public employees, including those who moved out of the SMA since 2012. Second, the public employers targeted in the 2012 relocation policy are prestigious employers in terms of socioeconomic status, job security, and compensation in South Korea. As individuals with high socioeconomic status are usually in better health than their peers with low status, the public employees may be able to address the negative shock due to having a long commute time better than others, thus reducing its negative health impacts.

Health Costs of the 2012 Relocation Policy

Using our baseline results, we conduct a back-of-the envelope calculation of the health costs associated with the 2012 relocation policy. The 2012 policy targeted approximately 70,000 workers located in the SMA for relocation (0.53% of the SMA workers in 2010). Regardless of whether or not they relocated, they are subject to frequent work-related travel and/or long commute time (see details in Section II). Suppose that those who remained in the SMA and those who moved to the newly built cities are subject to the same amount of travel time. In addition, we assume that long commute time negatively affects their health on average to the same extent as those who remained in the SMA, which is likely to be an underestimate (see the discussion above). Then, the 2012 relocation policy may increase the share of long commuters by 0.53% pts. This increased share of long commuters implies a 1.9% increase in the number of patients, a 2.1% increase in the number of hospital visits, a 2.8% increase in the medical expenses paid by workers, and a 4.8% increase in the expenses paid by the NHIS.

V.2 Medical Checkup Outcomes and Health-related Activities

In Panel A of Table 4, we examine the outcomes from medical checkups. Columns (1) and (2) reports the OLS estimates and the 2SLS estimates, respectively, whereas column (3) reports the average of the corresponding dependent variables. The effects of long commute time on health are mixed. The increase in commute time leads to a reduction in BMI, waist circumference, total cholesterol level, LDL and ALT, which are considered to reduce health risks. In contrast, the increase in commute time reduces good cholesterol (HDL), increasing fasting blood sugar levels

(FBS) as well as creatinine, which suggest increasing health risks.

In addition, we find that people in our sample respond to a long commute by reducing both bad and good health-related activities. See Panel B of Table 4. The increase in commute time leads to a reduction in the share of smokers (1.231% pts.) and the average number of cigarette consumed (0.308 cigarettes), which is health-improving. However, the long commute time also reduces the share of employees who exercise at least twice a week by 1.218% pts. (3.6%), which is health-damaging. Lastly, we find no statistically significant impact on drinking.

These mixed effects – both health-improving and health-harming – suggest the possibility that workers may try to diminish the negative health shock, namely, having a long commute time, through a better diet and reducing smoking. Their efforts appear to be partially successful in some outcomes (e.g., BMI, waist circumference, total cholesterol, LDL, ALT) but not in other outcomes (e.g., HDL, FBS, creatinine, ALT).

VI. Discussions and Robustness Check

VI.1 Types of Diseases

We investigate heterogeneous effects of long commute time, depending on types of diseases. The NHIS data provides diagnoses of patients based on the standard classification of diseases (Korean Standard Classification of Diseases, KSCD). We use a 1-digit classification and focus on the following six types of diseases: respiratory, digestive, musculoskeletal, circulatory, endocrine & metabolic, and pregnancy & childbirth-related diseases.[13] Note that the sequence of the diseases is based on their prevalence in terms of number of patients, in descending order. We examine respiratory diseases because air pollution is severe in South Korea, and thus a long commute can

[13] The diseases correspond to the KSCD categories of J, K, M, I, E, and O, respectively.

increase the risk of respiratory diseases due to greater exposure to pollutants on the road. The second through fourth diseases on the list are considered directly related to long commute time in the medical literature. For example, existing studies show a correlation between commute time and physiological illnesses such as lower-back pain, cardiovascular diseases, and gastric disorders (see Koslowsky et al., 2013). We additionally examine the last two types of diseases because of our finding that long commute leads to less exercise (see Section V) and because pregnant women may be more vulnerable to a long commute.

Long commute time significantly increases medical services usage for all diseases except for digestive and musculoskeletal diseases. See Table 5 for estimation results. Column (1) shows that a 1% pt. increase in the share of long commuters increases the number of patients who visit hospitals at least once to treat respiratory diseases by 3.1 persons (3.3%) and the number of hospital visits by 8.8 times (2.7%). For circulatory diseases, a 1% pt. increase in the share of long commuters increases the number of patients by 2.9 persons (17.1%) and the number of hospital visits by 15.5 times (17.2%). In terms of endocrine & metabolic diseases, a 1% pt. increase in the share of long commuters increases the number of patients by 1.1 persons (9.1%) and the number of hospital visits by 8.3 times (16.7%), whereas it leads to a 30.6% (1.0 persons) increase in the number of patients and a 32.3% (2.8 times) increase in the number of hospital visits.

