THREE ESSAYS ON FOOD AND HEALTH POLICIES IN SOUTH KOREA

by

KWANGHUN YEON

(Under the Direction of TRAVIS A. SMITH)

ABSTRACT

This dissertation contains three chapters on food and health policies in South Korea. All three chapters estimate the effects of the South Korean government's food policies on health. Chapter 1 evaluates the effects of an early warning system for food-borne outbreaks. In June 2006, over 3,600 students in 46 schools became sick due to a food-borne illness outbreak in the Seoul Capital Area in South Korea. In February 2008, the South Korean government responded by introducing an early warning system (EWS) for food-borne outbreaks in all schools. This study evaluates the effects of this warning system using monthly administrative panel data in South Korea from January 2002 through December 2020. Using the regression discontinuity, differences-in-discontinuities, and difference-in-differences approach, we could not find evidence that the EWS reduces the number of outbreaks and cases of food-borne illness.

Chapter 2 evaluates the effects of voluntary restaurant sanitary grades. Restaurants are the leading cause of food-borne illnesses in South Korea. In December 2009, Seoul City introduced voluntary restaurant sanitary grades to reduce the incidence of food-borne illness. In May 2017, the program became permanent in all regions in South Korea. This study evaluates the effects of the Seoul City restaurant sanitary grades using monthly administrative panel data in South Korea from January 2002 through April 2013. Using a difference-in-difference-in-differences approach, we could not find evidence that the Seoul city restaurant sanitary grades significantly reduce the number of outbreaks and cases of food-borne illness.

Chapter 3 estimates the effects of NutriPlus, a special supplemental nutrition program for women, infants, and children, on birth outcomes using an instrumental variable approach. We do not find statistically significant evidence that NutriPlus spending improves the incidence of low birth weight, very low birth weight, normal birth weight, and premature birth. However, we find that increasing NutriPlus spending by 100% reduces the average birth weight by 4.369 grams and the likelihood of high birth weight by approximately 0.021%. The results imply that the effect on the average is driven by a reduction in high birth weights.

.

INDEX WORDS: Food-borne Illness, Warning System, Restaurant, NutriPlus, Birth Weight

THREE ESSAYS ON FOOD AND HEALTH POLICIES IN SOUTH KOREA

by

KWANGHUN YEON

BE, Korea University, Republic of Korea, 2009

ME, Korea University, Republic of Korea, 2019

MS, Ghent University, Belgium, 2019

A Dissertation Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA

2023

THREE ESSAYS ON FOOD AND HEALTH POLICIES IN SOUTH KOREA

by

KWANGHUN YEON

Major Professor: Travis A. Smith
Committee: Chen Zhen
William Secor

Electronic Version Approved:

Ron Walcott
Vice Provost for Graduate Education and Dean of the Graduate School
The University of Georgia
May 2023

DEDICATION

I dedicate this dissertation to two people. Firstly, I would like to thank Professor Doo Bong Han, who supported me throughout my academic journey. His support and guidance during my Master's degree opened up opportunities for me to study in Europe and the U.S. I am forever grateful for his mentorship. Secondly, I dedicate this dissertation to my wife, Dalee, who has inspired and motivated me throughout this journey. I would not have made it this far without her constant encouragement and belief in me. I am eternally grateful for that.

ACKNOWLEDGEMENTS

I would like to express my gratitude to my advisor, Professor Travis A. Smith. He always encouraged me to write a good paper and guided me to become a good researcher. When I had problems in my study, he always gave me good advice and helped me to overcome the difficulties. I am also deeply grateful to Professor Chen Zhen and Professor William Secor for their insightful comments and suggestions on my dissertation. I would like to thank them for their time and efforts in serving on my committee.

TABLE OF CONTENTS

LIST OF TABLES

Page

LIST OF FIGURES

Page

CHAPTER 1

THE EFFECTS OF AN EARLY WARNING SYSTEM FOR FOOD-BORNE OUTBREAKS IN SCHOOLS: EVIDENCE FROM SOUTH KOREA

1.1. INTRODUCTION

In June 2006, 3,613 students suffered from food-borne illness in 46 schools in the Seoul Capital Area in South Korea.[1] These food-borne illness outbreaks were mainly caused by a large food company, CJ Food System.[2] In response, the South Korean government introduced the early warning system (EWS) for food-borne outbreaks in all schools in February 2008. The policy requires school employees to manually enter the list of food items and suppliers into the EWS when they purchase. Additionally, if food-borne illness outbreaks occur in a school, the school employees must enter the information about the outbreak (e.g., symptom and potentially hazardous foods (PHF)) into the system. Then, the EWS sends a warning text message to other schools' employees who purchased foods from the same food suppliers and the relevant government authorities. Therefore, other schools can discontinue providing students with the PHF, and the relevant government authorities can rapidly investigate the food suppliers to

[1] The South Korean government provides a universal free, eco-friendly school lunch program. Therefore, schools are one of the major food sources, which provide children with meals in South Korea. The percentage of schools offering school meals has been 100% since 2012. The fraction of primary and secondary school students participating in the school meal program has been 99.9% since 2015, except for 2017. It was 99.8% in 2017. (see https://www.index.go.kr/unity/potal/main/EachDtlPageDetail.do?idx_cd=1543 for details).

[2] The company changed its name to CJ Freshway in 2008 and still provides schools with food.

identify contaminated food items and pathogens. Ultimately, the EWS aims to prevent the students from eating already-contaminated foods.

Within a few years of its launch, the South Korean government upgraded the EWS by connecting it with two public electronic procurement systems. The EWS was connected with School Food e-Procurement System (also called eaT) in November 2013—henceforth, the first upgrade.[3] Subsequently, the EWS was connected with Korea ON-Line E-Procurement System (KONEPS) in September 2014—henceforth, the second upgrade.[4] In 2021, the South Korean government announced that they would expand the EWS to other institutions, including preschools, kindergartens, hospitals, and companies.[5]

This study estimates the effects of the EWS on the number of outbreaks and cases of food-borne illness using the monthly administrative panel data in South Korea from January 2002 through December 2020. However, the estimates would be biased by the potential endogeneity due to measurement errors or omitted variables. To address this problem, we use a regression discontinuity (RD), a differences-in-discontinuities, and a difference-in-differences (DID) approach. The main results show that the EWS is likely to reduce the number of outbreaks and cases of food borne illness by at least 77% and 65%, respectively, in the short run. We also find that the EWS is likely to reduce the number of outbreaks and cases of food borne illness by 90%–91% and 81%–85%, respectively, in the long run. However, they are not robust to different

---

[3] eaT is managed by Korea Agro-fisheries & Food Corporation. A new system NeaT was released in September 2022 (see https://ns.eat.co.kr/NeaT/eats/index.html for details).

[4] KONEPS is managed by Public Procurement Service (see https://www.g2b.go.kr/index.jsp for details).

[5] See the government press release for details :

https://www.mfds.go.kr/brd/m_99/view.do?seq=45041&srchFr=&srchTo=&srchWord=&srchTp=0&itm_seq_1=0&itm_seq_2=0&multi_itm_seq=0&company_cd=&company_nm=&page=88

identification strategies, model specifications, and study windows. Therefore, we could not find strong evidence that the EWS significantly reduces the incidence of food-borne illness.

Previous studies have estimated the impacts of similar programs. Scharff et al. (2016) and Brown et al. (2021) find that network systems for DNA fingerprinting have health and economic benefits. Unlike the above studies, the empirical studies about restaurant sanitary letter grades programs yield mixed results; Jin and Leslie (2003) and Jin and Leslie (2019) find that the programs significantly reduce the incidence of food-borne illness, while Ho et al. (2019) and Krinsky et al. (2022) do not find a significant effect.

The World Health Organization (WHO) reports that approximately 600 million people get sick, and 420,000 people die each year from food-borne illnesses (WHO, 2022). Food-borne illness also leads to $110 billion in medical and productivity costs yearly in low- and middle-income countries (WHO, 2022). Additionally, food-borne illness imposes an economic burden on developed countries as well. For example, the estimated medical cost of food-borne illness in the United States is from $60.9 billion to $97.4 billion (Scharff, 2018). In addition, food-borne illness affects food supply and market price (Arnade et al., 2016), which could aggravate food insecurity through the reduced food supply and high food prices. Government interventions have been implemented and studied in the response above problems. For example, researchers have studied the disclosure of restaurant hygiene scores (Jin & Leslie, 2003; Bederson et al., 2018; Ho et al., 2019; Jin & Leslie, 2019; Dai & Luca, 2020), food-safety standards (Alberini et al., 2008; Minor & Parrett, 2017; Ollinger & Bovay, 2018; Ollinger & Bovay, 2020; Adalja et al., 2022), and the PulseNet surveillance system (Scharff et al., 2016; Brown et al., 2021). Rapidly sharing information on the outbreaks of food-borne illness is essential to prevent the students from eating already-contaminated foods.[6] However, to our knowledge, no study has explored the EWS for food-borne outbreaks. To fill the research gap, this study aims to evaluate the impacts of the EWS using administrative data in South Korea.

The rest of the paper is organized as follows: Section 1.2 describes the data. Section 1.3 presents the empirical model. Section 1.4 presents the main results. Lastly, Section 1.5 discusses and concludes with some policy implications.

1.2. DATA

This study uses the monthly panel data on reported outbreaks and cases of food-borne illness from January 2002 through December 2020 across 17 regions in South Korea, as reported by the South Korean Ministry of Food and Drug Safety (MFDS). In South Korea, food-borne disease outbreaks (cases) are defined as the number of reported incidents (people) in which two or more people experience a similar illness resulting from ingesting a common food.[7] The MFDS data include information by pathogen type (e.g., salmonella, E. coli, etc), food source (e.g., home versus school), and region.[8]

---

[6] PulseNet is a national laboratory network to identify potential outbreaks of food-borne illness using DNA fingerprints of bacteria (see https://www.cdc.gov/pulsenet/index.html for details). Unlike PulseNet, the early warning system (EWS) is a network to urgently share information on potentially hazardous foods (PHF), the food suppliers, and the symptoms when food-borne illness outbreaks occur.

[7] See https://www.index.go.kr/unity/potal/main/EachDtlPageDetail.do?idx_cd=2761 for details. This definition is similar to that of CDC in the U.S. see https://www.cdc.asgov/foodsafety/outbreaks/investigating-outbreaks/confirming_diagnosis.html#:~:text=A%20foodborne%20disease%20outbreak%20is,Reporting%20System%20(NORS).** for details.

[8] The raw data is not publicly available. The South Korean government website (see https://www.foodsafetykorea.go.kr/main.do) only provides a subset of the raw data. Therefore, the raw data was obtained by request through a government website to disclose the information (see https://www.open.go.kr/com/main/mainView.do).

The number of reported outbreaks and cases of food-borne illness is likely underreported.[9] There are two reasons. First, the Food Sanitation Act mandates medical doctors and food service institutions must report the occurrence of food-borne illness to the local government authorities. Patients (or their guardians) can also report voluntarily. However, if they do not report it to the local government authorities, it is not reflected in the statistics. Second, as we mentioned above, the reported outbreaks and cases are the statistics when at least two people ate the same food and became ill. Therefore, if only one person eats some foods and then becomes ill, it is not reflected in the statistics. This underreporting issue is discussed in Section 1.3.

In the MFDS data, there are 15 types of food-borne illnesses. The top 10 food-borne illnesses in order of the number of outbreaks are unknow type, norovirus, pathogenic E. coli, salmonella, vibrio parahaemolyticus, staphylococcus aureus, protozoa, C. perfringens, campylobacter jejuni, and bacillus cereus.[10] This study uses both the total number of reported outbreaks and cases of food-borne illness, regardless of type, as well as the above top 10 food-borne illnesses including the unknown type.

The six food sources in the MFDS data are schools, institutions, restaurants, homes, other food sources, and unknown sources. Institutions indicate food service institutions (e.g., nursing homes) except for schools.

South Korea currently consists of 17 regions.[11] In July 2012, a new administrative district, Sejong, was established by merging areas of Chungcheongbuk-do and

---

[9] Bellemare and Nguyen (2018) also pointed out the underreporting issue when they used U.S. administrative data.

[10] The other 5 types are: other viruses, natural toxins, other bacteria, chemical substance, and clostridium botulinum.

[11] The 17 regions are Seoul, Gyeonggi-do, Incheon, Chungcheongbuk-do, Chungcheongnam-do, Daejeon, Sejong, Jeollabuk-do, Jeollanam-do, Gwangju, Gangwon-do, Gyeongsangbuk-do, Gyeongsangnam-do, Daegu, Ulsan, Busan, and Jeju-do.

Chungcheongnam-do, creating structural changes in these three regions. Therefore, we exclude the data on Sejong, Chungcheongbuk-do, and Chungcheongnam-do and construct monthly balanced panel data from January 2002 to December 2020 at the type-source-region level.

Control variables are obtained from the Korean Statistical Information Service provided by Statistics Korea.[12] While the unemployment rate is the region-month-year level variable, the other variables are the region-year level variables because of data availability.

Table 1.1 presents descriptive statistics. In any given month, the average region reports 0.545 outbreaks and 12.655 cases of food-borne illness per one million people. The number of reported outbreaks and cases in schools is calculated per million students. Looking at Table 1.1, schools are the primary food source of food-borne illnesses. The average number of reported outbreaks (cases) of food borne illness in school is 0.554 (36.394) per million. Table 1.1 also shows that norovirus, and pathogenic E. coli are the most common food-borne illnesses except for unknown type. The average number of reported outbreaks (cases) of norovirus is 0.073 (2.406) per million, and the average number of reported outbreaks (cases) of pathogenic E. coli is 0.069 (3.280) per million. The average number of reported outbreaks (cases) of unknown type is 0.214 (2.169) per million, comprising the largest portion of outbreaks.

1.3. EMPIRICAL MODEL

We aim to estimate the impact of the early warning system (EWS) on the number of outbreaks and cases of food-borne illnesses. We first use a regression discontinuity (RD) design with time as the running variable to evaluate the short-run effects with the following equation:

$$Y_{rt} = \alpha_1 + \beta_1 Post_t + \eta_1 f(t) + \theta_1 f(t) \times Post_t + \mu_m + \varepsilon_{rt} \qquad (1.1)$$

[12] Korean Statistical Information Service, see https://kosis.kr/index/index.do.

where Y_{rt} is the number of reported food-borne outbreaks (cases) per one million students in schools in region r in month t. $Post_t$ is an indicator variable equaling one if the date is February 2008 or later. $f(t)$ is a function of the running variable. μ_m is month fixed effects (i.e., dummy variables indicating February to December). ε_{rt} is an error term. In equation (1.1), our primary interest is to estimate β_1 indicating the effects of the EWS.

In equation (1.1), Y_{rt} is likely underreported, as we mentioned in Section 1.2. While measurement error in the dependent variable leads to inflated variances for estimated coefficients, it still provides unbiased estimates of the coefficients and valid statistical inference (Gujarati, 2015; Greene, 2018; Wooldridge, 2019).[13] In equation (1.1), we only use the linear and quadratic polynomial $f(t)$ according to the recommendation of Gelman and Imbens (2019).[14]

However, if an external shock occurred around the timing of the introduction of the EWS, RD estimates would be biased. To address this problem, we also employ the differences-in-discontinuities approach, which was empirically implemented by Grembi et al. (2016) and Hansen et al. (2020). We use institutions, homes, and other food sources as a control group with the following equation:

$$Y_{srt} = \alpha_2 + \beta_2 Post_t + \gamma_2 School_s + \delta_2 Post_t \times School_s + \omega_2 f(t) \times School_s + \eta_2 f(t) + \theta_2 f(t) \times Post_t + \kappa_2 f(t) \times Post_t \times School_s + \mu_m + \varphi School_s \times \mu_m + \varepsilon_{srt} \quad (1.2)$$

where Y_{srt} is the number of reported outbreaks (cases) of food-borne illness per million in food source s in region r in month-year level time t. $School_s$ is a dummy variable indicating one if

[13] Bellemare and Nguyen (2018) note that if the dependent variable is the number of food-borne illness outbreaks and cases, it is almost surely underreported, and it results in attenuation bias (i.e., the estimator is biased toward zero). However, attenuation bias is not caused by measurement errors in a dependent variable, but an independent variable (Gujarati, 2015; Greene, 2018; Wooldridge, 2019). Therefore, our model does not raise concern about attenuation bias.

[14] Gelman and Imbens (2019) argue that third or higher-degree polynomials of the running variable lead to “noisy estimates, sensitivity to the degree of the polynomial, and poor coverage of confidence intervals.”

the food source is school. $f(t)$ is a function of the running variables. In equation (1.2), we use the linear and quadratic polynomial $f(t)$, and the parameter of interest is δ_2.

However, parametric RD is sensitive to misspecifications: if the function of a running variable cannot accurately fit the time trend, the RD estimates would be biased (Angrist and Pischke, 2009). Unlike parametric RD, non-parametric RD does not depend on the correct model specification because it only uses data in a small enough neighborhood around the discontinuity (Angrist and Pischke, 2009). Therefore, we also use the non-parametric RD and non-parametric differences-in-discontinuities approach.[15] However, we do not have much data in this case, and the sample average is biased. (Angrist and Pischke, 2009). Therefore, we use the nonparametric version of regressions called local linear regression and local polynomial regression, which give more weight to points close to the cutoff.

To evaluate the long-run effects, we use the difference-in-differences (DID) model with the following equation:

$$Y_{srt} = \alpha_3 + \beta_3 Post_t \times School_s + X_{rt} + \sigma_s + \tau_t + \rho_r + \varepsilon_{srt} \qquad (1.3)$$

where X_{rt} is a vector of covariates including 20–44 years (%), 45–64 years (%), 65 years and over (%), male (%), middle school students (%), high school students (%), unemployment rate (%). σ_s is the food source fixed effects. τ_t is the time fixed effects. ρ_r is region fixed effects. ε_{srt} is an error term. In equation (1.3), β_3 is the parameter of interest.