These findings have important implications in the South Korean setting. Specifically, of the top 10 causes of death, the second, third, and ninth most frequent causes are circulatory diseases, and these deaths accounted for 21% of all deaths in 2017.[14] In 2017, respiratory diseases account for the fourth and eighth most frequent causes of deaths, comprising 9% of all deaths, while the

---

[14] In 2017, the top 10 most prevalent causes of deaths in South Korea were cancers (27.6%), heart diseases (10.8%), cerebrovascular diseases (8.0%), pneumonia (6.8%), suicide (4.4%), diabetes (3.2%), liver diseases (2.4%), Chronic Obstructive Pulmonary Diseases (2.4%), high blood pressure (2.0%), and traffic accidents (1.8%). See Statistics Korea (2018).

sixth most frequent cause of death is endocrine & metabolic disease, accounting for 3.2% of all deaths.

Regarding musculoskeletal diseases, we initially expect that having a long commute time would lead to more treatments for these ailments, based on the complaints of public employees reported in local media reports (see Section II.2). Our results are consistent with those complaints in that all point estimates reported in column (3) are positive, although they are not statistically significant at conventional levels.

VI.2 Alternative Specifications

For a robustness check, we examine two alternative specifications – one additionally controlling for cell fixed effects and the other changing the definition of long commute status from 2 hours or more to 1 hour or more. Columns (2) and (4) of Table 6 report the corresponding results, whereas column (1) reports the results from our baseline specifications for comparison. As for medical services usage, we find that almost all estimated impacts of long commute time are mostly larger in magnitude under the alternative specifications and statistically significant at the 1 or 5% level. For example, a 1% pt. increase in the share of long commuters leads to 8.9 more patients when we include the cell fixed effects and 10.1 more patients when we use the alternative definition of long commuters.

In contrast, the estimated effects on outcomes of medical checkups and health-related activities vary depending on specifications. Under the first alternative including cell fixed effects, we find that for only 3 out of 13 outcomes are the effects of long commute time statistically significant. For the second alternative specification, the estimates are comparable to the baseline results in terms of statistical power and larger in terms of magnitude. Despite these differences, the effect of long commute time is robust for two outcomes: creatinine level and the share of workers who regularly exercise. A 1% pts increase in the share of long commuters increases the average creatinine level by 0.031mg/dL under our baseline specification, 0.053mg/dL under the first alterative, and 0.059mg/dL under the second alterative, and all estimates are statistically significant at the 5% level. As for exercise, a 1% pts increase in the share of long commuters decreases the share of employees who exercise 30 minutes or longer at least two times per week by 1.218% pts. under our baseline specification, 1.461 % pts. under the first alterative, and 2.364% pts. under the second alterative, and the latter two estimates are statistically significant at the 10% level. Both the increase in creatinine level and the reduction in exercise may lead to poor health.

VI.3 Heterogeneous Effects

This section examines the possible heterogeneous effects of long commute time on health. Our baseline results reported in Table 3 show that health outcomes systematically vary by gender (Female), homeownership (Homeowner(%)), and employment status (Employee(%)), which suggests the possible heterogeneous effects of long commute time. Thus, we allow for the heterogeneous effects by each of these variables and report the results in Table 7. We additionally include the interaction term of "LongCommute" and the variable of interest (e.g., female) in the second stage equation. If the effect of long commute differs by the variable of interest, then the coefficient of the interaction term will be statistically different from zero. As we have two endogenous variables (e.g., LongCommute, and LongCommute x Female), we use the triple interaction term among the initial share of employees who are subject to the 2012 relocation policy, post period dummy, and the variable of interest (e.g., HR x Post x Female) as the additional instrument. Finally, in both the first and second stages, we include the interaction term between

HR and the variable of interest (HR x Female) as we include the triple interaction term (e.g., HR x Post x Female) as an instrument.

Panel A shows that men and women respond differently to long commute time in terms of the number of people who visited the hospital once or more per year (column (1)) and the number of hospital visits (column (2)). Specifically, a 1 percentage point increase in long commuters increases the number of male patients by 5.6 persons and the number of hospital visits among men by 61.5 times, while the same increase for women has no statistically significant impact in both outcomes (0.238=5.755-5.993, p-val(0.923), and -3.317=61.471-64.788, p-val(0.906)). There is no statistically differential effect on female outcomes in terms of copayments and NHIS expenses.

Our estimation results imply that although women on average use hospitals more often than men, they do not visit the hospital more in response to a long commute time. However, they do use more expensive medical services due to a long commute time, so ultimately, the effect of a long commute time on copayments and NHIS expenses is the same between men and women. This implication can be accounted for by the differential selection into employment by gender. In fact, the employment rate and the average commute time greatly differ between men and women in South Korea. For example, OECD reports that in 2009/2010, the average commute time is 101 minutes for men and 84 minutes for women while the share of people who report their commute time, indicating being employed, is 74% for men and only 50% for women (OECD Family Database, LMF2.6). This pattern suggests the possibility that the gap between working women with and without a long commute is wider than the gap between working men with and without a long commute. For example, working women subject to long commutes are on average much healthier than their female counterparts without long commutes, while working men subject to long commutes are not that much healthier than their male counterparts without long commutes.