The DID estimation relies on parallel trends assumption that the number of outbreaks (cases) of food-borne illness per million would follow the same trends in treatment food source (i.e., school) and control food source (i.e., institutions, homes, and other food sources) if there was no treatment (i.e., the introduction of the EWS). The event study design is commonly used

[15] The non-parametric differences-in-discontinuities approach was recently used by He et al. (2020) and Giambona and Ribas (2022).

to test if parallel trends assumption is likely to hold. We use the following specification (see details Clarke & Tapia-Schythe, 2021):

$$Y_{srt} = \alpha_4 + \sum_{l=-73}^{l=-2} \beta_l EWS_{l,st} + \sum_{l=0}^{l=154} \beta_l EWS_{l,st} + X_{rt} + \sigma_s + \tau_t + \rho_r + \varepsilon_{srt} \tag{1.4}$$

where $EWS_{l,st}$ are the event study dummies indicating one if the food source is school and the relative time to the introduction of the EWS (i.e., $t -$ February 2008) is l. In equation (1.4), the reference period is set as $l = -1$. Researchers jointly test if β_l is insignificantly different from zero in the pre-treatment period. However, failing to reject the null does not guarantee that parallel trends hold in the post-treatment period in the absence of intervention (Kahn-Lang & Lang, 2020; Bilinski & Hatfield, 2020). Additionally, high statistical power may detect even the slightest violations of parallel trends, even if it is not practically significant (Bilinski & Hatfield, 2020). Therefore, we compare the estimates from a DID model with and without allowing group-specific time trends according to the recommendation of Kahn-Lang & Lang (2020) and Bilinski & Hatfield (2020). We use the following specification with a group-specific trend:

$$Y_{srt} = \alpha_5 + \beta_5 Post_t \times School_s + \gamma_5 School_s \times t + X_{rt} + \sigma_s + \tau_t + \rho_r + \varepsilon_{srt} \tag{1.5}$$

We compare β_3 in equation (1.3) with β_5 in equation (1.5). If parallel trends hold, β_3 and β_5 would be similar. Therefore, if β_3 and β_5 are not similar, we prefer the DID model with allowing group-specific time trends (i.e., allowing for some violations of parallel trends assumption) to the DID model without allowing group-specific time trend (i.e., assuming parallel trends hold).

1.4. RESULTS

Figure 1.1 (Figure 1.2) illustrates trends in the average total outbreaks (cases) of food-borne illness per million for each food source in the full data period. We add a linear (quadratic) fitted line in Figures 1.1.1 and 1.2.1 (Figures 1.1.2 and 1.2.2) to illustrate the discontinuity in February 2008. In Figures 1.1.1 and 1.1.2 (Figure 1.2.1), we found a drop in the total number of outbreaks (cases) of food-borne illness in schools around February 2008, when the early warning system (EWS) was introduced. However, we also find that the total number of outbreaks (cases) of food-borne illness in each food source except for school is not smooth across the cut-off in Figures 1.1.1 and 1.1.2 (Figure 1.2.1). It implies an external shock, including seasonal effects, might be occurred around the timing of the introduction of the EWS. In Figure 1.2.2, we find no drop in the total number of cases of food-borne illness in schools around the cut-off. For robustness checks, we change the bandwidth to 12 months, 24 months, and 73 months (see Figures A.1.1 to A.1.6). We find that the discontinuity in the outbreaks and cases in school is largely affected by the bandwidth and the order of the polynomial function. In addition, seasonal effects are not adjusted in the Figures. Therefore, we also quantitatively estimate the effects of the EWS based on equations (1) and (2).

Panel A in Table 1.2 (Table 1.3) presents the regression discontinuity (RD) estimates of the effects of the EWS on the number of outbreaks (cases) of food borne illness. Panel B in Tables 1.2 and 1.3 presents the differences-in-discontinuities estimates. In Tables 1.2 and 1.3, each column presents estimation results by type. We choose 12 months as the baseline bandwidth. Using the linear running variable, we find that the EWS significantly reduces the number of outbreaks of food-borne illness of the total type, pathogenic E. coli, and unknown type by 126%, 124%–129%, and 148%–163%, respectively (see models 1 and 3 in Table 1.2).[16] However, we could not find significant effects when we used the quadratic running variable (see

[16] The RD estimates of model 1 in column (1) of Table 1.2 are interpreted as the EWS is likely to reduce 1.087 outbreaks of food-borne illness per million people in the same region-month. The policy impacts are calculated like : (-1.087 / 0.863) × 100 = -126%.

models 2 and 4 in Table 1.2). Using the linear running variable, we find that the EWS significantly reduces the number of cases of the total type and pathogenic E. coli by 65% and 88%–92%, respectively (see models 1 and 3 in Table 1.3). However, we cannot find any significant evidence that the EWS reduces the number of cases of food-borne illness when we used the quadratic running variable (see models 2 and 4 in Table 1.3).

We also estimate the effects using the 24 months bandwidth (see Tables A.1.1 and A.1.2). In addition, we use Poisson regression with the 12 months bandwidth (see Tables A.1.3 to A.1.4) and 24 months bandwidth (see Tables A.1.5 to A.1.6) for an additional robustness check.[17] However, we cannot find any robust evidence that the EWS significantly reduces the number of outbreaks and cases of food-borne illness.

We also use the non-parametric RD (see Tables 1.4 and 1.5). and non-parametric differences-in-discontinuities approach (see Tables 1.6 and 1.7). We employ three kernel functions (i.e., triangular, epanechnikov, and uniform kernels) and data-driven bandwidth (see details Calonico et al., 2014). We find that the EWS significantly reduces the number of total outbreaks by 77%–84% (see models 3 and 5 in Table 1.4). However, the results are not robust to different kernel functions and differences-in-discontinuities approach. From non-parametric RD and non-parametric differences-in-discontinuities approach, we cannot find any evidence that the EWS significantly reduces the number of total cases of food-borne illness.

[17] We can use the number of outbreaks (cases) as a dependent variable instead of the number of outbreaks (cases) per million. In this case, the dependent variable is the count variable. If the dependent variable is the count variable, Poisson regression can be an alternative because a linear model might not provide the best fit (Wooldridge, 2019). "No scaling is needed for the Poisson regression models, as binary regressors in Poisson regressions can be interpretated as semi-elasticities due to the underlying assumption about the conditional mean." (DeAngelo and Hansen, 2014). Thus, the RD estimates and differences-in-discontinuities estimates in our Poisson regression models can be directly interpreted as the percentage change (i.e., semi-elasticities).

To sum up, some results show that the EWS significantly reduces the number of total outbreaks and cases of food-borne illness by at least 77%–84% and 65%, respectively, in the short run. However, those results are not robust to diverse identification strategies and model specifications.

Before evaluating the long-run effects using the difference-in-differences (DID), we visually and statistically test parallel trends. Figure 1.3 displays the event study estimates from equation (1.4). The estimates show that the treatment and control groups evolved similarly in the pre-treatment period. Table A.1.7 illustrates the joint significance tests for parallel trends in the pre-treatment period. It shows that parallel trends do not hold in the pre-treatment period for the number of outbreaks of food-borne illness of the total types, norovirus, pathogenic E. coli, salmonella, vibrio parahaemolyticus, and unknown type. In addition, it does not hold for the number of cases of food-borne illness of total types, norovirus, pathogenic E. coli, Salmonella, and unknown type. However, as we discussed in Section 1.3, high statistical power may detect even the slightest violations of parallel trends. Therefore, as we mentioned, we compare the estimates from a DID model with and without allowing group-specific time trends.

Tables 1.8 and 1.9 present the DID estimates without allowing group-specific time trends based on equation (1.3). In Tables 1.8 and 1.9, Panel A presents the DID estimates in Period 1 (Jan. 2002 – Oct. 2013), Panel B shows the DID estimates in Period 2 (Jan. 2002 – Aug. 2014), and Panel C presents the DID estimates in Period 3 (Jan. 2002 – Dec. 2020). Each period indicates any time from January 2002 to the first upgrade, the second upgrade, and December 2020 (i.e., the last month of the data period), respectively. In Table 1.8, we cannot find significant evidence that the EWS reduces the number of outbreaks of food-borne illness. In Table 1.9, the DID estimates show that the EWS significantly reduces the number of cases of food-borne illness of total types and unknow type by 28%−33% and 46%−60%. However, the results are not robust to Poisson regression (see Tables A.1.9).[18]

Tables 1.10 and 1.11 present the DID estimates, allowing group-specific time trends based on equation (1.5). Looking at the DID estimates in Table 1.10, the EWS significantly reduces the number of outbreaks of total types and pathogenic E. coli by 90%–91% and 172%−177%, respectively.[19] Looking at the DID estimates in Table 1.11, the EWS significantly reduces the number of cases of total types, norovirus, pathogenic E. coli, and vibrio parahaemolyticus by 81%–85%, 99%–117%, 161%–167%, and 103%, respectively. Comparing the DID estimates from Tables 1.8 to 1.11, we find that DID estimates with and without allowing group-specific time trends are not similar. Therefore, our preferred results are the DID estimates based on equation (1.5) relaxing the parallel trends assumption. However, the results are not robust to Poisson regression (see Tables A.1.10 and A.1.11). To summarize, we find some significant results that the EWS reduces the number of total outbreaks and cases of food-borne illness by 90%–91% and 81%–85%, respectively, in the long run. However, those results are not robust to diverse model specifications and the study windows.

1.5. DISCUSSION AND CONCLUSIONS

Food-borne illness not only threatens health but also imposes an economic burden worldwide. Additionally, food borne illness can inflict shock on food supply and market price,

[18] We use Poisson regression with the Mundlak approach because it provides a fully robust estimation (Wooldridge, 2021).

[19] If the outbreaks of food-borne illness in schools decrease while the outbreaks of food-borne illness in the control food sources increase before and after the policy change, the policy impact could be less than -100%. For example, [(90-100)-(300-200)]/100 = -110/100 = -1.1 = -110%. In addition, if the outbreaks of food-borne illness in the control food sources increase much more than that of schools, the policy impact could be less than -100%. For instance, [(120-100)-(400-200)]/100 = -180/100 = -1.8 = -180%.

aggravating food insecurity. Government interventions to control food-borne illness have been implemented and studied in response to this problem. The EWS is a government intervention implemented by the South Korean government to prevent the spread of food-borne illnesses. However, to our knowledge, there are no studies about the early warning system (EWS). To fill the research gap, this study evaluates the impacts of the EWS on the number of outbreaks and cases of food-borne illness using the monthly administrative panel data in South Korea for January 2002 – December 2020. We exploit the regression discontinuity (RD), difference-in-discontinuities, and difference-in-differences (DID) approaches to address endogeneity problems.

We find some results that the EWS significantly reduces the number of outbreaks and cases of food-borne illness in the short and long run. However, those results are not robust to diverse identification strategies, model specifications, and the study windows. In conclusion, we could not find strong evidence that the EWS significantly reduces the number of outbreaks and cases of food-borne illness. Why the EWS does not significantly and robustly prevent food-borne illness from spreading is unclear. However, below three hypotheses may provide plausible explanations.

First, the manual warning system may operate well to control the spread of food-borne illnesses. Before the introduction of the EWS, the school employees could report the occurrence of food-borne illnesses to other schools or relevant government authorities via e-mail, phone call, or mobile text messaging. If the advantage of introducing the EWS is not significantly larger than manual warning systems, it is hard to detect the significant effects of the EWS.

Second, insufficient school employee training for operating the EWS may cause insignificant effects of the EWS. Before the second upgrade, the South Korean government required school employees to manually enter the list of food items and suppliers into the EWS when they purchased. In addition, even after the second upgrade, the school employees still have to enter the information about the outbreak (e.g., symptoms and potentially hazardous foods (PHF)) into the system if food borne illness outbreaks occur. If school employees fail to take action promptly, the EWS could not send a warning text message or might provide insufficient information to other schools or relevant government authorities.

Third, if PHF is supplied to many schools on the same day and the students eat the PHF at the same time, the EWS is not likely to prevent the occurrence of food-borne illness. For example, a food company provided 17 elementary schools and kindergartens with school meals in Daejeon in South Korea on January 13, 2022.[20] After the children had lunch on the same day, they developed diarrhea and stomach cramps. On January 18, 2022, the MFDS announced that about 50 children from 9 elementary schools and kindergartens were infected with C. perfringens from the incidence.[21]

Our findings allow us to suggest several policy implications. First, if the manual warning system via e-mail, phone calls, or mobile text messaging is effectively operated, it may be an alternative to introducing the EWS. We could not find strong evidence that the EWS significantly reduces the incidence of food-borne illness in the short and long run. It does not necessarily mean that the introduction of the EWS systems has no effect on preventing the spread of food-borne illnesses. However, it suggests a possibility that there is no difference between the impacts of the manual warning system and that of the EWS. Therefore, if the cost of

---

[20] See news https://www.asiae.co.kr/article/2022011613520049893

[21] See press release

https://www.mfds.go.kr/brd/m_99/view.do?seq=46079&srchFr=&srchTo=&srchWord=%EC%8B%9D%EC%A4%91%EB%8F%85&srchTp=0&itm_seq_1=0&itm_seq_2=0&multi_itm_seq=0&company_cd=&company_nm=&Data_stts_gubun=C9999&page=2

introducing the EWS places a large financial burden on a government, they could consider maintaining the manual warning system, instead of introducing the EWS.

Second, school employee training for operating the EWS may help reduce the incidence of food-borne illness. As discussed above, operating the EWS requires school employees to take action in a timely manner. Therefore, if school employees do not have sufficient skills and know-how to operate the EWS, the effects of the EWS could be less effective. It implies that education or training for school employees may improve the performance of the EWS, although it is not testable in our study.

Third, the EWS is not likely to prevent the spread of food-borne illness if PHF is supplied and consumed on the same day in many schools. Therefore, it is important to reduce the risk of PHF before they are supplied to students. For example, the government can reinforce regular sanitary inspections of food companies, school kitchens, and cafeterias. Additionally, they can tighten the food safety laws to make food companies thoroughly inspect food safety themselves. Those precautionary measures may lower the risk of food-borne illness in the early stage.

Lastly, as we mentioned, our results should not be interpreted as the EWS cannot lower the incidence of food-borne illness at all. Instead, we should ask why we could not detect strong evidence that the EWS reduces the incidence of food-borne illness. It implies that future studies need to explore what factors weaken the impacts of the EWS. As we discussed, the factors may be school employees' skills and technical know-how or PHF supplied and consumed on the same day among many schools.

Table 1.1. Descriptive statistics

Variable	Mean	Std. Dev.	Min	Max
Reported outbreaks per million				
Total	0.545	0.915	0.000	17.264
Food source				
Schools	0.554	1.827	0.000	28.189
Institutions	0.034	0.152	0.000	2.658
Restaurants	0.302	0.691	0.000	17.264
Homes	0.022	0.145	0.000	3.568
Other food sources	0.076	0.243	0.000	3.307
Unknown food sources	0.033	0.176	0.000	3.547
Type				
Norovirus	0.073	0.257	0.000	3.568
Pathogenic E. coli	0.069	0.256	0.000	3.458
Salmonella	0.047	0.203	0.000	3.626
Vibrio parahaemolyticus	0.037	0.245	0.000	7.847
Staphylococcus aureus	0.024	0.128	0.000	1.784
Protozoa	0.020	0.127	0.000	3.117
Campylobacter jejuni	0.015	0.087	0.000	1.078
C. perfringens	0.016	0.105	0.000	1.806
Bacillus cereus	0.014	0.101	0.000	1.621
Unknown type	0.214	0.498	0.000	7.847
Reported cases per million				
Total	12.655	33.606	0.000	638.100
Food source				
Schools	36.394	160.300	0.000	3508.678
Institutions	1.489	10.293	0.000	275.206
Restaurants	3.762	16.361	0.000	377.543
Homes	0.155	1.475	0.000	39.887
Other food sources	1.537	9.515	0.000	232.741
Unknown food sources	0.509	5.207	0.000	175.592

Type				
Norovirus	2.406	14.236	0.000	274.453
Pathogenic E. coli	3.280	21.094	0.000	402.197
Salmonella	1.401	9.276	0.000	232.358
Vibrio parahaemolyticus	0.623	5.511	0.000	180.151
Staphylococcus aureus	0.784	6.640	0.000	182.910
Protozoa	0.112	0.805	0.000	18.703
Campylobacter jejuni	0.616	6.110	0.000	172.881
C. perfringens	0.628	7.199	0.000	209.029
Bacillus cereus	0.219	2.576	0.000	83.924
Unknown type	2.169	7.785	0.000	140.402
Male (%)	50.092	0.501	48.629	51.554
Under 20 years (%)	22.846	3.682	14.538	31.504
20−44 years (%)	37.974	4.498	27.312	47.282
45−64 years (%)	27.275	4.334	17.433	34.794
65 years and over (%)	11.905	4.079	4.341	23.541
Elementary school students (%)	47.747	3.797	40.766	57.258
Middle school students (%)	25.845	1.580	22.131	28.494
High school students (%)	26.408	2.985	19.699	31.542
Unemployment rate (%)	3.267	1.069	0.800	7.100
Population (millions of individuals)	3.327	3.222	0.549	13.427
Students (millions of individuals)	0.451	0.434	0.079	1.853

Notes: This table indicates descriptive statistics for 14 regions in South Korea for the period Jan. 2002 – Dec. 2020 (N = 3,192).

Table 1.2. RD and differences-in-discontinuities estimates of the EWS on the number of outbreaks of food-borne illness per million (bandwidth: 12 months)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Total	Norovirus	E. coli	Salmonella	Vibrio	S. aureus	C. jejuni	Perfringens	B. cereus	Unknown
Panel A. RD estimates										
Model 1. linear running variable										
Post	-1.087**	-0.152	-0.387*	-0.010	-0.026	-0.118	0.063	-0.170	-0.002	-0.283*
	(0.429)	(0.258)	(0.208)	(0.010)	(0.026)	(0.129)	(0.071)	(0.184)	(0.046)	(0.134)
% Change from Mean	-126%	-79%	-124%	-74%	-229%	-98%	261%	-	-13%	-163%
Obs.	350	350	350	350	350	350	350	350	350	350
Model 2. quadratic running variable										
Post	-0.860	-0.554	0.168	0.017	-0.003	0.031	-0.047	-0.093	-0.175	-0.203
	(0.845)	(0.387)	(0.239)	(0.018)	(0.003)	(0.194)	(0.096)	(0.211)	(0.124)	(0.284)
% Change from Mean	-100%	-287%	54%	127%	-26%	26%	-195%	-	-1,156%	-117%
Obs.	350	350	350	350	350	350	350	350	350	350
Panel B. differences-in-discontinuities estimates										
Model 3. linear running variable										
Post × School	-1.089**	-0.158	-0.401*	-0.015	-0.024	-0.131	0.064	-0.168	-0.003	-0.257*
	(0.422)	(0.255)	(0.210)	(0.018)	(0.029)	(0.128)	(0.071)	(0.182)	(0.045)	(0.121)
% Change from Mean	-126%	-82%	-129%	-112%	-212%	-109%	265%	-	-20%	-148%
Obs.	1,400	1,400	1,400	1,400	1,400	1,400	1,400	1,400	1,400	1,400
Model 4. quadratic running variable										
Post × School	-0.839	-0.565	0.177	0.015	-0.008	0.035	-0.050	-0.095	-0.181	-0.167
	(0.848)	(0.386)	(0.238)	(0.019)	(0.010)	(0.192)	(0.095)	(0.209)	(0.123)	(0.273)
% Change from Mean	-97%	-293%	57%	112%	-71%	29%	-207%	-	-1,195%	-96%
Obs.	1,400	1,400	1,400	1,400	1,400	1,400	1,400	1,400	1,400	1,400
Pre-Period Outcome Mean	0.863	0.193	0.312	0.013	0.011	0.120	0.024	0.000	0.015	0.174

Notes: This table indicates OLS estimates. Columns indicate different outcome variables. Panel A reports RD estimates from Eq. (1.1). Panel B reports differences-in-discontinuities estimates from Eq. (1.2). We do not have enough variation in Protozoa. Standard errors clustered at the region level are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 1.3. RD and differences-in-discontinuities estimates of the EWS on the number of cases of food-borne illness per million (bandwidth: 12 months)

	(1) Total	(2) Norovirus	(3) E. coli	(4) Salmonella	(5) Vibrio	(6) S. aureus	(7) C. jejuni	(8) Perfringens	(9) B. cereus	(10) Unknown
Panel A. RD estimates										
Model 1. linear running variable										
Post	-29.013	6.325	-17.699*	-0.077	-3.322	-4.678	-3.857	-2.795	0.113	-3.023
	(16.852)	(8.846)	(8.964)	(0.078)	(3.391)	(9.052)	(5.678)	(3.972)	(0.743)	(5.009)
% Change from Mean	-63%	121%	-88%	-72%	-229%	-78%	-73%	-	15%	-43%
Obs.	350	350	350	350	350	350	350	350	350	350
Model 2. quadratic running variable										
Post	28.688	-1.005	28.127	0.140	-0.398	-3.414	10.665	12.472	-8.803	-9.096
	(36.902)	(14.566)	(17.541)	(0.143)	(0.407)	(12.810)	(12.999)	(17.723)	(7.634)	(10.929)
% Change from Mean	62%	-19%	140%	130%	-27%	-57%	203%	-	-1163%	-129%
Obs.	350	350	350	350	350	350	350	350	350	350
Panel B. differences-in-discontinuities estimates										
Model 3. linear running variable										
Post × School	-29.775*	6.382	-18.494*	0.019	-3.408	-4.683	-3.841	-2.787	0.108	-3.137
	(16.488)	(8.694)	(9.068)	(0.267)	(3.392)	(8.968)	(5.622)	(3.928)	(0.736)	(5.157)
% Change from Mean	-65%	122%	-92%	18%	-235%	-78%	-73%	-	14%	-45%
Obs.	1,400	1,400	1,400	1,400	1,400	1,400	1,400	1,400	1,400	1,400
Model 4. quadratic running variable										
Post × School	29.684	-0.701	29.064	-0.207	-0.261	-3.443	10.632	12.302	-8.849	-8.857
	(36.909)	(14.538)	(17.522)	(0.275)	(0.549)	(12.666)	(12.862)	(17.545)	(7.550)	(11.007)
% Change from Mean	65%	-13%	144%	-193%	-18%	-58%	202%	-	-1169%	-126%
Obs.	1,400	1,400	1,400	1,400	1,400	1,400	1,400	1,400	1,400	1,400
Pre-Period Outcome Mean	45.995	5.242	20.148	0.107	1.453	5.981	5.266	0.000	0.757	7.040

Notes: This table indicates OLS estimates. Columns indicate different outcome variables. Panel A reports RD estimates from Eq. (1.1). Panel B reports differences-in-discontinuities estimates from Eq. (1.2). We do not have enough variation in Protozoa. Standard errors clustered at the region level are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 1.4. Non-parametric RD estimates of the EWS on the number of outbreaks of food-borne illness per million (data-driven bandwidth)

	(1) Total	(2) Norovirus	(3) E. coli	(4) Salmonella	(5) Vibrio	(6) S. aureus	(7) C. jejuni	(8) Perfringens	(9) B. cereus	(10) Unknown
Panel A. linear running variable										
Model 1. Triangular kernel										
RD estimates	-0.486	0.191	-0.353*	-0.019***	-0.003	-0.039	0.029	0.040	0.044	-0.068
	(0.338)	(0.141)	(0.190)	(0.005)	(0.031)	(0.103)	(0.039)	(0.073)	(0.069)	(0.114)
% Change from Mean	-60%	83%	-146%	-259%	-17%	-45%	248%	266%	702%	-32%
Pre-Period Outcome Mean	0.815	0.231	0.242	0.007	0.017	0.087	0.012	0.015	0.006	0.216
Bandwidth	18.74	17.74	27.41	22.17	22.80	28.01	39.02	27.58	29.93	25.22
Eff. Number of obs.	518	490	770	630	630	798	1106	770	826	714
Model 2. Epanechnikov kernel										
RD estimates	-0.465	0.252*	-0.362*	-0.021***	-0.005	-0.045	0.032	0.043	0.046	-0.067
	(0.361)	(0.143)	(0.193)	(0.007)	(0.034)	(0.114)	(0.050)	(0.067)	(0.069)	(0.133)
% Change from Mean	-58%	106%	-138%	-287%	-25%	-50%	266%	276%	683%	-27%
Pre-Period Outcome Mean	0.805	0.239	0.262	0.007	0.020	0.091	0.012	0.016	0.007	0.245
Bandwidth	16.46	16.48	25.33	22.26	19.71	27.76	38.71	26.31	27.42	22.65
Eff. Number of obs.	462	462	714	630	546	770	1078	742	770	630
Model 3. Uniform kernel										
RD estimates	-0.635**	-0.049	-0.287	0.010	0.000	-0.069	0.027	0.063	0.031	-0.209*
	(0.299)	(0.161)	(0.194)	(0.011)	(0.041)	(0.125)	(0.047)	(0.112)	(0.061)	(0.113)
% Change from Mean	-77%	-23%	-92%	-	0%	-88%	298%	-	426%	-130%
Pre-Period Outcome Mean	0.823	0.214	0.312	0.000	0.021	0.078	0.009	0.000	0.007	0.160
Bandwidth	25.55	23.57	21.89	8.22	18.34	32.98	32.35	11.72	25.13	13.84
Eff. Number of obs.	714	658	602	238	518	910	910	322	714	378
Panel B. quadratic running variable										
Model 4. Triangular kernel										
RD estimates	-0.574	0.226	-0.379*	0.010	-0.008	0.021	0.032	0.028	0.035	-0.206***
	(0.418)	(0.154)	(0.226)	(0.020)	(0.025)	(0.092)	(0.030)	(0.116)	(0.082)	(0.076)
% Change from Mean	-64%	122%	-168%	130%	-80%	26%	254%	214%	329%	-103%
Pre-Period Outcome Mean	0.895	0.186	0.225	0.008	0.010	0.080	0.013	0.013	0.011	0.200
Bandwidth	23.66	27.77	30.91	21.70	38.68	33.00	23.39	31.08	39.04	27.96
Eff. Number of obs.	658	770	854	602	1078	938	658	882	1106	770
Model 5. Epanechnikov kernel										
RD estimates	-0.809*	0.226	-0.432*	0.010	-0.002	0.005	0.040	0.015	0.042	-0.020
	(0.481)	(0.164)	(0.241)	(0.024)	(0.034)	(0.105)	(0.030)	(0.129)	(0.083)	(0.113)

% Change from Mean	-84%	110%	-172%	124%	-20%	6%	304%	100%	832%	-12%
Pre-Period Outcome Mean	0.961	0.205	0.252	0.008	0.010	0.082	0.013	0.015	0.005	0.169
Bandwidth	21.35	24.82	26.91	20.01	37.28	35.57	22.82	27.00	36.95	32.01
Eff. Number of obs.	602	686	742	574	1050	994	630	770	1022	910
Model 6. Uniform kernel										
RD estimates	-0.832	0.307**	-0.602**	0.005	0.011	0.010	0.031	0.007	0.026	-0.014
	(0.628)	(0.154)	(0.252)	(0.011)	(0.037)	(0.108)	(0.025)	(0.134)	(0.075)	(0.120)
% Change from Mean	-96%	149%	-193%	68%	87%	11%	235%	43%	250%	-7%
Pre-Period Outcome Mean	0.867	0.205	0.312	0.007	0.013	0.087	0.013	0.016	0.010	0.216
Bandwidth	19.74	24.16	21.12	22.26	30.50	28.22	22.06	25.02	40.35	25.42
Eff. Number of obs.	546	686	602	630	854	798	630	714	1134	714

Notes: Columns indicate different outcome variables. We adopted mserd bandwidth selector (See Calonico et al., 2017). Seasonal effects were controlled in all models. We do not have enough variation in Protozoa. Standard errors clustered at the region level are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 1.5. Non-parametric RD estimates of the EWS on the number of cases of food-borne illness per million (data-driven bandwidth)

	(1) Total	(2) Norovirus	(3) E. coli	(4) Salmonella	(5) Vibrio	(6) S. aureus	(7) C. jejuni	(8) Perfringens	(9) B. cereus	(10) Unknown
Panel A. linear running variable										
Model 1. Triangular kernel										
RD estimates	24.155	42.817***	-6.048	-0.052	-1.904	2.148	-5.480***	6.463	2.890	-6.636*
	(27.879)	(10.004)	(13.355)	(0.051)	(3.325)	(7.993)	(1.517)	(11.409)	(4.583)	(3.976)
% Change from Mean	36%	282%	-22%	-113%	-173%	40%	-182%	756%	1145%	-82%
Pre-Period Outcome Mean	67.308	15.206	27.910	0.046	1.098	5.326	3.009	0.855	0.252	8.131
Bandwidth	22.70	15.73	24.72	28.38	28.19	43.74	21.36	35.01	36.96	22.21
Eff. Number of obs.	630	434	686	798	798	1218	602	994	1022	630
Model 2. Epanechnikov kernel										
RD estimates	10.873	45.044***	-5.096	-0.034	-1.773	2.605	-7.289**	6.585	2.955	-8.286*
	(29.214)	(10.411)	(14.903)	(0.044)	(3.627)	(8.419)	(3.567)	(10.744)	(4.541)	(4.851)
% Change from Mean	22%	282%	-17%	-79%	-133%	41%	-277%	682%	1073%	-106%
Pre-Period Outcome Mean	50.504	15.948	30.447	0.043	1.337	6.378	2.633	0.965	0.275	7.812
Bandwidth	19.27	14.99	22.67	30.63	23.59	34.43	24.90	31.27	33.59	20.10
Eff. Number of obs.	546	406	630	854	658	966	686	882	938	574
Model 3. Uniform kernel										
RD estimates	-17.242	6.716	-1.757	-0.011	-1.519	1.053	-5.754	10.033	3.002	-6.997
	(25.530)	(7.026)	(16.135)	(0.270)	(3.602)	(7.901)	(5.211)	(12.634)	(4.244)	(4.354)
% Change from Mean	-32%	50%	-8%	-9%	-99%	19%	-296%	-	958%	-114%
Pre-Period Outcome Mean	53.190	13.558	22.549	0.129	1.538	5.585	1.947	0.000	0.313	6.131
Bandwidth	31.41	30.19	19.56	10.99	20.15	41.84	36.34	22.07	29.75	19.63
Eff. Number of obs.	882	854	546	294	574	1162	1022	630	826	546
Panel B. quadratic running variable										
Model 4. Triangular kernel										
RD estimates	15.491	37.287***	-7.394	-0.096	-2.075	-0.188	5.550	11.839	3.682	-3.028
	(35.030)	(12.481)	(10.658)	(0.433)	(3.511)	(7.724)	(8.185)	(19.667)	(5.854)	(4.057)
% Change from Mean	26%	206%	-32%	-208%	-324%	-3%	176%	1306%	563%	-50%
Pre-Period Outcome Mean	60.428	18.070	23.050	0.046	0.641	6.258	3.159	0.907	0.654	6.085
Bandwidth	25.34	21.99	32.27	28.25	48.96	33.70	20.47	33.85	40.78	33.27
Eff. Number of obs.	714	602	910	798	1358	938	574	938	1134	938
Model 5. Epanechnikov kernel										
RD estimates	-12.774	50.738***	0.447	0.015	-2.157	-0.025	5.068	11.519	3.944	-5.697
	(40.583)	(11.013)	(13.673)	(0.539)	(3.316)	(8.998)	(7.367)	(20.087)	(6.018)	(4.480)

% Change from Mean	-18%	419%	2%	28%	-365%	0%	160%	1155%	1649%	-92%
Pre-Period Outcome Mean	70.513	12.108	22.254	0.054	0.591	6.378	3.159	0.997	0.239	6.169
Bandwidth	21.14	19.46	34.26	24.47	52.54	34.07	20.86	30.88	38.64	29.74
Eff. Number of obs.	602	546	966	686	1470	966	574	854	1078	826
Model 6. Uniform kernel										
RD estimates	1.177	35.600***	-7.950	-0.428	-2.416	2.293	1.717	12.526	3.199	-0.128
	(35.345)	(12.344)	(12.473)	(0.354)	(3.749)	(9.207)	(2.879)	(19.892)	(5.441)	(6.309)
% Change from Mean	2%	271%	-30%	-730%	-283%	40%	57%	1088%	514%	-2%
Pre-Period Outcome Mean	50.504	13.121	26.793	0.059	0.854	5.725	3.009	1.151	0.623	5.345
Bandwidth	19.77	31.06	25.05	22.34	36.91	40.94	21.76	26.36	42.28	17.20
Eff. Number of obs.	546	882	714	630	1022	1134	602	742	1190	490

Notes: Columns indicate different outcome variables. We adopted mserd bandwidth selector (See Calonico et al., 2017). Seasonal effects were controlled in all models. We do not have enough variation in Protozoa. Standard errors clustered at the region level are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 1.6. Non-parametric differences-in-discontinuities estimates of the EWS on the number of outbreaks of food-borne illness per million (data-driven bandwidth)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Total	Norovirus	E. coli	Salmonella	Vibrio	S. aureus	C. jejuni	Perfringens	B. cereus	Unknown
Panel A. linear running variable										
Model 1. Triangular kernel										
Diff-in-disc estimates	-0.352	0.111	-0.321	0.016	-0.012	-0.023	0.025	0.049	0.019	-0.040
	(0.293)	(0.152)	(0.198)	(0.011)	(0.026)	(0.103)	(0.045)	(0.085)	(0.069)	(0.094)
% Change from Mean	-47%	51%	-128%	149%	-92%	-25%	235%	302%	314%	-15%
Pre-Period Outcome Mean	0.750	0.218	0.252	0.011	0.013	0.091	0.011	0.016	0.006	0.263
Bandwidth	28.99	18.63	26.71	15.10	29.14	27.68	43.14	25.40	30.99	20.78
Eff. Number of obs.	3192	2072	2968	1736	3304	3080	4872	2856	3416	2296
Model 2. Epanechnikov kernel										
Diff-in-disc estimates	-0.386	0.124	-0.326	0.019	-0.004	-0.039	0.026	0.041	0.023	-0.081
	(0.295)	(0.142)	(0.203)	(0.013)	(0.033)	(0.107)	(0.048)	(0.090)	(0.070)	(0.094)
% Change from Mean	-50%	57%	-125%	165%	-22%	-43%	222%	-	354%	-56%
Pre-Period Outcome Mean	0.778	0.218	0.262	0.012	0.018	0.091	0.012	0.000	0.006	0.145
Bandwidth	27.38	18.30	25.18	14.06	21.26	27.09	39.43	19.52	28.95	18.25
Eff. Number of obs.	3080	2072	2856	1624	2408	3080	4424	2184	3192	2072
Model 3. Uniform kernel										
Diff-in-disc estimates	-0.178	0.182	-0.315	0.008	-0.012	-0.007	0.047	0.041	0.013	-0.099
	(0.361)	(0.144)	(0.198)	(0.012)	(0.022)	(0.115)	(0.050)	(0.097)	(0.074)	(0.105)
% Change from Mean	-19%	71%	-111%	-	-86%	-6%	486%	-	143%	-62%
Pre-Period Outcome Mean	0.961	0.255	0.285	0.000	0.014	0.109	0.010	0.000	0.009	0.160
Bandwidth	21.03	15.95	23.67	7.52	27.65	21.48	30.57	14.44	20.42	13.68
Eff. Number of obs.	2408	1736	2632	840	3080	2408	3416	1624	2296	1512
Panel B. quadratic running variable										
Model 4. Triangular kernel										
Diff-in-disc estimates	-0.364	0.121	-0.323	0.013	-0.005	0.035	-0.000	0.031	0.004	-0.084
	(0.441)	(0.214)	(0.220)	(0.022)	(0.030)	(0.107)	(0.034)	(0.123)	(0.076)	(0.086)
% Change from Mean	-40%	61%	-138%	121%	-52%	45%	0%	214%	45%	-39%
Pre-Period Outcome Mean	0.917	0.197	0.234	0.011	0.010	0.078	0.015	0.014	0.009	0.216
Bandwidth	22.85	25.78	28.17	15.63	39.73	32.46	19.77	28.10	47.32	26.00
Eff. Number of obs.	2520	2856	3192	1736	4424	3640	2184	3192	5320	2856
Model 5. Epanechnikov kernel										