If so, the working women with long commutes may not visit the hospital more often than their counterparts, but their health conditions may become still worse and thus require more expensive medical services when they do visit a hospital.

We further examine the possible heterogeneous effects with respect to homeownership status (the share of homeowners in a cell, Panel B) and employment status (the share of employees in a cell, Panel C). However, we do not find any statistically different effects of long commute time.

VII. Conclusion

We examine the effect of long commutes on workers' medical services usage, health outcomes, and health-related activities, by exploiting a large-scale policy change in South Korea. The policy aimed to disperse the concentrated population in the capital area to the rest of the country by relocating 170 public sector employers. Despite the policy's intention, a large share of their workers kept their residences in the capital area, and spend long hours on a daily commute. Using this shock to commuting time, we estimate the health impacts of a long commute time. Our estimation results show that having a long commute time leads to a substantial increase in medical services usage. However, it has a mixed impact on health measures and health-related activities.

Our findings suggest some important policy implications applicable to South Korea as well as other developing countries, where checks and balances against the public sector and politicians are limited relative to developed countries. First, rigorous scientific assessment of individuals' decision-making processes is crucial for a policy to achieve its goals. When the South Korean president proposed the relocation plan in 2005, many opposed the plan based on the expectation that its impact on population relocation would be limited. This expectation was based on the fact that those who were married and had children would likely choose to keep their residence in Seoul and its surroundings for the sake of their children's education and the couple's dual careers. Despite this concern, the Cabinet conducted no scientific investigations to assess people's residential location decisions and forcefully relocated national government agencies and for-profit employers to less developed areas.[15] Not surprisingly, the population concentration in Seoul and its surroundings has persisted. Furthermore, those less developed areas have not been able to attract enough residents, amounting only to 60% of the policy target. Making matters worse, a large share of their incoming residents were from neighboring areas that had already been suffering from a population drain, not from the capital area (Sedaily, 2018). For these reasons, we regard the relocation policy as having failed to achieve its stated goal.

Second, our finding that unintended consequences of a policy can gravely harm the welfare of individuals and households highlights the importance of establishing protocols for evidence-based policymaking (EBPM). Several developed countries, such as the U.S and Japan, have introduced legislations for their governments to establish plans to collect data and evaluate a policy's impacts based on ex-post outcomes. The EBPM protocols have been applied to international aid projects and social programs. In contrast, other countries, including South Korea, have not adopted the EBPM protocols, and policy adoptions are heavily subject to political decision making. Such a practice may exacerbate the agency problem, risking the welfare of the general public and wasting national resources. For example, despite failing to meet its goal, the relocation policy maintains its legacy, producing subsequent and related development projects in South Korea. Establishing quantitative evaluations based on the EBPM can enhance a

[15] Since 1999, Korean law has required the Minister of Finance to assess the needs of any large-scale public project based on a cost-benefit analysis (Article 38, the National Finance Act, *GukgaJaejungBub* in Korean). However, the relocation policy was not thus assessed.

government's accountability for policy decisions and reduce a country's inefficiency, which many developing countries suffer from.

References

Burd, C., Burrows, M., and McKenzie, B. (2021). 'Travel time to work in the United States: 2019', *American Community Survey Reports*.

CDC(Center for Disease Control and Prevention). (2021, August). Diabetes Tests, retrieved from https://www.cdc.gov/diabetes/basics/getting-tested.html.

Gangwisch, J. E., Malaspina, D., Boden-Albala, B., and Heymsfield, S. B. (2005). 'Inadequate sleep as a risk factor for obesity: analyses of the NHANES I', *Sleep*, vol. 28(10), pp. 1289-96.

Goerke, L., and Lorenz, O. (2017). 'Commuting and sickness absence', SOEPpaper No.946.

Hankook. (2019, May). 'Nobody knows exactly how much government officials got paid', Retrieved from https://www.hankookilbo.com/News/Read/201903101471797630.

Henderson, V. (2002). 'Urbanization in developing countries', *The World Bank Research Observer*, vol. 17(1), pp. 89-112.

Innovation City. (2021, July). Retrieved from http://innocity.molit.go.kr/v2/.

Jacob, N., Munford, L., Rice, N., and Roberts, J. (2021). 'Does commuting mode choice impact health?', *Health Economics*, Vol. 30(2), pp. 207-30.

Kageyama, T., Nishikido, N., Kobayashi, T., Kurokawa, Y., Kaneko, T., and Kabuto, M. (1998). 'Long Commuting Time, Extensive Overtime, and Sympathodominant State Assessed in Terms of Short-term Heart Rate Variability Among Male White-collar Workers in the Tokyo Megalopolis', *Industrial Health,* vol. 36(3), pp. 209-17.

KERI(Korea Economic Research Institute). (2006). 'Key Issues of Regulations on the Seoul Metropolitan Areas', Issue Paper 15.

KIPA(Korea Institute of Public Administration). (2013). 'Analysis on Current Policy Communication Practices by the Decentralization of Central Administrative Organ into Sejong City and Improvement Measures'.

KIPA(Korea Institute of Public Administration). (2014). 'Survey of Government Officials on Relocation to Sejong'.

KIPA(Korea Institute of Public Administration). (2017). 'Is it Worth for the National Parliament to Establish its Branch in Sejong?'.

KIPF(Korea Institute of Public Finance). (2013). 'Status and Challenges for Public Sector Hiring'.

Koslowsky, M., Kluger, A. N., and Reich, M. (2013). 'Commuting stress: Causes, Effects, and Methods of Coping', *Springer Science and Business Media.*

Künn-Nelen, A. (2016). 'Does Commuting Affect Health?', *Health Economics*, vol. 25(8), pp. 984-1004.

Lee H. (2014). 'Legal Improvements for the Effective Introduction of Wage Peak System According to Mandatory Retirement age of 60', *Dong-A Law Review*, vol. 63, pp. 1-37.

Martin, A., Goryakin, Y., and Suhrcke, M. (2014). 'Does Active Commuting Improve Psychological Wellbeing? Longitudinal Evidence from Eighteen Waves of the British Household Panel Survey', *Preventive Medicine*, vol. 69, pp. 296-303.

Ministry of the Interior and Safety. (2019, June). Retrieved from https://www.korea.kr/archive /expDocView.do?docId=38557.

Ministry of the Interior and Safety. (2005, Oct). Retrieved from https://url.kr/harzc1.

Ministry of the Interior and Safety. (2010, Aug). Retrieved from https://url.kr/ni7eps.

Ministry of the Interior and Safety. (2015, Oct). Retrieved from https://url.kr/mvs3t6.

MT(Money Today). (2015, May). Retrieved from https://news.mt.co.kr/mtview.php?no=2015 022614511769080&outlink=1&ref=https%3A%2F%2Fsearch.naver.com.

Moovitapp. (2021, July). Retrieved from https://moovitapp.com/insights/en/Moovit_Insights_ Public_Transit_Index-commute-time.

NABO(National Assembly Budget Office). (2016). 'Assessment of the Relocation Policy'.

National Finance Act, Article 38. (2016, Dec). https://url.kr/yjo528.

Neumark, D., and Simpson, H. (2014). 'Place-Based Policies (No. w20049)', *National Bureau of Economic Research.*

NICE biz info. (2020, June). Retrieved from https://www.nicebizinfo.com/cm/CM0100M001GE .nice.

NHLBI(National Heart, Lung, and Blood Institute). (2005). 'High Blood Cholesterol What you need to Know', *NIH Publication No.05-3290*.

NHLBI(National Heart, Lung, and Blood Institute). (2020, July). Retrieved from https://www. nhlbi.nih.gov/health/educational/lose_wt/risk.htm.

Novaco, R. W., Stokols, D., and Milanesi, L. (1990). 'Objective and Subjective Dimensions of Travel Impedance as Determinants of Commuting Stress', *American Journal of Community Psychology*, vol. 18(2), pp. 231-57.

OECD Family Database. (2020, Oct). Retrieved from https://www.oecd.org/els/family/database.htm.

Pfeifer, C. (2018). 'An Empirical Note on Commuting Distance and Sleep during Workweek and Weekend', *Bulletin of Economic Research*, vol. 70(1), pp. 97-102.

Razvdauskas, F. V. (2018). ‘Megacities: Developing Country Domination’, *Euromonitor International.*

Schaefer, M. H., Street, S. W., Singer, J. E., and Baum, A. (1988). ‘Effects of Control on the Stress Reactions of Commuters’, *Journal of Applied Social Psychology*, Vol. 18(11), pp. 944-57.

Sedaily. (2018, Sept). Retrieved from https://m.sedaily.com/NewsView/1S4STSGDL3#cb.

Seoul Metropolitan City. (2020). Retrieved from https://data.seoul.go.kr/dataList/229/S/2/datasetView.do#.

Special Act on the Construction of Administrative City. [Act No.7391, Mar 18, 2005].

Special Act on the Construction and Support of Innovation Cities Following Relocation of Public Agencies. [Act No.8238, Jan.11, 2007].

Statistics Korea. (2009, Feb). Policy Briefing, Retrieved from https://www.korea.kr/archive/expDocView.do?docId=10770.

Statistics Korea. (2020, Dec). Policy Briefing, Retrieved from https://www.korea.kr/news/policyBreifingView.do?newsId=156428803.

UCSF Health. (2022). ‘Medical Tests’, Retrieved from https://www.ucsfhealth.org/medical-tests.