Diff-in-disc estimates	-0.424	0.159	-0.360	0.015	-0.003	0.024	0.010	0.029	0.010	-0.127
	(0.462)	(0.217)	(0.239)	(0.025)	(0.033)	(0.109)	(0.030)	(0.122)	(0.075)	(0.105)
% Change from Mean	-44%	77%	-149%	130%	-29%	30%	69%	179%	111%	-57%
Pre-Period Outcome Mean	0.961	0.205	0.242	0.012	0.011	0.081	0.014	0.016	0.009	0.225
Bandwidth	21.61	24.26	27.54	14.78	36.27	31.51	20.38	25.70	46.66	24.58
Eff. Number of obs.	2408	2744	3080	1624	4088	3528	2296	2856	5208	2744
Model 6. Uniform kernel										
Diff-in-disc estimates	-0.474	0.189	-0.514*	0.027	-0.000	0.010	0.008	0.053	0.014	-0.340
	(0.523)	(0.209)	(0.276)	(0.034)	(0.035)	(0.113)	(0.035)	(0.131)	(0.080)	(0.275)
% Change from Mean	-55%	84%	-157%	218%	0%	11%	47%	-	247%	-147%
Pre-Period Outcome Mean	0.867	0.224	0.327	0.012	0.012	0.087	0.017	0.000	0.006	0.232
Bandwidth	19.47	22.37	20.35	13.38	31.98	28.03	17.32	15.63	32.02	19.94
Eff. Number of obs.	2184	2520	2296	1512	3528	3192	1960	1736	3640	2184

Notes: Columns indicate different outcome variables. We adopted CCT bandwidth selector (See Calonico et al., 2014). Seasonal effects were controlled in all models. We do not have enough variation in Protozoa. Standard errors clustered at the region level are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 1.7. Non-parametric differences-in-discontinuities estimates of the EWS on the number of cases of food-borne illness per million (data-driven bandwidth)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Total	Norovirus	E. coli	Salmonella	Vibrio	S. aureus	C. jejuni	Perfringens	B. cereus	Unknown
Panel A. linear running variable										
Model 1. Triangular kernel										
Diff-in-disc estimates	23.759	22.569**	-4.972	-1.355***	-2.860	3.118	-3.395	11.914	2.778	-4.580
	(23.151)	(9.252)	(10.776)	(0.151)	(2.860)	(8.689)	(5.270)	(14.590)	(4.903)	(4.353)
% Change from Mean	47%	148%	-22%	-	-326%	52%	-181%	1035%	978%	-75%
Pre-Period Outcome Mean	50.504	15.206	22.549	0.000	0.879	5.963	1.876	1.151	0.284	6.131
Bandwidth	19.47	15.07	19.96	7.79	35.21	31.93	53.02	26.32	32.99	19.43
Eff. Number of obs.	2184	1736	2184	840	3976	3528	5992	2968	3640	2184
Model 2. Epanechnikov kernel										
Diff-in-disc estimates	26.330	24.057***	-5.734	-1.476***	-2.613	3.428	-3.349	11.773	3.250	-4.492
	(25.199)	(9.263)	(12.008)	(0.172)	(3.419)	(8.997)	(5.386)	(13.935)	(4.895)	(4.930)
% Change from Mean	37%	151%	-17%	-	-221%	54%	-168%	905%	1073%	-83%
Pre-Period Outcome Mean	70.513	15.948	33.492	0.000	1.183	6.375	1.988	1.301	0.303	5.439
Bandwidth	21.44	14.74	20.08	7.29	26.54	29.93	50.74	23.88	30.03	18.64
Eff. Number of obs.	2408	1624	2296	840	2968	3304	5656	2632	3416	2072
Model 3. Uniform kernel										
Diff-in-disc estimates	20.316	15.605*	15.105	-1.668***	-1.548	5.478	-3.386	11.572	3.544	-4.494
	(25.720)	(9.446)	(21.272)	(0.349)	(3.332)	(10.176)	(5.398)	(14.177)	(4.967)	(5.050)
% Change from Mean	40%	661%	45%	-	-106%	78%	-172%	-	858%	-74%
Pre-Period Outcome Mean	50.408	2.362	33.492	0.000	1.464	7.061	1.974	0.000	0.413	6.058
Bandwidth	15.27	10.49	20.47	5.82	21.61	21.29	39.85	18.75	22.72	15.60
Eff. Number of obs.	1736	1176	2296	616	2408	2408	4424	2072	2520	1736
Panel B. quadratic running variable										
Model 4. Triangular kernel										
Diff-in-disc estimates	22.368	15.141	-11.313	-1.418***	-2.179	3.632	7.376	14.841	2.402	-5.691
	(31.430)	(12.048)	(13.026)	(0.240)	(3.270)	(8.952)	(7.621)	(21.736)	(5.482)	(4.813)
% Change from Mean	31%	92%	-39%	-1649%	-340%	69%	152%	1339%	413%	-76%
Pre-Period Outcome Mean	71.928	16.499	29.123	0.086	0.641	5.235	4.860	1.108	0.581	7.454
Bandwidth	20.33	23.19	23.50	15.62	48.10	49.31	13.17	27.37	45.94	24.64
Eff. Number of obs.	2296	2632	2632	1736	5432	5544	1512	3080	5096	2744
Model 5. Epanechnikov kernel										

Diff-in-disc estimates	30.795	16.423	-18.653	-1.521***	-2.247	3.938	9.019	14.883	2.773	-6.117
	(34.087)	(10.707)	(16.500)	(0.274)	(3.446)	(9.357)	(10.985)	(21.917)	(5.615)	(5.524)
% Change from Mean	59%	136%	-61%	-1651%	-314%	71%	214%	1144%	445%	-82%
Pre-Period Outcome Mean	52.277	12.108	30.447	0.092	0.715	5.577	4.212	1.301	0.623	7.454
Bandwidth	18.68	19.92	22.10	14.89	43.80	46.16	15.58	23.88	42.28	24.67
Eff. Number of obs.	2072	2184	2520	1624	4872	5208	1736	2632	4760	2744
Model 6. Uniform kernel										
Diff-in-disc estimates	36.667	9.729	7.204	-1.620***	-2.196	5.421	4.404	14.658	3.459	-4.679
	(33.674)	(19.957)	(14.309)	(0.332)	(3.546)	(9.651)	(5.424)	(21.231)	(5.382)	(6.170)
% Change from Mean	78%	51%	27%	-1507%	-257%	93%	126%	-	1370%	-86%
Pre-Period Outcome Mean	47.258	18.974	26.793	0.107	0.854	5.861	3.510	0.000	0.252	5.439
Bandwidth	16.48	20.90	25.45	12.29	36.38	37.03	18.50	21.01	36.35	18.09
Eff. Number of obs.	1848	2296	2856	1400	4088	4200	2072	2408	4088	2072

Notes: Columns indicate different outcome variables. We adopted CCT bandwidth selector (See Calonico et al., 2014). Seasonal effects were controlled in all models. We do not have enough variation in Protozoa. Standard errors clustered at the region level are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 1.8. DID estimates of the EWS on the number of outbreaks of food-borne illness per million (without differential linear time trends)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Total	Norovirus	E. coli	Salmonella	Vibrio	S. aureus	C. jejuni	Perfringens	B. cereus	Unknown
Panel A. Period 1 (Jan. 2002 – Oct. 2013)										
Post × School	-0.037	0.050	0.036	-0.017	-0.006	-0.039	0.022*	0.001	0.010	-0.051
	(0.123)	(0.038)	(0.047)	(0.010)	(0.004)	(0.029)	(0.011)	(0.015)	(0.010)	(0.052)
% Change from Mean	-7%	44%	25%	-83%	-85%	-69%	233%	4%	113%	-34%
Obs.	7,952	7,952	7,952	7,952	7,952	7,952	7,952	7,952	7,952	7,952
Panel B. Period 2 (Jan. 2002 – Aug. 2014)										
Post × School	-0.000	0.057	0.055	-0.016	-0.005	-0.038	0.021**	0.007	0.008	-0.049
	(0.131)	(0.040)	(0.062)	(0.010)	(0.004)	(0.029)	(0.008)	(0.015)	(0.009)	(0.048)
% Change from Mean	0%	50%	39%	-78%	-71%	-68%	222%	30%	90%	-33%
Obs.	8,512	8,512	8,512	8,512	8,512	8,512	8,512	8,512	8,512	8,512
Panel C. Period 3 (Jan. 2002 – Dec. 2020)										
Post × School	-0.024	0.017	0.060	-0.013	-0.004	-0.039	0.037***	-0.003	0.009	-0.050
	(0.118)	(0.036)	(0.040)	(0.011)	(0.003)	(0.028)	(0.009)	(0.013)	(0.010)	(0.040)
% Change from Mean	-4%	15%	42%	-63%	-57%	-69%	391%	-13%	102%	-34%
Obs.	12,768	12,768	12,768	12,768	12,768	12,768	12,768	12,768	12,768	12,768
Pre-Period Outcome Mean	0.573	0.114	0.142	0.021	0.007	0.056	0.009	0.023	0.009	0.149

Notes: This table indicates OLS estimates. Columns indicate different outcome variables. Each panel reports DID estimates from Eq. (1.3). We do not have enough variation in Protozoa. Standard errors clustered at the region level are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 1.9. DID estimates of the EWS on the number of cases of food-borne illness per million (without differential linear time trends)

	(1) Total	(2) Norovirus	(3) E. coli	(4) Salmonella	(5) Vibrio	(6) S. aureus	(7) C. jejuni	(8) Perfringens	(9) B. cereus	(10) Unknown
Panel A. Period 1 (Jan. 2002 – Oct. 2013)										
Post × School	-13.129*	-1.333	-1.086	-1.388	-1.062	-2.512	1.348	-0.013	0.059	-3.586*
	(7.256)	(3.989)	(3.191)	(0.891)	(0.690)	(2.383)	(1.872)	(1.205)	(0.898)	(1.979)
% Change from Mean	-28%	-13%	-8%	-88%	-98%	-59%	99%	-1%	7%	-46%
Obs.	7,952	7,952	7,952	7,952	7,952	7,952	7,952	7,952	7,952	7,952
Panel B. Period 2 (Jan. 2002 – Aug. 2014)										
Post × School	-10.546	-1.493	0.703	-0.699	-1.056	-2.675	1.374	0.367	-0.058	-3.680*
	(7.554)	(3.839)	(5.009)	(1.312)	(0.690)	(2.327)	(1.525)	(1.200)	(0.869)	(1.869)
% Change from Mean	-22%	-15%	5%	-44%	-98%	-63%	101%	17%	-7%	-48%
Obs.	8,512	8,512	8,512	8,512	8,512	8,512	8,512	8,512	8,512	8,512
Panel C. Period 3 (Jan. 2002 – Dec. 2020)										
Post × School	-15.757**	-3.365	-0.580	-0.856	-1.041	-3.122	1.829**	-0.750	-0.203	-4.664***
	(6.906)	(3.386)	(3.731)	(1.106)	(0.690)	(2.165)	(0.734)	(0.802)	(0.831)	(1.520)
% Change from Mean	-33%	-33%	-4%	-54%	-96%	-74%	134%	-34%	-23%	-60%
Obs.	12,768	12,768	12,768	12,768	12,768	12,768	12,768	12,768	12,768	12,768
Pre-Period Outcome Mean	47.348	10.131	14.526	1.582	1.084	4.242	1.362	2.190	0.878	7.729

Notes: This table indicates OLS estimates. Columns indicate different outcome variables. Each panel reports DID estimates from Eq. (1.3). We do not have enough variation in Protozoa. Standard errors clustered at the region level are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 1.10. DID estimates of the EWS on the number of outbreaks of food-borne illness per million (with differential linear time trends)

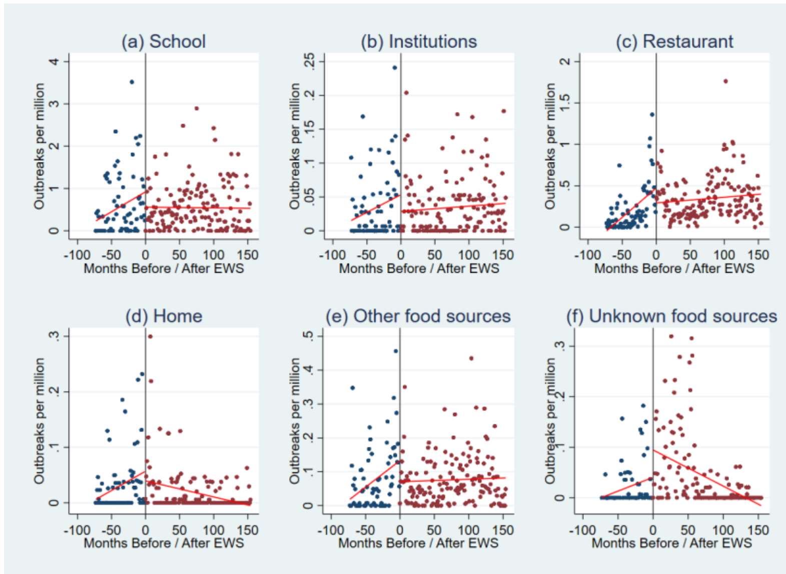
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Total	Norovirus	E. coli	Salmonella	Vibrio	S. aureus	C. jejuni	Perfringens	B. cereus	Unknown
Panel A. Period 1 (Jan. 2002 – Oct. 2013)										
Post × School	-0.513*	-0.143	-0.245**	-0.016	-0.012	-0.059	0.009	0.013	0.011	-0.074
	(0.271)	(0.085)	(0.098)	(0.011)	(0.011)	(0.079)	(0.028)	(0.029)	(0.019)	(0.118)
% Change from Mean	-90%	-126%	-172%	-78%	-170%	-105%	95%	56%	124%	-50%
Obs.	7,952	7,952	7,952	7,952	7,952	7,952	7,952	7,952	7,952	7,952
Panel B. Period 2 (Jan. 2002 – Aug. 2014)										
Post × School	-0.521*	-0.126	-0.252**	-0.018	-0.012	-0.059	0.013	-0.000	0.016	-0.075
	(0.261)	(0.072)	(0.104)	(0.010)	(0.010)	(0.073)	(0.029)	(0.030)	(0.017)	(0.118)
% Change from Mean	-91%	-111%	-177%	-87%	-170%	-105%	137%	0%	181%	-50%
Obs.	8,512	8,512	8,512	8,512	8,512	8,512	8,512	8,512	8,512	8,512
Panel C. Period 3 (Jan. 2002 – Dec. 2020)										
Post × School	-0.103	0.060	-0.040	-0.016	-0.009	-0.034	0.008	0.016	0.004	-0.059
	(0.176)	(0.045)	(0.080)	(0.010)	(0.006)	(0.046)	(0.013)	(0.021)	(0.016)	(0.082)
% Change from Mean	-18%	53%	-28%	-78%	-127%	-60%	85%	69%	45%	-40%
Obs.	12,768	12,768	12,768	12,768	12,768	12,768	12,768	12,768	12,768	12,768
Pre-Period Outcome Mean	0.573	0.114	0.142	0.021	0.007	0.056	0.009	0.023	0.009	0.149

Notes: This table indicates OLS estimates. Columns indicate different outcome variables. Each panel reports DID estimates from Eq. (1.5). We do not have enough variation in Protozoa. Standard errors clustered at the region level are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

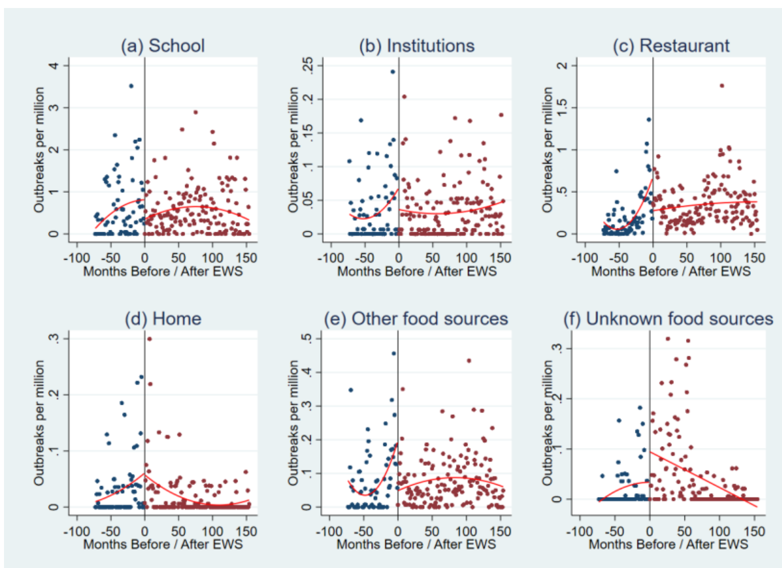
Table 1.11. DID estimates of the EWS on the number of cases of food-borne illness per million (with differential linear time trends)

	(1) Total	(2) Norovirus	(3) E. coli	(4) Salmonella	(5) Vibrio	(6) S. aureus	(7) C. jejuni	(8) Perfringens	(9) B. cereus	(10) Unknown
Panel A. Period 1 (Jan. 2002 – Oct. 2013)										
Post × School	-38.356*	-11.873*	-23.414**	-1.453	-1.183	-2.733	-3.827	3.260	1.761	0.327
	(18.414)	(5.768)	(8.667)	(0.978)	(1.071)	(5.927)	(3.503)	(4.153)	(0.996)	(4.817)
% Change from Mean	-81%	-117%	-161%	-92%	-109%	-64%	-281%	149%	201%	4%
Obs.	7,952	7,952	7,952	7,952	7,952	7,952	7,952	7,952	7,952	7,952
Panel B. Period 2 (Jan. 2002 – Aug. 2014)										
Post × School	-40.337*	-10.032*	-24.280**	-2.959	-1.178	-2.411	-3.117	2.166	1.754	-0.076
	(19.899)	(5.230)	(11.114)	(1.892)	(0.945)	(5.566)	(3.514)	(4.094)	(0.998)	(4.779)
% Change from Mean	-85%	-99%	-167%	-187%	-109%	-57%	-229%	99%	200%	-1%
Obs.	8,512	8,512	8,512	8,512	8,512	8,512	8,512	8,512	8,512	8,512
Panel C. Period 3 (Jan. 2002 – Dec. 2020)										
Post × School	-12.455	-1.656	-5.290	-1.270	-1.112*	-1.751	0.242	2.058	0.584	-1.442
	(8.530)	(4.157)	(4.313)	(0.902)	(0.602)	(3.554)	(2.248)	(2.571)	(0.985)	(3.096)
% Change from Mean	-26%	-16%	-36%	-80%	-103%	-41%	18%	94%	67%	-19%
Obs.	12,768	12,768	12,768	12,768	12,768	12,768	12,768	12,768	12,768	12,768
Pre-Period Outcome Mean	47.348	10.131	14.526	1.582	1.084	4.242	1.362	2.190	0.878	7.729

Notes: This table indicates OLS estimates. Columns indicate different outcome variables. Each panel reports DID estimates from Eq. (1.5). We do not have enough variation in Protozoa. Standard errors clustered at the region level are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.



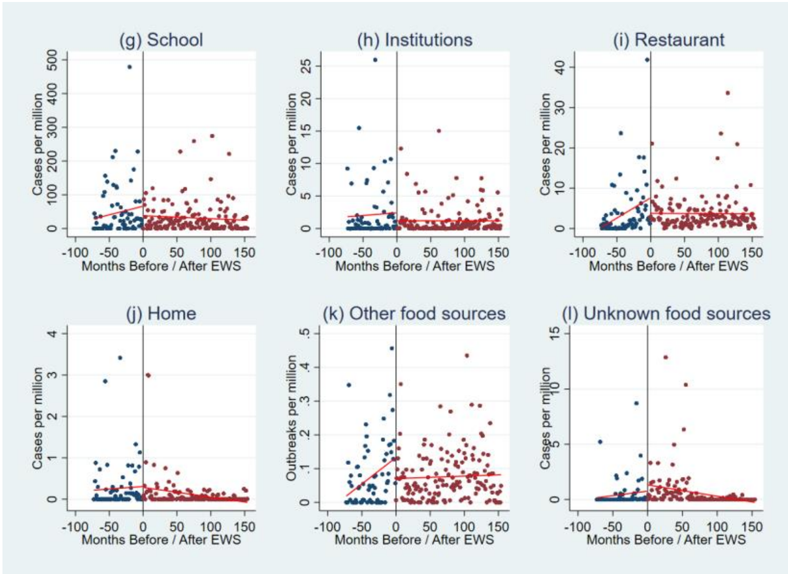
1. Linear fit



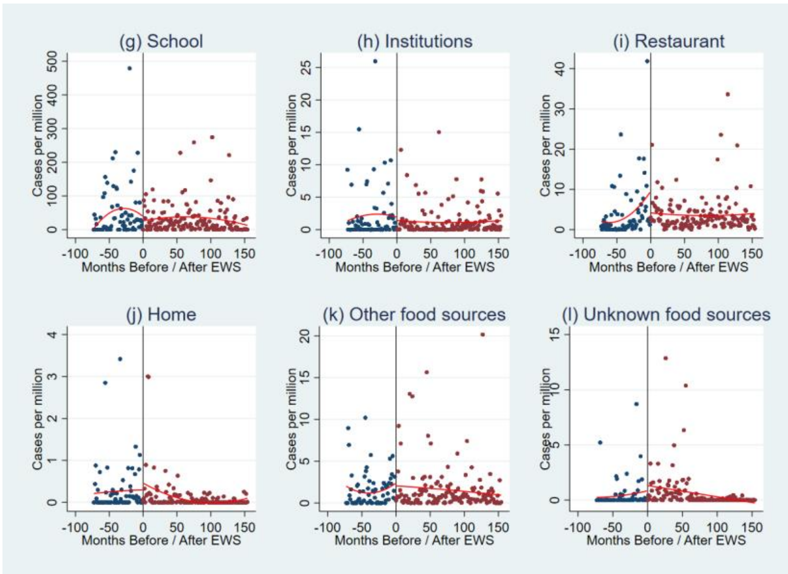
2. Quadratic fit

Figure 1.1. Discontinuity of the outbreaks of food-borne illness (full data period)

Notes: These figures illustrate trends in the average total outbreaks of food-borne illness per million. The horizontal axis is the month relative to the introduction of the EWS (i.e., Feb. 2008). The bandwidth is full data period (i.e., from Jan. 2002 to Dec. 2020). Month fixed effects are not controlled in Figure 1.1.



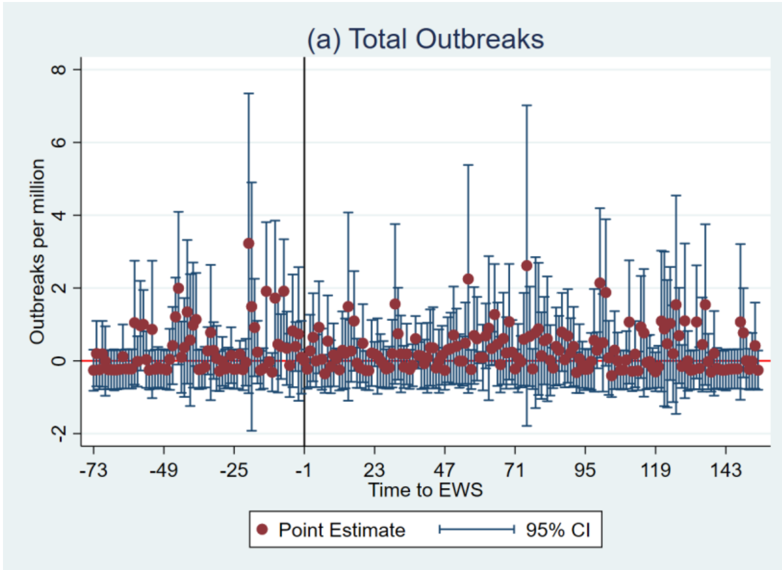
1. Linear fit



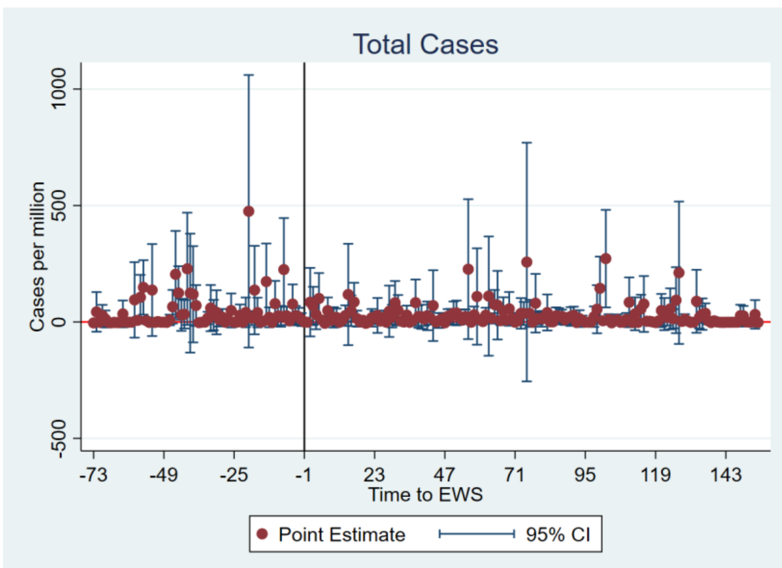
2. Quadratic fit

Figure 1.2. Discontinuity of the cases of food-borne illness (full data period)

Notes: These figures illustrate trends in the average total cases of food-borne illness per million. The horizontal axis is the month relative to the introduction of the EWS (i.e., Feb. 2008). The bandwidth is full data period (i.e., from Jan. 2002 to Dec. 2020). Month fixed effects are not controlled in Figure 1.2.



1. Outbreaks



2. Cases

Figure 1.3. Event Study Estimates

Note: Each figure reports coefficients estimated from Eq. (1.4). Standard errors are clustered at the region level.

CHAPTER 2

THE EFFECTS OF VOLUNTARY RESTAURANT SANITARY GRADES: EVIDENCE FROM SOUTH KOREA

2.1. INTRODUCTION

In 2021, the number of outbreaks and cases of food-borne illnesses was 245 and 5,160 in South Korea, respectively. According to the South Korean Ministry of Food and Drug Safety (MFDS), the annual social and economic costs of food-borne illness in South Korea are approximately 1.8 trillion Korean won (nearly 1.4 billion U.S. dollars).[22] Restaurants are the leading food source causing food-borne illness in South Korea. For example, in South Korea, 48.6% of outbreaks and 52.4% of cases of food-borne illness were from restaurants in 2021 (see Figure 2.1). However, few empirical studies investigated the impacts of the South Korean government's restaurant sanitary grades.

In December 2009, Seoul City introduced voluntary restaurant sanitary grades as a permanent policy to reduce the incidence of food-borne illness. Other regional governments introduced this program as a pilot test from May 2013 through December 2013[23]; 483 restaurants participated in the pilot test.[24] Each regional government started to expand this pilot program in 2015. In May 2017, the program became permanent in all regions in South Korea. However, in

---

[22] The MFDS calculated the costs by analyzing the data from 2016 to 2018. See press release https://www.mfds.go.kr/docviewer/skin/doc.html?fn=20220620090338671.hwpx&rs=/docviewer/result/ntc0021/46458/1/202302

[23] See http://repository.kihasa.re.kr/handle/201002/20396 for details.

[24] See https://www.mfds.go.kr/brd/m_209/view.do?seq=36562 for details.

August 2022, only 10,931 restaurants (1.2%) among 909,893 restaurants participated in the letter grading program in South Korea.[25]

A restaurant owner can apply to this program voluntarily without any cost. Two inspectors visit the restaurant to evaluate the sanitary condition if an application to the program is received. Before they visit, the inspectors have to check with the restaurant's owner to schedule the date of their visit. In addition, they have to notify the restaurant's owner of the information on the inspection, including the procedures, via phone call or text message. Each restaurant can post letter grades corresponding to sanitary inspection scores. An inspection score of 90 or more points is an "excellent," 85 to 90 points is a "very good," and 80 to 85 points is a "good." Customers can see the grades on food delivery apps, mobile map apps, the government's website or app, and restaurants' front doors or windows. The South Korean government encourages restaurant owners to participate in this program by providing various benefits. For example, the government exempts the restaurant sanitary inspection for two years for the restaurant that received a grade.[26] In addition, the South Korean government provides restaurant owners with a loan to renovate their restaurant facilities.

---

[25] The MFDS provided Youngin Ko, a member of the National Assembly, with data about the letter grading program. He released it via press release. See https://www.theminjoo.kr/board/view/inspection/1068951?page=68

[26] If a customer reports that a restaurant with a letter grade violated the sanitation law, government officials visit the restaurant to inspect the violation. In addition, if a food-borne illness occurs in a restaurant with a letter grade and it is reported, the inspectors visit the restaurant to investigate it. In South Korea, the MFDS oversees the sanitation of restaurants, and each regional administration has the right to implement the inspections. In practice, when the MFDS orders the regional governments to inspect restaurants, the regional governments' officials visit restaurants and inspect the sanitation. The inspection date or target restaurants is different depending on the issues. Therefore, the inspections outside of the voluntary program are closed to irregular inspections than regular inspections. For example, in July 2022, the MFDS announced that they would inspect restaurants in water parks, beaches, and camping sites with the regional governments between July 18 and 26, 2022. See press release https://www.mfds.go.kr/docviewer/skin/doc.html?fn=20220711090609534.hwpx&rs=/docviewer/result/ntc0021/46521/1/202303

Lee and Baek (2019) find that the Seoul City restaurant sanitary grades significantly decrease the number of food-borne illnesses. However, they do not handle the endogeneity problem. In addition, they use yearly panel data instead of monthly panel data. It does not capture the exact treatment timing. For example, Seoul City introduced the program in December 2009. However, they assume that the program was introduced in 2009. Moreover, they do not check the robustness of their results using pathogen-level data.

Empirical studies analyzing the effects of restaurant sanitary grades on food-borne illness have yielded mixed results. An infamous study by Jin and Leslie (2003) showed that the Los Angeles County restaurant sanitary grades significantly decrease the number of food-borne illness hospitalizations.[27] Jin and Leslie (2019) support the original conclusion of Jin and Leslie (2003). However, Ho et al. (2019) argue that the findings of Jin and Leslie (2003) do not hold up under the improved data and methods, as well as errors in the interpretation of their specification. In addition, Krinsky et al. (2022) find that the New York City restaurant sanitary grades do not significantly decrease the incidence of Salmonellosis.[28] Researchers have also studied food safety standards (Alberini et al., 2008; Minor & Parrett, 2017; Ollinger & Bovay, 2018; Ollinger & Bovay, 2020; Adalja et al., 2022), the PulseNet surveillance system (Scharff et al., 2016)[29], and the GenomeTrakr Whole Genome Sequencing (WGS) Network (Brown et al., 2021).[30]

---

[27] The program was mandatory in some cities within Los Angeles County and voluntary in other cities.

[28] The New York City restaurant sanitary grades are mandatory.

[29] PulseNet is a national laboratory network using DNA fingerprints of bacteria to predict potential outbreaks of food-borne illness (see https://www.cdc.gov/pulsenet/index.html for details).

[30] In the past, PulseNet used pulsed-field gel electrophoresis (PFGE) for DNA fingerprinting. However, PulseNet has used WGS after public health agencies adopted it from 2013–2019.

To fill in the research gaps, this study estimates the effects of the Seoul City restaurant sanitary grades on the number of outbreaks and cases of food-borne illness using the monthly administrative panel data in South Korea from January 2002 through April 2013. However, the estimates would be biased if the error term is correlated with both the voluntary decision to receive a grade (i.e., “treatment variable”) and the amount of food-borne illness (i.e., “dependent variable”). We use the difference-in-difference-in-differences (DDD) approach to address this potential endogeneity problem. We find that the Seoul City restaurant sanitary grades significantly reduce the number of food-borne illness outbreaks by 46%–66%. However, the results are not robust to different time windows. In addition, we could not find significant evidence that the Seoul City restaurant sanitary grades reduce the number of cases of food-borne illness.

The rest of the paper is organized as follows: Section 2.2 describes the data. Section 2.3 presents the empirical model. Section 2.4 presents the main results. Lastly, Section 2.5 discusses and concludes with some policy implications.

2.2. DATA

We requested data on the incidence of food-borne illness from the South Korean Ministry of Food and Drug Safety (MFDS).[31] This monthly panel data includes the number of reported food-borne outbreaks and cases by pathogen type (e.g., salmonella, E. Coli, etc.), food source (e.g., home versus school), and geographic region. It is well known that food-borne illness is underreported.[32] In South Korea, medical doctors and food service institutions (e.g., nursing

---

[31] We used a government website to request the MFDS data (see https://www.open.go.kr/com/main/mainView.do).

[32] The following example shows how the CDC estimates the annual number of food-borne illness episodes from the underreporting issue. See https://www.cdc.gov/foodborneburden/pdfs/scallan-estimated-illnesses-foodborne-

homes) are mandated to report the occurrence of food-borne illnesses to local government authorities. Therefore, like many countries, an illness is not counted if the person does not seek medical attention. However, the obligation to report is confined to when the institution confirms that two or more people became ill from the same source. We will discuss this underreporting issue in Section 2.3.

The MFDS data has information on 15 types of food-borne illnesses. The top 10 food borne illnesses are pathogenic E. coli, norovirus, unknown type, salmonella, staphylococcus aureus, vibrio parahaemolyticus, C. perfringens, campylobacter jejuni, bacillus cereus, and other bacteria in order of the number of cases.[33] This study analyzes the outbreaks and cases of the total types and each of the above top 10 types. The MFDS data also specify food sources depending on where people consumed contaminated foods or beverages. The six food sources are schools, institutions, restaurants, homes, other food sources, and unknown. In addition, the MFDS data includes information on the administrative regions.[34]

Finally, monthly balanced panel data (a type-source-region level) were constructed from January 2002 to December 2020. However, we analyze the data from only January 2002 through April 2013 because the control group (i.e., the regions outside of Seoul) has been polluted since May 2013, when they started the restaurant sanitary grades as a pilot test.

pathogens.pdf

[33] The rank is based on the data from January 2002 through April 2013. We will discuss why we choose the data period in detail below. The other five types are other viruses, natural toxins, chemical substances, clostridium botulinum, and protozoa.

[34] South Korea is divided into 17 administrative regions; Seoul, Gyeonggi-do, Incheon, Chungcheongbuk-do, Chungcheongnam-do, Daejeon, Sejong, Jeollabuk-do, Jeollanam-do, Gwangju, Gangwon-do, Gyeongsangbuk-do, Gyeongsangnam-do, Daegu, Ulsan, Busan, and Jeju-do.

We obtained control variables from the Korean Statistical Information Service.[35] The unemployment rate is the region-month-year level variable. However, male (%), under 20 years (%), 20-44 years (%), 45-64 (%), 65 years and over (%), and population are the region-year level variables.