URMC(University of Rochester Medical Center Rochester, NY). (2022). ‘Health Encyclopedia’, Retrieved from https://www.urmc.rochester.edu/encyclopedia/content.aspx?contenttypeid=167&contentid=alt_sgpt.

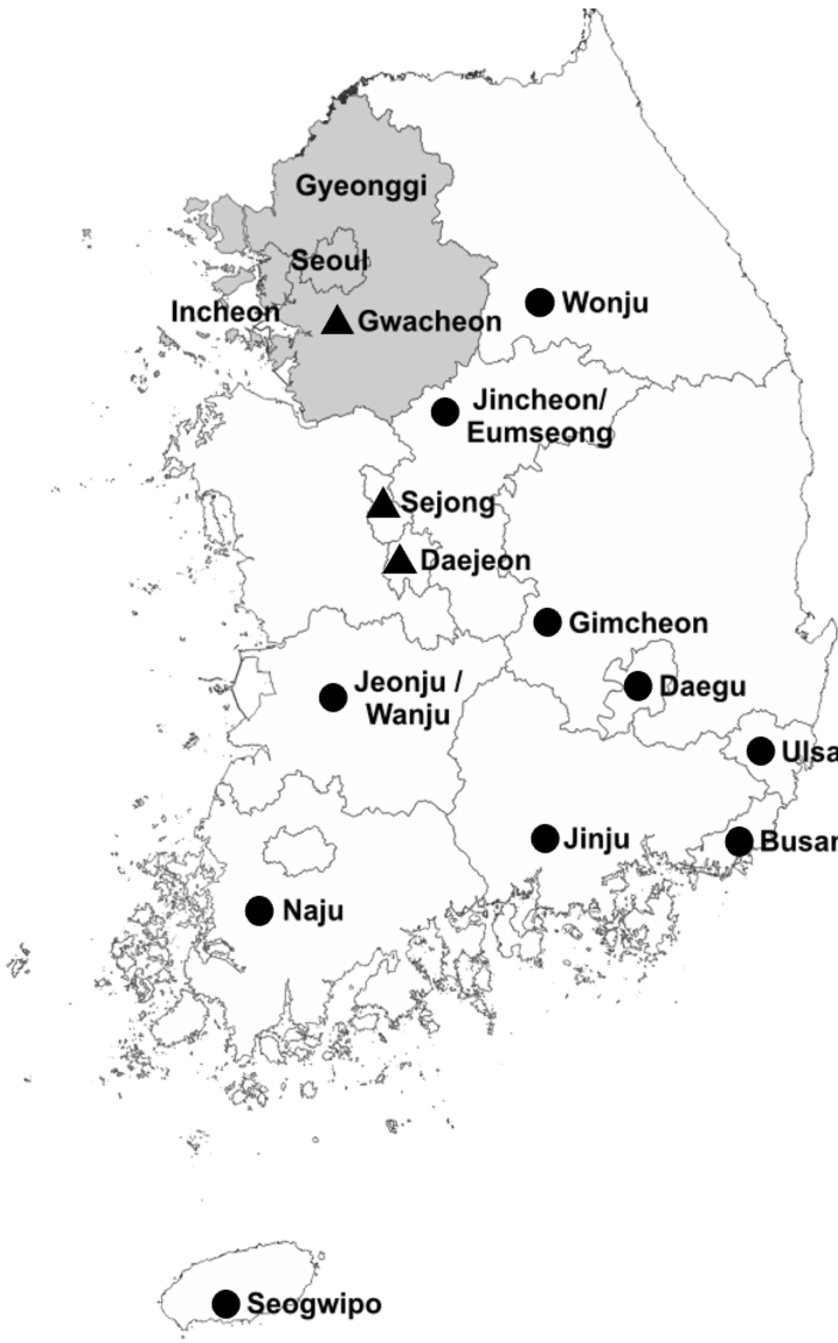
Statistics Korea. (2018, Sep). ′Causes of Deaths, 2017′.

Stutzer, A. and Frey, B. (2008). ‘Stress that Doesn't pay: The Commuting Paradox’, *The Scandinavian Journal of Economics*, vol. 110(2), pp. 339-66.

Walsleben, J. A., Norman, R. G., Novak, R. D., O'Malley, E. B., Rapoport, D. M., and Strohl,

K. P. (1999). ‘Sleep Habits of Long Island Rail-road Commuters’, *Sleep*, vol. 22(6), pp. 728-34.

Figure 1. Map of South Korea and Relocation Destinations



Note: This map illustrates the territories of South Korea and the boundaries of major administrative units (total of 17). Seoul, the capital city, is located in the northwest corner of South Korea, and Incheon, a key port city, is located west of Seoul. Gyeonggi Province surrounds these two key cities. These three areas are referred to as the Seoul Metropolitan Area (SMA, highlighted in grey) in this manuscript. The national government agencies are located in Seoul, Gwacheon, Sejong, and Daejeon; the latter three are marked with triangles. Locations highlighted with circles indicate the 11 cities that have been chosen to host public sector employers since 2012.