Table 2.1 summarizes descriptive statistics. Table 2.1 shows that the average region-month-year cell reports 0.506 outbreaks and 14.679 cases of food borne illness per million. Looking at Table 2.1, restaurant is the second primary food source of food-borne illnesses.[36] The average number of reported outbreaks (cases) of food borne illness in the restaurant is 0.237 (3.700) per million people. Table 2.1 also shows that Norovirus and Pathogenic E. coli are the primary pathogens causing food-borne illnesses in South Korea. The mean of reported outbreaks (cases) of norovirus is 0.079 (3.133) per million people, and the mean of reported outbreaks (cases) of Pathogenic E. coli is 0.059 (3.538) per million people.

2.3. EMPIRICAL MODEL

We aim to estimate the impact of the Seoul City restaurant sanitary grades on the number of outbreaks and cases of food-borne illnesses. We use the difference-in-difference-in-differences (DDD) approach to address potential endogeneity problems, following Liu et al. (2021) and Geng et al. (2021). We consider three differences: before and after the introduction of the Seoul City restaurant sanitary grades (before and after December 2009); administrative

[35] Korean Statistical Information Service, see https://kosis.kr/index/index.do.

[36] In Table 2.1, we divided the number of reported outbreaks (cases) by the number of students to calculate the number of reported outbreaks (cases) per million in school. In contrast, we divided the number of reported outbreaks (cases) by population to calculate the number of reported outbreaks (cases) per million in a food source excluding schools. As discussed in Section 2.1, if we see the count of outbreaks (cases) instead of outbreaks (cases) per million, restaurants are the leading food source causing food-borne illness in South Korea.

region differences (Seoul vs. other regions);[37] and food source differences (restaurants vs. institutions, homes, and other food sources).[38] The specification is as follows:

$$Y_{srt} = \alpha_1 + \beta_1 Seoul_r \times Restaurant_s \times Post_t + \delta_1 Seoul_r \times Restaurant_s + \gamma_1 Seoul_r \times Post_t + \omega_1 Restaurant_s \times Post_t + X_{rt} + \varphi_r + \delta_s + \tau_t + \varepsilon_{srt} \quad (2.1)$$

where Y_{srt} is the number of reported outbreaks (cases) of food-borne illness per million people in food source s in region r in month-year level time t. $Seoul_r$ is a dummy variable indicating one if the region is Seoul. $Restaurant_s$ is a dummy variable indicating one if the food source is the restaurant. $Post_t$ is a dummy variable indicating one if the date is December 2009 or later. X_{rt} is a vector of covariates including 20–44 years (%), 45–64 years (%), 65 years and over (%), male (%), and unemployment rate (%). φ_r is the region fixed effects. δ_s is the food source fixed effects. τ_t is the time fixed effects. ε_{srt} is an error term. In equation (2.1), β_1 is the parameter of interest.

As we mentioned in Section 2.2, Y_{srt} in equation (2.1) is likely to be underreported. However, it still provides unbiased estimates and valid statistical inferences (Gujarati, 2015; Greene, 2018; Wooldridge, 2019).[39]

The DDD model requires the treatment and control groups to satisfy the common trend assumption, but it is not directly testable. The alternative researchers usually choose to test the

[37] We only use the data on 14 regions among 17 regions excluding Sejong, Chungcheongbuk-do, and Chungcheongnam-do because the South Korean government established a new administrative division, Sejong, by merging some parts of areas in Chungcheongbuk-do and Chungcheongnam-do in July 2012.

[38] School is not used as a control group because there was a policy change in schools in February 2008 (i.e., the introduction of the early warning system for food-borne outbreaks in schools).

[39] Bellemare and Nguyen (2018) are concerned about attenuation bias (i.e., the estimator is biased toward zero) because the dependent variable in their study is the number of food-borne illness outbreaks and cases, and the dependent variable is almost surely underreported. However, attenuation bias is unrelated to measurement errors in a dependent variable because it is only caused by measurement errors in an independent variable (Gujarati, 2015; Greene, 2018; Wooldridge, 2019). Therefore, our model does not cause attenuation bias.

common trend assumption is the event study design. Therefore, we use the event study approach (see details Clarke & Tapia-Schythe, 2021). The specification is as follows:

$$Y_{srt} = \alpha_2 + \sum_{l=-95}^{l=-2} \beta_l Grades_{l,rst} + \sum_{l=0}^{l=40} \beta_l Grades_{l,rst} + \delta_2 Seoul_r \times Restaurant_s + \gamma_2 Seoul_r \times Post_t \\ + \omega_2 Restaurant_s \times Post_t + X_{rt} + \varphi_r + \delta_s + \tau_t + \varepsilon_{srt} \tag{2.2}$$

where $Grades_{l,rst}$ are the event study dummies indicating one if the region is Seoul, the food source is the restaurant, and the relative time to the introduction of the Seoul City restaurant sanitary grades (i.e., $t - \text{December } 2009$) is l. In equation (2.2), the reference period is set as $l = -1$. We can jointly test if β_l is insignificantly different from zero in the pre-treatment period (i.e., from January 2002 to November 2009). However, testing if the coefficients of the time dummies in the pre-treatment period are all statistically insignificant does not guarantee that the trend is common to both groups in the post-treatment period in the absence of treatment (Kahn-Lang & Lang, 2020; Bilinski & Hatfield, 2020). In addition, high statistical power may detect practically insignificant violations of parallel trends (Bilinski & Hatfield, 2020).

Therefore, we compare the estimates from a DDD model with and without allowing group-specific time trends, following Kahn-Lang & Lang (2020) and Bilinski & Hatfield (2020). The specification with a group-specific trend is as follows:

$$Y_{srt} = \alpha_3 + \beta_3 Seoul_r \times Restaurant_s \times Post_t + \delta_3 Seoul_r \times Restaurant_s + \gamma_3 Seoul_r \times Post_t \\ + \omega_3 Restaurant_s \times Post_t + \mu_3 Seoul_r \times Restaurant_s \times t + X_{rt} + \varphi_r + \delta_s + \tau_t \\ + \varepsilon_{srt} \tag{2.3}$$

We compare β_1 in equation (2.1), with β_3 in equation (2.3). If parallel trends hold, the two estimates would be similar. It implies that if two estimates are not similar, the DDD model with allowing group-specific time trends would be more appropriate.

2.4. RESULTS

Before evaluating the policy impacts using the difference-in-difference-in-differences (DDD), we statistically tested parallel trends from equation (2.2). Table B.1.1 illustrates the joint significance tests for parallel trends in the pre-treatment period. It shows that parallel trends do not hold in the pre-treatment period for the number of outbreaks of food-borne illness of the total types, norovirus, pathogenic E. coli, salmonella, vibrio parahaemolyticus, S. aureus, C. jejuni, and unknown type. In addition, it does not hold for the number of cases of food-borne illness of the total types, norovirus, pathogenic E. coli, salmonella, vibrio parahaemolyticus, S. aureus, and unknown type. However, as discussed in Section 2.3, high statistical power may detect even the slightest violations of parallel trends. In addition, we have a long pre-treatment period which is 95 months. Thus, if trends differ in even a month among 94 months in the pre-treatment period, the common trends assumption does not hold.[40] Therefore, as we mentioned, we compare the estimates from a DDD model with and without allowing group-specific time trends.

Table 2.2 presents the DDD estimates with and without allowing group-specific time trends. Panel A presents the DDD estimates without allowing group-specific time trends. Panel B presents the DDD estimates with allowing group-specific time trends. The outcome variables are total outbreaks per million people in columns 1–4 and total cases per million people in columns 5–8, respectively. In columns 1–4, we find that the DDD estimates in Panel A are not similar to that of Panel B. Therefore, we prefer the DDD estimates in Panel B, allowing group-specific time trends. The DDD estimates in Panel B show that the Seoul City restaurant sanitary grades significantly reduce the number of food-borne illness outbreaks by 46%–66%. The results are

[40] In our event study design model, the reference is a month before the introduction of the Seoul City restaurant sanitary grades; thus, we jointly test if the 94 months, except for the reference in the pre-treatment period, are insignificantly different from zero.

robust to Poisson regression (see Tables B.1.2). However, they are not robust to different time windows (see Table B.1.3).[41] In columns 5–8 of Table 2.2, we could not find significant evidence that the Seoul City restaurant sanitary grades reduce the number of cases of food-borne illness.

The Seoul City restaurant sanitary grades may also have heterogeneous effects on the outbreaks and cases of food-borne illness among pathogens. Therefore, we run the estimation separately for pathogens. The results are shown in Tables 2.3 and 2.4, respectively. Table 2.3 shows that the Seoul City restaurant sanitary grades significantly reduce the outbreaks of S. aureus, C. jejuni, C. Perfringens, and unknown types by 132%, 318%, 304%, and 57%, respectively.[42] However, the results are not robust to Poisson regression (see Table B.1.4). In addition, they are sensitive to the different time windows (see Tables B.1.5 and B.1.6). Table 2.4 shows that the Seoul City restaurant sanitary grades significantly reduce the cases of C. jejuni by 320%.[43] However, the results are not robust to the different time windows (see Tables B.1.8 and B.1.9).

To summarize, we find some significant results that the Seoul City restaurant sanitary grades reduce the number of outbreaks and cases of food-borne illness. However, those results are not robust to diverse model specifications and the study windows.

[41] In our baseline model, the sample period is Jan 2002 to April 2013, 136 months. The long-term sample period may have remaining unobserved heterogeneity. Thus, we check the robustness by shorting the time windows following Geng et al. (2021).

[42] The low incidence of food-borne illness in the pre-term period drives those large estimates. For example, the policy impact on C. Perfringens is calculated as “DDD estimates / pre-period outcome mean of C. Perfringens × 100%.” The DDD estimates are approximately -0.015, and the pre-period outcome mean is approximately 0.005. Thus, the policy impact is about -304%.

[43] We do not have enough variation in C. jejuni from the Poisson regression (see Table B.1.7).

2.5. DISCUSSION AND CONCLUSIONS

South Korea’s annual social and economic costs of food-borne illness are approximately 1.8 trillion Korean won (nearly 1.4 billion U.S. dollars). Restaurants are the leading food source causing food-borne illness in South Korea. The Seoul City restaurant sanitary grades are one of the policies responding to the above issues. However, there are few studies about the restaurant letter grading system. To fill the research gap, we evaluate the effects of the Seoul City restaurant sanitary grades on the number of outbreaks and cases of food-borne illness using the monthly administrative panel data in South Korea for January 2002 – April 2013. We exploit the difference-in-difference-in-differences (DDD) approach to address endogeneity problems.

The DDD results show that the Seoul City restaurant sanitary grades significantly reduce the number of outbreaks of total types, S. aureus, C. jejuni, C. Perfringens, and unknown types. We also find that the restaurant sanitary grades significantly reduce the cases of C. jejuni. However, the results are not robust to Poisson regression or different time windows. Therefore, we could not find strong evidence that the Seoul City restaurant sanitary grades significantly reduce the number of outbreaks and cases of food-borne illness. The results are similar to Ho et al. (2019) and Krinsky et al. (2022). They could not find the significant effects of the restaurant sanitary grades in Los Angeles County and New York City.

Why the Seoul City restaurant sanitary grades do not significantly and robustly reduce the outbreaks and cases of food-borne illness is unclear. However, below two hypotheses may provide plausible explanations.

First, while this policy may encourage restaurants to participate in the restaurant sanitary grades, it may also have adverse consequences. Restaurants could neglect sanitary regulations after receiving a letter grade; thus, it may alleviate the program’s effectiveness in reducing the incidence of food-borne illnesses. For instance, according to the South Korean Ministry of Food and Drug Safety (MFDS), between January 2018 and August 2022, the number of restaurants with a letter grade was 10,931. During this period, 346 restaurants with a letter grade violated sanitary regulations, and foreign substances were found in 164 restaurants. Surprisingly, among the 164 restaurants, 98 restaurants (59.8%) had an 'Excellent' grade.[44] It implies that even an 'Excellent' grade does not guarantee food safety and may provide consumers with incorrect information. It also suggests that the South Korean government failed to provide the precise information on a restaurant's food safety necessary for the letter grading program to succeed. To ensure the program's success, it is important for relevant authorities to consistently monitor and enforce the regulatory framework to maintain the program's effectiveness in reducing the incidence of food-borne illnesses.

Second, the low rate of restaurants participating in the letter grading program seems to be one reason for undermining the policy impacts. For example, there were 909,893 restaurants in South Korea in August 2022. Only 10,931 restaurants (1.2%) participated in the letter grading program. Even New York City's mandatory restaurant sanitary grades do not significantly decrease the incidence of food-borne illness (Krinsky et al., 2022). Considering the results, it is difficult to expect that the South Korean voluntary restaurant sanitary grades will significantly impact reducing food borne illnesses. Thus, increasing restaurant participation in the letter grading program is necessary to promote food safety in South Korea.

[44] See press release https://www.theminjoo.kr/board/view/inspection/1068951?page=68

Table 2.1. Descriptive statistics

Variable	Mean	Std. Dev.	Min	Max
Reported outbreaks per million				
Total	0.506	0.855	0	10.705
Food source				
Schools	0.544	1.770	0	24.599
Institutions	0.032	0.150	0	2.658
Restaurants	0.237	0.548	0	7.252
Homes	0.031	0.181	0	3.568
Other food sources	0.069	0.230	0	2.145
Unknown food sources	0.052	0.222	0	3.547
Type				
Norovirus	0.079	0.289	0	3.568
Pathogenic E. coli	0.059	0.244	0	3.313
Salmonella	0.049	0.217	0	3.626
Vibrio parahaemolyticus	0.041	0.205	0	3.626
Staphylococcus aureus	0.035	0.155	0	1.784
Campylobacter jejuni	0.010	0.076	0	1.078
C. perfringens	0.012	0.101	0	1.806
Bacillus cereus	0.012	0.090	0	1.389
Other bacteria	0.005	0.053	0	1.045
Unknown type	0.185	0.420	0	5.337
Reported cases per million				
Total	14.679	38.016	0	638.1
Food source				
Schools	40.961	170.515	0	3508.678
Institutions	1.734	12.414	0	275.206
Restaurants	3.700	16.760	0	377.543
Homes	0.240	1.894	0	39.887
Other food sources	1.735	10.482	0	218.769
Unknown food sources	0.819	6.716	0	175.592

Type				
Norovirus	3.133	17.681	0	274.453
Pathogenic E. coli	3.538	22.363	0	377.543
Salmonella	1.466	9.391	0	218.769
Vibrio parahaemolyticus	0.771	5.505	0	116.549
Staphylococcus aureus	1.230	8.509	0	182.910
Campylobacter jejuni	0.548	6.655	0	172.881
C. perfringens	0.563	6.593	0	165.874
Bacillus cereus	0.237	2.845	0	83.924
Other bacteria	0.257	4.577	0	163.621
Unknown type	2.611	9.388	0	140.402
Male (%)	50.132	0.442	49.417	51.497
Under 20 years (%)	25.096	2.636	18.819	31.504
20–44 years (%)	40.316	3.671	31.064	47.282
45–64 years (%)	24.481	3.144	17.433	32.047
65 years and over (%)	10.108	3.359	4.341	19.377
Elementary school students (%)	48.806	3.901	40.766	57.258
Middle school students (%)	26.215	1.516	22.131	28.494
High school students (%)	24.979	2.662	19.699	31.032
Unemployment rate (%)	3.176	1.073	0.800	6.400
Population (millions of individuals)	3.273	3.108	0.549	12.144
Students (millions of individuals)	0.497	0.461	0.086	1.853

Notes: This table indicates descriptive statistics for 14 regions in South Korea for the period Jan. 2002 – Apr. 2013 (N = 2,072).