Table 1. Summary Statistics

	Average	S.D.	Min	Max
	(1)	(2)	(3)	(4)
A. Relocation Policy and Commute				
Year 2010 (Pre relocation)				
- % long commuters	29.394	(11.188)	0	57.607
- % high-risk workers	1.106	(0.507)	0.852	6.866
Year 2015 (Post relocation)				
- % long commuters	31.734	(11.549)	1.687	61.162
- % high-risk workers	1.022	(0.407)	0.233	5.508
B. Demographics				
Age	39.873	(9.225)	25	59
Female	0.422	(0.494)	0	1
4-year college graduates (%)	47.256	(19.443)	0	89.390
Owning a house (%)	51.959	(12.020)	20.873	90.323
Being married (%)	62.374	(25.202)	5.705	97.638
Having a white-collar job (%)	59.878	(17.117)	7.653	90.751
Employee (%)	78.355	(11.462)	23.252	97.347
C. Health Outcomes				
No. of workers who visited a hospital	148.850	(86.384)	1	595
No. of workers never visiting a hospital	21.765	(18.608)	0	103
No. of visits	1,422.583	(786.950)	6	5,048
Copayments*	17.908	(10.955)	0.060	74.107
NHIS expenses*	48.064	(35.981)	0.143	259.861
Body Mass Index (BMI)	23.640	(1.445)	18.550	27.917
Waist circumference: Men	84.224	(1.424)	77.400	92.417
Waist circumference: Women	73.201	(2.927)	62.334	86.667
Total cholesterol (mg/dL)	193.586	(9.965)	159.000	243.167
HDL (“good”: mg/dL)	56.870	(5.773)	42.400	94.533
LDL (“bad”: mg/dL)	112.424	(10.808)	71.333	282.220
Fasting Blood Sugar (FBS, mg/dL)	95.105	(6.042)	78	131.400
Creatinine (mg/dL)	0.964	(0.225)	0.600	3.350
ALT (mg/dL)	25.640	(8.135)	9	90.191
% Drink twice + /week	30.920	(15.531)	0	73.913
% Smokers	27.299	(21.727)	0	84.615
Cigarette consumption per day (no.)	3.876	(3.316)	0	13.875
% 30 min+ workout/week twice +	33.983	(8.561)	0	100

Note: The unit of observation is cell by year (2010, 2015). Our sample includes 1,050 cells defined by sex, age group, location. Observations are weighted by their number of employees. “% long commuters” refers to the share of workers who spend at least two hours commuting for work per day, and “% high-risk workers” refers to the share of workers subject to the 2012 relocation policy * Unit: 1 million won (approximately US$827).

Table 2. First-Stage and Validity of Identification Strategy

Types	**First-Stage**	**First-Stage**	**Falsification**	**Falsification**
Dep.V.	LCommute	LCommute	Patients	Hospital Visits
Sample	Baseline	(1) – 10 cells	(1) – 10 cells	(1) – 10 cells
	(1)	(2)	(3)	(4)
LongCommute(%)	-	-	-1.171	4.357
			(0.813)	(9.828)
HR x Post	1.877***	1.911***	-	-
	(0.347)	(0.346)		
HR	-1.448***	-1.494***	-4.040**	-2.635
	(0.413)	(0.427)	(1.774)	(24.376)
Female	-5.881***	-5.906***	-28.053***	-68.308
	(0.638)	(0.638)	(4.782)	(53.763)
Married (%)	-0.021	-0.020	-0.018	1.864
	(0.039)	(0.039)	(0.178)	(1.956)
College (%)	0.398***	0.397***	1.039***	-0.206
	(0.047)	(0.047)	(0.329)	(4.063)
Homeowner (%)	0.310***	0.310***	0.083	-4.091
	(0.051)	(0.051)	(0.249)	(3.190)
White collar (%)	0.043	0.044	0.592***	5.681***
	(0.031)	(0.031)	(0.098)	(1.072)
Employee (%)	-0.085*	-0.083*	-0.552*	-4.951
	(0.047)	(0.047)	(0.319)	(3.143)
Fixed Effects	Loca.x Year	Loca.x Year	Loca.x Year	Loca.x Year
F-stat	30.514	31.873	-	-
Mean Dep.	30.570	30.576	49.239	516.712
No Obs.	2,100	2,090	2,090	2,090
R-squared	0.895	0.895	0.818	0.790

Note: The unit of observation is cell by year. Observations are weighted by their number of employees. We additionally include dummies for 7 age groups and location by year fixed effects. Robust standard errors, clustered at location levels (total of 75), are reported in parentheses. Symbols ***, **, * indicate statistical significance at the 1%, 5% and 10% level, respectively.

Table 3. Effects of Long Commute on Medical Services Usage and Medical Expenses

Dep.Var.	Patients (1)	Hospital Visits (2)	Copayments (3)	NHIS Expenses (4)
A. 2SLS				
LongCommute (%)	5.201**	57.260***	0.947**	4.357***
	(2.031)	(21.059)	(0.430)	(1.407)
Female	28.350*	668.875***	6.873***	28.144***
	(14.896)	(129.467)	(2.443)	(7.456)
Married (%)	0.500	4.533	0.067	0.223
	(0.598)	(4.408)	(0.066)	(0.224)
College (%)	2.081**	11.116	0.060	-0.607
	(0.852)	(10.213)	(0.185)	(0.646)
Homeowner (%)	-2.603***	-22.692***	-0.347**	-1.395***
	(0.919)	(8.680)	(0.163)	(0.527)
White collar (%)	-0.598*	2.347	0.008	0.078
	(0.362)	(2.703)	(0.042)	(0.154)
Employee (%)	4.200***	38.084***	0.450***	1.293***
	(1.062)	(9.033)	(0.128)	(0.367)
Mean Dep.	148.850	1,422.583	17.908	48.064
No Obs.	2,100	2,100	2,100	2,100
R-squared	0.793	0.743	0.693	0.427
B. OLS				
LongCommute (%)	0.366	1.186	0.028	0.341**
	(0.485)	(3.747)	(0.052)	(0.157)

Note: The unit of observation is cell by year. Observations are weighted by their number of employees. We additionally include dummies for 7 age groups and location by year fixed effects in Panels A and B. Robust standard errors, clustered at location levels (total of 75), are reported in parentheses. Symbols ***, **, * indicate the corresponding coefficient is statistically different than zero at the 1%, 5% and 10% level, respectively. The unit for copayments and NHIS expenses is 1 million won (approximately US$827).

Table 4. Medical Checkup Outcomes and Health-Related Activities

Dep. Var.	OLS (1)	IV (2)	Mean Dep. (3)
A. Medical Checkup Outcomes			
Body Mass Index (BMI)	-0.004	-0.104***	23.640
	(0.005)	(0.036)	
Waist circumference	0.004	-0.307***	79.536
	(0.012)	(0.086)	
Total cholesterol	0.218***	-1.449***	193.586
	(0.041)	(0.376)	
HDL (good)	-0.004	-0.250**	56.870
	(0.014)	(0.127)	
LDL (bad)	0.168***	-1.661***	112.424
	(0.060)	(0.462)	
FBS	0.013	0.429***	95.105
	(0.016)	(0.163)	
CRTN	0.000	0.031**	0.964
	(0.001)	(0.014)	
ALT	0.110***	-0.511***	25.640
	(0.023)	(0.198)	
B. Health-related Activities			
% Drink twice + /week	0.050	-0.265	30.920
	(0.043)	(0.303)	
Alcohol consumption per week	0.011***	-0.022	3.857
	(0.003)	(0.030)	
% Smokers	0.143***	-1.231***	27.230
	(0.042)	(0.435)	
Cigarette consumption per day	0.024***	-0.308***	3.876
	(0.007)	(0.087)	
% 30 min+ workout/week: twice +	-0.185***	-1.218***	33.983
	(0.037)	(0.446)	
Fixed Effects	Loca.xYear	Loca.xYear	

Note: Each entry in columns (1) and (2) is based on a separate regression analysis and reports the estimated coefficient of “LongCommute(%).” The unit of observation is cell by year. Observations are weighted by their number of employees. We additionally include dummies for 7 age groups and location by year fixed effects in Panels A and B. Robust standard errors, clustered at location levels (total of 75), are reported in parentheses. Symbols ***, **, * indicate the corresponding coefficient is statistically different than zero at the 1%, 5% and 10% level, respectively.

Table 5. Impacts of Long Commute Depending on Disease Types

Types of Diseases	Respiratory	Digestive	Musculoskeletal	Circulatory	Endocrine & Metabolic	Pregnancy & Childbirth
	(1)	(2)	(3)	(4)	(5)	(6)
A. No. of patients						
- Mean	94.365	45.913	43.480	17.038	12.346	3.394
- Estimated Effect	3.107**	0.879	1.082	2.922***	1.126**	1.038***
(s.e.)	(1.264)	(0.705)	(0.831)	(0.815)	(0.475)	(0.247)
B. No. of visits						
- Mean	330.670	103.541	194.514	90.366	49.719	8.554
- Estimated Effect	8.832**	-0.117	1.498	15.527***	8.316***	2.762***
(s.e.)	(4.102)	(1.729)	(3.956)	(4.906)	(2.595)	(0.761)
C. Copayments						
- Mean	1.943	1.796	2.310	1.166	0.699	0.397
- Estimated Effect	0.058	0.070	0.015	0.186*	0.004	0.109***
(s.e.)	(0.044)	(0.087)	(0.076)	(0.112)	(0.053)	(0.035)
D. NHIS Expenses						
- Mean	4.178	4.098	5.527	4.011	1.016	2.755
- Estimated Effect	0.169*	0.072	0.086	1.342*	0.142**	0.698***
(s.e.)	(0.089)	(0.374)	(0.216)	(0.692)	(0.059)	(0.225)

Note: The unit of observation is cell by year. Observations are weighted by their number of employees. We additionally include dummies for 7 age groups and location by year fixed effects in Panels A and B. Robust standard errors, clustered at location levels (total of 75), are reported in parentheses. Symbols ***, **, * indicate the corresponding coefficient is statistically different than zero at the 1%, 5% and 10% level, respectively. The unit for copayments and NHIS expenses is 1 million won (approximately US$827).