Table 2.2. The effects of the Seoul City restaurant sanitary grades

Dependent Variables	Total outbreaks per million people				Total cases per million people			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: DDD estimates without allowing group-specific time trends								
Seoul × Rest. × Post	0.018	0.018	0.018	0.018	1.335	1.335	1.335	1.335
	(0.033)	(0.036)	(0.033)	(0.035)	(0.915)	(0.951)	(0.924)	(0.975)
Pre-Period Outcome								
Mean	0.237	0.237	0.237	0.237	3.700	3.700	3.700	3.700
% Change from Mean	8%	8%	8%	8%	36%	36%	36%	36%
Obs.	7,616	7,616	7,616	7,616	7,616	7,616	7,616	7,616
Panel B: DDD estimates with allowing group-specific time trends								
Seoul × Rest. × Post	-0.156***	-0.120***	-0.056	-0.110**	0.227	0.734	1.157	0.139
	(0.033)	(0.042)	(0.045)	(0.046)	(0.915)	(1.056)	(1.139)	(1.373)
Pre-Period Outcome								
Mean	0.237	0.237	0.237	0.237	3.700	3.700	3.700	3.700
% Change from Mean	-66%	-51%	-24%	-46%	6%	20%	31%	4%
Obs.	7,616	7,616	7,616	7,616	7,616	7,616	7,616	7,616
Seoul	Y	Y	N	N	Y	Y	N	N
Restaurant	Y	Y	N	N	Y	Y	N	N
Post	Y	Y	N	N	Y	Y	N	N
Covariates	N	Y	N	Y	N	Y	N	Y
Region FE	N	N	Y	Y	N	N	Y	Y
Source FE	N	N	Y	Y	N	N	Y	Y
Time FE	N	N	Y	Y	N	N	Y	Y

Notes: This table indicates OLS estimates. Column (4) of Panel A reports DDD estimates from Eq. (2.1). Column (8) of Panel B reports DDD estimates from Eq. (2.3). Seoul × Restaurant dummy, Seoul × Post dummy, Restaurant × Post dummy, and constant are included in all models. Seoul × Restaurant × Time is included in all models in Panel B. The "Rest." in this and all subsequent tables refers to "Restaurant." Standard errors in parentheses are clustered at the region-source level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 2.3. Heterogeneity analysis (Dependent variable: outbreaks per million people)

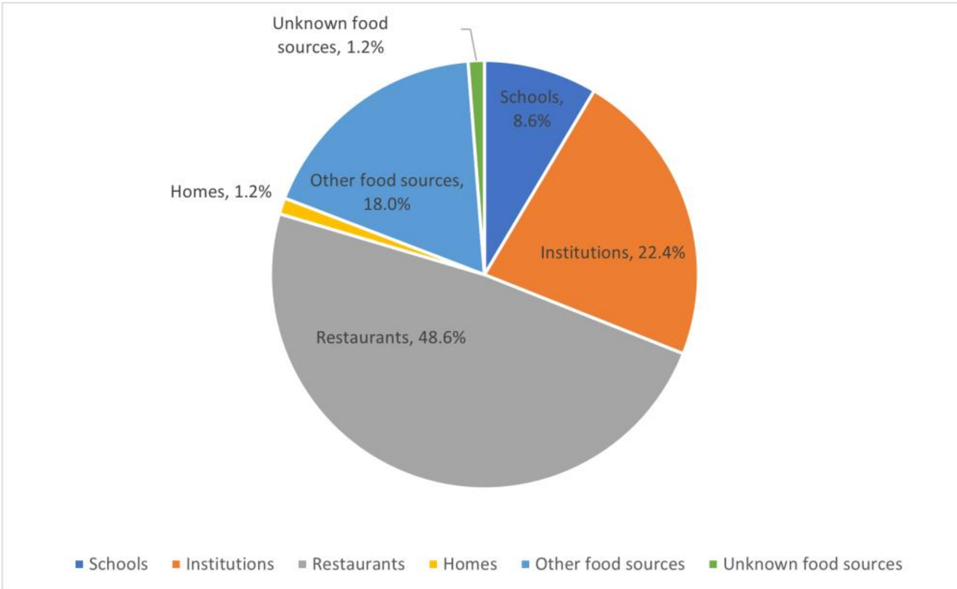
Dependent Variables	Outbreaks per million people									
	Norovirus	E. coli	Salmonella	Vibrio	S. aureus	C. jejuni	C. Perfringens	B. cereus	Other bacteria	Unknown
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Panel A: DDD estimates without allowing group-specific time trends										
Seoul × Rest. × Post	0.018	-0.003	-0.013*	0.002	-0.009	-0.004	-0.003	-0.009	-0.000	0.037*
	(0.011)	(0.008)	(0.008)	(0.007)	(0.006)	(0.002)	(0.004)	(0.006)	(0.001)	(0.018)
Pre-Period Outcome Mean	0.028	0.019	0.025	0.026	0.016	0.003	0.005	0.007	0.000	0.104
% Change from Mean	63%	-16%	-51%	8%	-57%	-159%	-61%	-136%	0%	35%
Obs.	7,616	7,616	7,616	7,616	7,616	7,616	7,616	7,616	7,616	7,616
Panel B: DDD estimates with allowing group-specific time trends										
Seoul × Rest. × Post	0.008	-0.005	-0.006	0.014	-0.021***	-0.008***	-0.015*	-0.012	-0.001	-0.059**
	(0.013)	(0.010)	(0.010)	(0.010)	(0.006)	(0.003)	(0.008)	(0.007)	(0.001)	(0.024)
Pre-Period Outcome Mean	0.028	0.019	0.025	0.026	0.016	0.003	0.005	0.007	0.000	0.104
% Change from Mean	28%	-27%	-24%	53%	-132%	-318%	-304%	-181%	-1854%	-57%
Obs.	7,616	7,616	7,616	7,616	7,616	7,616	7,616	7,616	7,616	7,616

Notes: The estimates in each panel report coefficients from OLS regression. Panel A reports DDD estimates from Eq. (2.1). Panel B reports DDD estimates from Eq. (2.3). Standard errors in parentheses are clustered at the region-source level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

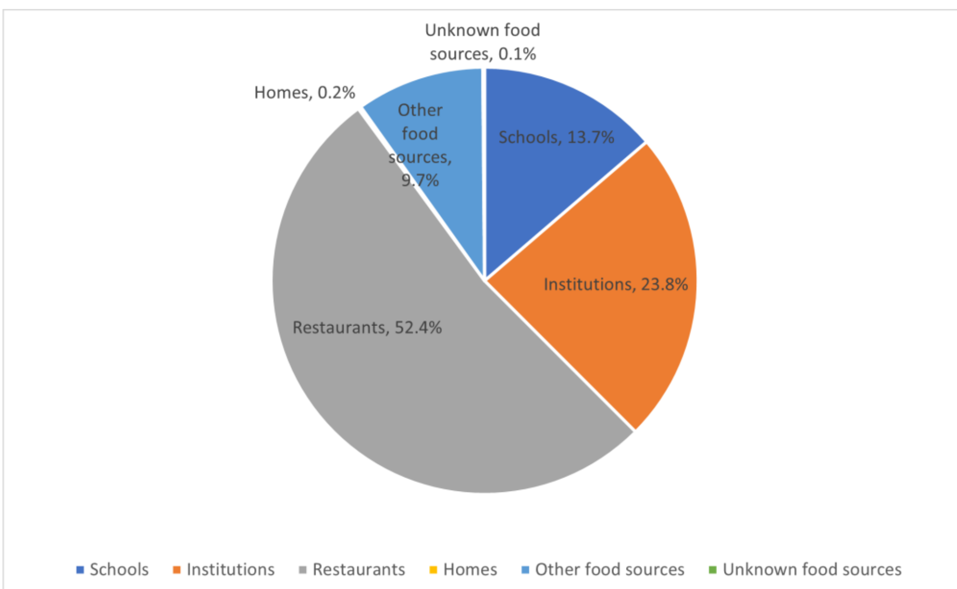
Table 2.4. Heterogeneity analysis (Dependent variable: cases per million people)

Dependent Variables	Cases per million people									
	Norovirus	E. coli	Salmonella	Vibrio	S. aureus	C. jejuni	C. Perfringens	B. cereus	Other bacteria	Unknown
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Panel A: DDD estimates without allowing group-specific time trends										
Seoul × Rest. × Post	1.040*	0.093	0.024	0.177	-0.086	-0.078	-0.012	-0.023	-0.037	0.195
	(0.596)	(0.505)	(0.302)	(0.175)	(0.118)	(0.061)	(0.128)	(0.038)	(0.036)	(0.169)
Pre-Period Outcome Mean	0.594	0.658	0.719	0.470	0.233	0.054	0.112	0.038	0.009	0.795
% Change from Mean	175%	14%	3%	38%	-37%	-145%	-11%	-61%	-408%	25%
Obs.	7,616	7,616	7,616	7,616	7,616	7,616	7,616	7,616	7,616	7,616
Panel B: DDD estimates with allowing group-specific time trends										
Seoul × Rest. × Post	0.977	-0.001	-0.283	0.260	0.038	-0.173*	-0.131	-0.044	-0.113	-0.380
	(0.865)	(0.850)	(0.429)	(0.183)	(0.135)	(0.099)	(0.170)	(0.055)	(0.093)	(0.363)
Pre-Period Outcome Mean	0.594	0.658	0.719	0.470	0.233	0.054	0.112	0.038	0.009	0.795
% Change from Mean	164%	0%	-39%	55%	16%	-320%	-117%	-116%	-1247%	-48%
Obs.	7,616	7,616	7,616	7,616	7,616	7,616	7,616	7,616	7,616	7,616

Notes: The estimates in each panel report coefficients from OLS regression. Panel A reports DDD estimates from Eq. (2.1). Panel B reports DDD estimates from Eq. (2.3). Standard errors in parentheses are clustered at the region-source level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.



1. Outbreaks



2. Cases

Figure 2.1. The food-borne illnesses in South Korea in 2021

Notes: These figures show the relative proportion of food sources in the outbreaks and cases of food-borne illnesses in South Korea in 2021. The source is the author's calculation based on the South Korean Ministry of Food and Drug Safety data (MFDS).

CHAPTER 3

DOES A FOOD ASSISTANCE PROGRAM FOR PREGNANT WOMEN IMPROVE BIRTH OUTCOMES? EVIDENCE FROM NUTRIPLUS IN SOUTH KOREA

3.1. INTRODUCTION

NutriPlus is a special supplemental nutrition program for South Korean women, infants, and children. Its goal is to improve the nutritional well-being of low-income pregnant and postpartum women, infants, and children under six. It is based on the Women, Infants, and Children (WIC) program in the United States. NutriPlus was first implemented as a pilot program from 2005 to 2007 in 20 public health centers. In 2008, NutiPlus became permanent and was rolled out across the country. By 2011, nearly all 256 public health centers in South Korea had implemented NutiPlus.

In 2020, the South Korean government appropriated $6.2 million to fund NutriPlus, and the program served 73,333 people, of which 10.4% are pregnant women, 23.1% are postpartum or breastfeeding women, 35.8% are infants (up to age 1), and 30.7% are children (age 1-6).[45] In South Korea, only about 2.9% of pregnant women who gave birth in 2020 participated in NutriPlus (about 39.6% of women who gave birth in the United States in 2016 participated in WIC).[46] In 2016, only about 10% of women, infants, and children with NutriPlus eligibility

[45] The fraction of NutriPlus participants who lived in Seoul in 2020 was 18.6%, and NutriPlus spending in Seoul in 2020 was \$1,161,922. Using this information, we roughly calculated NutriPlus spending in South Korea in 2020 as $\$1{,}161{,}922 \times \frac{100}{18.6} = \$6{,}246{,}892$. NutriPlus information is available on the Korea Health Promotion Institute website; see https://www.khealth.or.kr/kps.

[46] The 2.9% was calculated by authors using the data from the Korea Health Promotion Institute and the birth

participated in this program because of the limited budget. The number of NutriPlus participants has decreased since 2015 (see Figure 3.1).

NutriPlus provides tailored food packages, including rice, eggs, and infant formula, once or twice per month via a food delivery service (note: WIC provides electronic vouchers/EBT cards for grocery stores). Similar to WIC, NutriPlus offers nutrition education at least once per month, as well as nutritional check-ups during the initial certification period, recertification period, and on the last day of participation.[47]

NutriPlus is a means-tested program. Eligibility rules require participants to live in households with family incomes below 80% of the standard median income and to be at nutritional risk.[48] However, pregnant women are eligible regardless of nutritional risk.

In South Korea, the average birth weight has declined from 3,265g in 1998 to 3,170g in 2020 (see Figure 3.2). Correspondingly, during this period, the fraction of low birth weight (less than 2,500g) increased from 3.5% to 6.8%, and the fraction of very low birth weight (less than 1,500g) increased from 0.2% to 0.8%. Premature birth also increased from 3.5% in 1998 to 8.6% in 2020.[49] Poor birth outcomes are associated with future medical costs, health problems, and

---

certificate data from the "microdata integrated service" provided by Statistics Korea. WIC information is available on the CDC website; see https://www.cdc.gov/nchs/products/databriefs/db298.htm.

[47] In 2019, 94.5 % of nutrition education was conducted in-person learning (see the report of the Korea Health Promotion Institute, https://www.khepi.or.kr/kps/publish/view?menuId=MENU00891&page_no=B2017004&pageNum=9&siteId=&srch_text=&srch_cate=&srch_type=&str_clft_cd_list=&str_clft_cd_type_list=&board_idx=10442). However, in 2020, only 49.7% of nutrition education was conducted in person because of the Coronavirus Pandemic (see the report of the Korea Health Promotion Institute, https://www.khepi.or.kr/kps/publish/view?menuId=MENU00891&page_no=B2017004&pageNum=6&siteId=&srch_text=&srch_cate=&srch_type=&str_clft_cd_list=&str_clft_cd_type_list=&board_idx=10723).

[48] Nutritional risk is determined by a health professional at a public health center. Re-certification period is every six months.

[49] Premature birth indicates a birth that occurs before the 37th week of pregnancy. The authors calculated the average birth weight, the fraction of low birth weight, very low birth weight, and premature birth using the birth certificate data from the "microdata integrated service" provided by Statistics Korea. Using the same data, we calculated the trends of maternal age in South Korea. The fraction of women whose maternal age is 35 or above increased from 6.2% in 1998 to 33.9% in 2020. Correspondingly, during this period, the fraction of women whose

behavioral deficits (Almond et al., 2005; Datta Gupta et al., 2013; Hummer et al., 2014). Further, adverse birth outcomes may be associated with lower educational attainment and wages (Behrman & Rosenzweig, 2004; Black et al., 2007; Fletcher, 2011; Pehkonen et al., 2021).

Thus, improving birth outcomes is an important policy issue in South Korea. As such, the food packages and nutritional services for pregnant women are considered a major tool to (potentially) improve birth outcomes in South Korea. However, there are limited empirical studies concerning the effects of NutriPlus. Moreover, to our knowledge, no study has investigated the relationship between NutriPlus and birth outcomes (note: in the United States, a substantial body of literature has examined the effects of WIC on birth outcomes).[50] Thus, this study aims to answer the question: Does NutriPlus spending improve birth outcomes?

Estimating the impact of government spending on NutriPlus is difficult because of the endogeneity problem. The estimates would be biased if the error term is correlated with both the government spending on NutriPlus (i.e., “treatment variable”) and the birth outcomes (i.e., “dependent variable”). To address this problem, we use an instrumental variable approach. We use the average NutriPlus spending in neighboring counties (i.e., geographically adjacent counties) as an instrumental variable for the NutriPlus spending in a domestic county.

In our main analysis, we do not find statistically significant evidence that NutriPlus spending improves the incidence of low birth weight, very low birth weight, normal birth weight, and premature birth. However, we find that increasing NutriPlus spending by 100% reduces the

maternal age is 30 or above increased from 29.8% to 77.5%. According to Koo et al. (2012), maternal age 35 or above is associated with the incidence of low birth weight, and maternal age 30 or above is associated with preterm birth. Thus, the increase in maternal age may explain the aggravation of birth outcomes in South Korea during the period.

[50] Chorniy et al. (2020) argue that the most credible studies report WIC improves average birth weight and the likelihood of low birth weight while Sonchak (2016) argues that evidence of WIC’s effectiveness is mixed, with substantial and minimal health improvements.

average birth weight by 4.369g (0.134%) and the fraction of high birth weight by approximately 0.021%. We could not find a clear explanation for why NutriPlus reduces the average birth weight in our sample. However, the decreased likelihood of high birth weight may be associated with the reduced average birth weight. Additionally, using linear interpolation, we find that NutriPlus significantly reduces the likelihood of very low birth weight by 0.031%. Overall, our main results are robust to including the linear interpolation except for the likelihood of very low birth weight.

3.2. DATA

3.2.1. NutriPlus spending data

The policy variable of interest is real per capita county government spending on NutriPlus in 25 counties in Seoul, where one-fifth of the South Korean population lives.[51] Since the data is unavailable from a unified source, we compile information from closing financial statements posted on each county's website. We used closing financial statements for 19 counties over 25 counties because the other six counties do not have separate NutriPlus expenditure account codes or did not publicize the closing financial statements on their websites. Each county's expenditures on NutriPlus were coded as zero before each county implemented the program. Some counties over the 19 counties have missing values in NutriPlus expenditures in specific years. Therefore, Nutriplus spending data is county-level unbalanced panel data. For example, Gangnam-gu introduced NutriPlus in 2008, and we imported information on the county's NutriPlus expenditures from 2013 to 2020. Thus, there are missing values from 2009 to

[51] In 2021, 9.5 million out of 51.6 million people live in Seoul. There are 25 administrative districts called gu in Seoul (e.g., Gangnam-gu). From now on, we will refer to gu as the county.

2012. More details are included in Table C.1.1. To fill in the missing years, we also use linear interpolation. However, the linear interpolation may not fit the true NutriPlus spending trends because some counties may have ramped up spending at an increasing rate, while others at a decreasing rate. Therefore, we use linear interpolation not for our main analysis but for robustness checks.

This study uses the average NutriPlus spending in neighboring counties as an instrumental variable for the NutriPlus spending in a domestic county. For example, in the case of Gangnam-gu, the value is computed from two neighboring counties, Seocho-gu and Songpa-gu. More details are included in Table C.1.2.

3.2.2. Birth certificates data

In this study, we collected individual-level birth outcomes (e.g., birth weight, sex, and multiple births) and maternal characteristics (e.g., age, educational attainment, and gestational age) in Seoul from 1997 through 2020 from the "microdata integrated service" provided by Statistics Korea.[52] The data is coded from birth certificates and available beginning in 1997. The data represent a 100% sample of births and include approximately 2.3 million observations. The problem is that multiple births increase the likelihood of low birth weight (Cho & Lee, 2021). However, the 1997 data does not include information on multiple births. Therefore, we exclude the 1997 data in our analyses. We also eliminate observations with multiple births and missing values. Therefore, the final microdata includes approximately 2.0 million singleton births observations from 1998 through 2020. However, this microdata does not include the information

[52] Microdata integrated service website, see https://mdis.kostat.go.kr/index.do.

on who was on NutriPlus. Thus, we transform this microdata into county-level balanced panel data, then combine it with NutriPlus spending data.

3.2.3. The final data

We use Seoul's real gross regional domestic product (GRDP) per capita as a control variable. We collected the data from the "Korean Statistical Information Service" provided by Statistics Korea.[53] However, Statistics Korea does not provide county-level real GRDP per capita data. Thus, we use this data to control income variation in Seoul between years. We combine this real GRDP per capita data with the NutriPlus spending and birth certificate data. Therefore, the analysis uses yearly county-level unbalanced panel data compiled across 19 counties from 1998–2020. Table 3.1 presents descriptive statistics.

3.3. EMPIRICAL MODEL

We estimate the impact of government NutriPlus spending on birth outcomes. Specifically, we estimate the following baseline model:

$$Y_{it} = \alpha_1 + \beta_1 S_{it} + \delta_1 X_{it} + \tau_i + \eta_i \times year_t + \epsilon_{it} \qquad (3.1)$$

In equation (3.1), the endogenous variable S_{it} is government NutriPlus spending per capita in a county i in year t. The dependent variable, Y_{it} is birth outcomes in a county i in year t. The main outcomes of interest are the average birth weight and the normal, low, very low, and high birth weight fractions. The vector X_{it} contains the maternal characteristics (e.g., age, educational attainment, and gestational age), newborns' sex, and real GRDP per capita. In all specifications, we include county fixed effects τ_i. We examine the sensitivity to including county-specific linear time trends η_i or time fixed effects, which are not shown in equation (3.1).