Table 6. Robustness Check

Dep. Var.	Baseline (1)	Adding Cell FE estimate (2)	Adding Cell FE (s.e.) (3)	LongCommute (1hour+) estimate (4)	LongCommute (1hour+) (s.e.) (5)
A. Medical Services Usage & Expenses					
No. patients	5.201**	8.900***	(3.454)	10.090**	(4.222)
No. hospital visits	57.260***	65.434**	(31.781)	111.075**	(47.962)
Copayments	0.947**	0.799	(0.601)	1.838**	(0.846)
NHIS expenses	4.357***	4.728**	(2.193)	8.452**	(3.368)
B. Medical Checkup Outcomes					
Body Mass Index (BMI)	-0.104***	-0.011	(0.047)	-0.202*	(0.105)
Waist circumference	-0.307***	-0.148	(0.116)	-0.596**	(0.276)
Total cholesterol	-1.449***	-0.257	(0.441)	-2.810**	(1.221)
HDL (good)	-0.250**	-0.277	(0.227)	-0.484	(0.304)
LDL (bad)	-1.661***	-0.654	(0.695)	-3.222**	(1.302)
FBS	0.429***	0.070	(0.260)	0.832**	(0.423)
Creatinine	0.031**	0.053**	(0.022)	0.059**	(0.025)
ALT	-0.511***	0.068	(0.361)	-0.992**	(0.481)
C. Health-related Activities					
% Drink twice + /week	-0.265	-1.329**	(0.620)	-0.514	(0.614)
Alcohol consumption	-0.022	0.047	(0.056)	-0.043	(0.064)
% Smokers	-1.231***	0.474	(0.550)	-2.388**	(1.061)
Cigarette consumption	-0.308***	-0.054	(0.100)	-0.598**	(0.236)
% 30 min+	-1.218***	-1.461*	(0.844)	-2.364*	(1.276)
Fixed Effects	Loca.xYear	Cell and Loca.xYear		Loca.xYear	

Note: Each entry in columns (1), (2), and (4) is based on a separate regression analysis and reports the estimated coefficient of "LongCommute(%)." The unit of observations is cell by year. Observations are weighted by their number of employees Robust standard errors, clustered at location levels (total of 75), are reported in parentheses. Symbols ***, **, * indicate the corresponding coefficient is statistically different than zero at the 1%, 5% and 10% level, respectively.

Table 7. Heterogeneous Effects

Dep.Var.	Patients (1)	Hospital Visits (2)	Copayments (3)	NHIS Expenses (4)
A. Female				
LongCommute (%)	5.755**	61.472**	0.956**	4.173***
	(2.259)	(24.471)	(0.453)	(1.399)
LongCommute x Female	-5.993***	-64.788***	-0.488	-0.319
	(1.974)	(23.749)	(0.407)	(1.790)
Female	138.558***	1,792.996***	14.457*	25.926
	(41.769)	(481.243)	(8.653)	(36.044)
B. Homeownership				
LongCommute (%)	3.851**	50.700***	0.728**	3.150**
	(1.624)	(17.430)	(0.319)	(1.278)
LongCommute x Homeown	-0.350	8.181	-0.084	-1.048
	(0.888)	(10.723)	(0.181)	(0.773)
Homeowner(%)	-2.563***	-28.054***	-0.325***	-0.946**
	(0.834)	(7.866)	(0.120)	(0.464)
C. Employment				
LongCommute (%)	5.332***	64.383***	0.912***	3.910***
	(1.679)	(23.100)	(0.346)	(1.269)
LongCommute x Employee	-0.655	-38.842	0.186	2.419
	(3.273)	(32.832)	(0.566)	(2.768)
Employee(%)	3.996***	33.942***	0.483**	1.597**
	(1.343)	(11.396)	(0.192)	(0.752)

Note: The unit of observation is cell by year. Observations are weighted by their number of employees. We additionally include dummies for 7 age groups and location by year fixed effects in Panels A and B. Robust standard errors, clustered at location levels (total of 75), are reported in parentheses. Symbols ***, **, * indicate the corresponding coefficient is statistically different than zero at the 1%, 5% and 10% level, respectively. The unit for copayments and NHIS expenses is 1 million won (approximately US$827).