[53] Korean Statistical Information Service, see https://kosis.kr/index/index.do.

The term ϵ_{it} is an error term.

Although we use the baseline model, the endogenous variable may be correlated with the error term because of measurement errors or omitted variable bias. To overcome this endogeneity problem, we use an instrumental variables approach. The instrumental variable must be correlated with the endogenous variable (i.e., the relevance condition) and uncorrelated with the error term (i.e., the exogeneity condition). In other words, the instrument variable must affect the outcome variable only through the endogenous variable (i.e., exclusion restriction).

We use the average NutriPlus spending in neighboring counties as an instrumental variable for the NutriPlus spending in a domestic county. This strategy assumes that the fiscal design has spill-over effects across counties. In other words, government expenditures in a county mirror those in neighboring counties. This is similar to the strategy used by Collier and Hoeffler (2004), Bodea et al. (2016), and Justino and Martorano (2018). Collier and Hoeffler (2004) and Bodea et al. (2015) use the level of government military spending in neighboring countries to instrument the level of government military spending in a given country. Justino and Martorano (2018) use the level of government welfare spending in neighboring countries to instrument the level of government welfare spending in a given country. In Section 3.4, we show that the instrumental variable is strongly correlated to the endogenous variable. Further, we assume that the average NutriPlus spending in neighboring counties affects birth outcomes only through the domestic county's fiscal design. It is plausible because women can only participate in NutriPlus at a public health center in the county where they live. Specifically, we estimate the following model. First stage:

$$\mathrm{S}_{it} = \alpha_2 + \beta_2 \mathrm{N}_{it} + \delta_2 X_{it} + \tau_i + \eta_i \times year_t + \epsilon_{it} \tag{3.2}$$

In equation (3.2), N_{it} indicates the average NutriPlus spending per capita in county i's neighboring counties in year t. Second stage:

$$Y_{it} = \alpha_3 + \beta_3 \hat{S}_{it} + \delta_3 X_{it} + \tau_i + \eta_i \times year_t + \epsilon_{it} \quad (3.3)$$

In equation (3.3), $\hat{S}_{it}$ indicates the predicted value of the endogenous variable. We cluster the standard errors at the county level to account for any variation within counties. In the same way, we estimate the impact of NutriPlus spending on premature birth. In this case, the vector X_{it} in equations (3.1) and (3.3) do not contain gestational age.

3.4. RESULTS

Table 3.2 shows the first-stage relationship between the endogenous variable (i.e., the average NutriPlus spending per capita in each domestic county) and the instrumental variable (i.e., the average NutriPlus spending per capita in the neighboring counties) is significantly positive in all specifications. We perform statistical tests using the first-stage F-statistic. The null hypothesis is that the instruments are weak. It is not a weak instrument if it is greater than 10 (Gujarati, 2015). Table 3.2 shows that F-statistics associated with the first stage are above 10 in regressions (1) and (4). This relationship is robust to include county-specific linear time trends (see regressions (3) and (6)). However, the relationship is not robust to include year fixed effects (see regressions (2) and (5)). If an instrument is weak, the instrumental variable estimator can be severely biased. This issue motivates our choice of including county-specific linear time trends instead of year fixed effects as the preferred specification.

Table 3.3 reports the OLS estimates for birth outcomes in regressions (1), (2), and (3) and the instrumental variable estimates in regressions (4), (5), and (6), respectively. Regressions (1) and (4) do not include year fixed effects or county-specific linear time trends. Regressions (2) and (5) control year fixed effects, and regressions (3) and (6) control county-specific linear time trends. Results are presented separately for each birth outcome: (i) average birth weight, (ii) the fraction of low birth weight, (iii) the fraction of very low birth weight, (iv) the fraction of high birth weight, (v) the fraction of normal birth weight, and (vi) the fraction of premature birth.

The six columns in panel (i) report the impact of NutriPlus spending per capita on the average birth weight. The coefficients are negative in all specifications, and they are statistically significant in regressions (1), (3), (4), and (6). According to regression (6), our preferred specification, increasing NutriPlus spending by 1,000 KRW per capita, reduces the average birth weight by around 13.1 grams. We also want to know the impact of increasing NutriPlus spending by 1% (labeled ‘‘1% spending’’).[54] The results indicate that increasing NutriPlus spending by 1% reduces the average birth weight by 0.044 grams. The estimate expressed as a percentage of mean birth weight (labeled ‘‘1% Impact’’) is -0.001%.[55] It implies that increasing NutriPlus spending by 1% reduces the average birth weight by 0.001%. Additionally, increasing NutriPlus spending by 100% (labeled ‘‘100% spending’’) reduces the average birth weight by 4.369 grams. The estimate expressed as a percentage of mean birth weight (labeled ‘‘100% Impact’’) is 0.134. It means that increasing NutriPlus spending by 100% reduces the average birth weight by 0.134%.

Panel (iv) reports that the coefficients are negative in all specifications, and they are statistically significant in regressions (1), (3), (4), and (6). Regression (6), our preferred specification, shows that increasing NutriPlus spending by 1,000 KRW per capita reduces the likelihood of high birth weight by around 0.064%. The results indicate that increasing NutriPlus

---

[54] This is calculated as coefficient × average NutriPlus spending × 0.01.

[55] This is calculated by dividing the “1% spending” by mean birth weight.

spending by 1% (100%) reduces the likelihood of high birth weight by approximately 0.000% (0.021%).

Panels (ii), (iii), (v), and (vi) show that NutriPlus spending does not significantly affect the likelihood of low birth weight, very low birth weight, normal birth weight, and premature birth in all specifications.

We also conduct robustness tests using linear interpolation to address the missing years in NutriPlus spending data. Table 3.4 shows that when we use the linear interpolation, NutriPlus spending per capita in each domestic county and average NutriPlus spending per capita in neighboring counties have a significantly and strongly positive relationship in all specifications. Moreover, the F-statistics associated with the first stage are above 10 in all specifications. In Table 3.5, our preferred specification, regression (6), shows that NutriPlus spending significantly reduces the average birth weight, the fraction of very low birth weight, and the fraction of high birth weight. It implies that our main results in Table 3.3 are robust to linear interpolation.

3.5. DISCUSSION AND CONCLUSIONS

NutriPlus is a representative food assistance program for women, infants, and children in South Korea. As such, the food packages and nutritional services for pregnant women are considered a major tool to (potentially) improve birth outcomes. However, no study has investigated the relationship between NutriPlus and birth outcomes. Therefore, this study analyzes the impact of NutriPlus spending on birth outcomes using yearly county-level unbalanced panel data compiled across 19 counties from 1998–2020.

We use the average NutriPlus spending per capita in neighboring counties as an instrumental variable for the NutriPlus spending per capita in a domestic county. In our main analysis, we do not find statistically significant evidence that NutriPlus spending improves the fraction of low birth weight, very low birth weight, normal birth weight, and premature birth. However, we find that NutriPlus spending significantly decreases the average birth weight and the fraction of high birth weight. The results are robust to linear interpolation.

We could not find a clear explanation for why NutriPlus reduces the average birth weight. However, it seems that a reduction in high birth weights drives the effect on the average. It is supported by the fact that NutriPlus spending does not significantly affect the fraction of low birth weight, very low birth weight, and normal birth weight.

Poor birth outcomes are associated with future medical costs, health problems, lower educational attainment, and wages (Almond et al., 2005; Behrman & Rosenzweig, 2004; Black et al., 2007; Datta Gupta et al., 2013; Fletcher, 2011; Hummer et al., 2014; Pehkonen et al., 2021). Especially high birth weight is associated with the likelihood of type 2 diabetes and obesity (Johnsson et al., 2015). Therefore, the results imply that NutriPlus may reduce potential economic and health problems by improving birth outcomes. However, the effects we found are relatively small. Moreover, there is little clinical evidence that supplemental nutrition for pregnant women improves birth outcomes in developed countries (Joyce et al., 2005). However, we do not argue that NutriPlus is unnecessary for pregnant women because even the effectiveness of the WIC program that has been studied over 30 years is not consistent and clear. In addition, the effects of NutriPlus on birth outcomes could be averaged out because only a small fraction of the population is even on NutriPlus. For example, approximately 2.9% of pregnant women who gave birth in 2020 participated in NutriPlus, and only 10% of women, infants, and children with NutriPlus eligibility participated in this program because of the limited budget. Thus, the results indicate the lower bound of the effects of NutriPlus. Therefore, more evidence-based studies about the impact of NutriPlus would contribute to discussing the expansion of NutriPlus eligibility and enrollment.

Table 3.1. Descriptive Statistics

Variable	Mean	Std. Dev.	Min	Max
NutriPlus spending per capita (1,000 KRW)	0.245	0.374	0.000	2.693
NutriPlus spending per capita in NBR (1,000 KRW)	0.284	0.397	0.000	2.006
Real GRDP per capita (1,000 KRW)	26948.541	9330.273	13402.000	45859.000
Mother's age (%)				
age $<$ 20	0.328	0.193	0.000	1.487
20 $\leq$ age $<$ 35	82.867	9.329	54.751	94.536
age $\geq$ 35	16.805	9.371	5.045	45.199
College (%)	66.952	17.891	24.261	96.924
Male newborn (%)	51.598	0.916	47.361	54.406
Gestation				
Gestational age (weeks)	39.035	0.253	38.426	39.535
Premature birth (%)	3.970	0.648	2.444	6.216
Birth weight (in grams)	3259.459	24.831	3188.528	3317.362
Birth weight (%)				
Very low birth weight (birth weight $<$ 1500g)	0.344	0.136	0.000	0.933
Low birth weight (birth weight $<$ 2500g)	3.316	0.511	2.005	5.146
Normal birth weight (2500g $\leq$ birth weight $\leq$ 4500g)	96.421	0.480	94.490	97.719
High birth weight (birth weight $>$ 4500g)	0.263	0.118	0.000	1.084

Notes: This table indicates descriptive Statistics (N = 575). Statistics are weighted using the number of births in the cell. "NutriPlus spending per capita in NBR" indicates the average NutriPlus spending in neighboring counties. See Table C.1.3 for unweighted descriptive statistics.

Table 3.2. First-stage regression

	OLS					
	(1)	(2)	(3)	(4)	(5)	(6)
NBR (1,000 KRW)	0.618***	0.193*	0.656***	0.621***	0.202*	0.668***
	(0.089)	(0.099)	(0.092)	(0.093)	(0.097)	(0.101)
Gestational age	Yes	Yes	Yes	No	No	No
County fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	No	Yes	No	No	Yes	No
County × Time trends	No	No	Yes	No	No	Yes
Observations	354	354	354	354	354	354
F-test on instrument	47.92	3.79	51.27	44.56	4.28	44.10

Notes: Table 3.2 reports estimates from Eq. (3.2). The dependent variable is the NutriPlus spending per capita in a domestic county (1,000 KRW). Estimates are weighted using the number of births in the cell. Robust standard errors clustered by county are reported in parentheses. The number of observations in this and all subsequent tables refers to county-year cells. "NBR" indicates the average NutriPlus spending per capita in neighboring counties. All specifications include age below 20 (%), age 35 or above (%), college (%), male newborns (%), and the log of real GRDP per capita. See Table C.1.4 for details. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 3.3. Effects of NutriPlus spending on birth outcomes

	OLS	OLS	OLS	IV	IV	IV
	(1)	(2)	(3)	(4)	(5)	(6)
(i) Dependent variable: Birth weight (in grams)						
NutriPlus spending	-8.116**	-1.130	-9.038***	-16.459***	-19.059	-13.134***
	(2.997)	(3.507)	(2.752)	(3.464)	(29.252)	(3.574)
1% spending	-0.027	-0.004	-0.030	-0.055	-0.063	-0.044
1% impact	-0.001%	0.000%	-0.001%	-0.002%	-0.002%	-0.001%
100% spending	-2.7	-0.376	-3.007	-5.475	-6.34	-4.369
100% impact	-0.083%	-0.012%	-0.092%	-0.168%	-0.195%	-0.134%
R^2	0.841	0.888	0.884	0.837	0.878	0.884
(ii) Dependent variable: Low birth weight (%)						
NutriPlus spending	0.114	0.162	0.156	0.091	-0.157	0.069
	(0.081)	(0.117)	(0.113)	(0.140)	(0.607)	(0.152)
1% spending	0.000	0.001	0.001	0.000	-0.001	0.000
100% spending	0.038	0.054	0.052	0.030	-0.052	0.023
R^2	0.635	0.675	0.667	0.642	0.670	0.674
(iii) Dependent variable: Very low birth weight (%)						
NutriPlus spending	-0.009	0.029	-0.023	-0.044	-0.074	-0.050
	(0.019)	(0.029)	(0.023)	(0.027)	(0.127)	(0.031)
1% spending	0.000	0.000	0.000	0.000	0.000	0.000
100% spending	-0.003	0.010	-0.008	-0.015	-0.024	-0.017
R^2	0.512	0.550	0.546	0.514	0.541	0.550
(iv) Dependent variable: High birth weight (%)						
NutriPlus spending	-0.061**	-0.034	-0.041*	-0.081***	-0.085	-0.064**
	(0.022)	(0.030)	(0.024)	(0.023)	(0.113)	(0.025)
1% spending	0.000	0.000	0.000	0.000	0.000	0.000
100% spending	-0.020	-0.011	-0.014	-0.027	-0.028	-0.021
R^2	0.487	0.531	0.508	0.485	0.524	0.508
(v) Dependent variable: Normal birth weight (%)						
NutriPlus spending	-0.053	-0.128	-0.114	-0.010	0.242	-0.005

	(1)	(2)	(3)	(4)	(5)	(6)
	(0.092)	(0.127)	(0.127)	(0.149)	(0.656)	(0.162)
1% spending	0.000	0.000	0.000	0.000	0.001	0.000
100% spending	-0.018	-0.042	-0.038	-0.003	0.081	-0.002
R^2	0.537	0.587	0.577	0.545	0.577	0.585
(vi) Dependent variable: Premature birth (%)						
NutriPlus spending	0.118	0.096	0.152	0.106	-0.467	0.157
	(0.092)	(0.159)	(0.090)	(0.174)	(0.981)	(0.134)
1% spending	0.000	0.000	0.001	0.000	-0.002	0.001
100% spending	0.039	0.032	0.051	0.035	-0.155	0.052
R^2	0.697	0.729	0.726	0.700	0.714	0.728
Year fixed effects	No	Yes	No	No	Yes	No
County × Time trends	No	No	Yes	No	No	Yes
Observations	437	437	437	437	437	437

Notes: Columns (1)–(3) report the estimates from Eq. (3.1). Columns (4)–(6) report the estimates from Eq. (3.3). Instrumental variable is the average NutriPlus spending per capita in neighboring counties. Estimates are weighted using the number of births in the cell. Robust standard errors clustered by county are reported in parentheses. The age below 20 (%), age 35 or above (%), college (%), male newborn (%), the log of real GRDP per capita, and county fixed effects are included in panels (i) to (vi). The gestational age is included only in panels (i) to (v). See Tables C.1.5 to C.1.10 for details. 1% spending is calculated as coefficient × average NutriPlus spending per capita × 0.01. 1% impact is calculated as 1% spending / mean birth weight × 100%. 100% spending is calculated as coefficient × average NutriPlus spending per capita. 100% impact is calculated as 100% spending / mean birth weight × 100%. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 3.4. First-stage regression using linear interpolation

	OLS					
	(1)	(2)	(3)	(4)	(5)	(6)
NBR (1,000 KRW)	0.513***	0.289***	0.547***	0.514***	0.289***	0.559***
	(0.084)	(0.091)	(0.088)	(0.088)	(0.091)	(0.096)
Gestational age	Yes	Yes	Yes	No	No	No
Year fixed effects	No	Yes	No	No	Yes	No
County × Time trends	No	No	Yes	No	No	Yes
Observations	437	437	437	437	437	437
F-test on instrument	37.54	10.10	38.29	34.51	10.05	34.02

Notes: Table 3.4 reports estimates from Eq. (3.2). The dependent variable is the NutriPlus spending per capita in a domestic county (1,000 KRW). Estimates are weighted using the number of births in the cell. Robust standard errors clustered by county are reported in parentheses. "NBR" indicates the average NutriPlus spending per capita in neighboring counties. All specifications include age below 20, age 35 or above, college, male newborns, and real GRDP per capita. See Table A11 for details. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 3.5. Effects of NutriPlus spending on birth outcomes using linear interpolation

	OLS	OLS	OLS	IV	IV	IV
	(1)	(2)	(3)	(4)	(5)	(6)
(i) Dependent variable: Birth weight (in grams)						
NutriPlus spending	-7.122**	1.449	-7.933***	-18.399***	1.426	-16.007***
	(2.867)	(2.895)	(2.440)	(4.941)	(13.434)	(4.262)
1% spending	-0.025	0.005	-0.027	-0.063	0.005	-0.055
1% impact	-0.001	0.000	-0.001	-0.002	0.000	-0.002
100% spending	-2.45	0.498	-2.729	-6.330	0.491	-5.508
100% impact	-0.075	0.015	-0.084	-0.195	0.015	-0.169
R^2	0.816	0.877	0.867	0.807	0.877	0.864
(ii) Dependent variable: Low birth weight (%)						
NutriPlus spending	0.117	0.116	0.143	-0.041	-0.460	-0.056
	(0.078)	(0.071)	(0.106)	(0.157)	(0.321)	(0.161)
1% spending	0.000	0.000	0.000	0.000	-0.002	0.000
100% spending	0.040	0.040	0.049	-0.014	-0.158	-0.019
R^2	0.629	0.661	0.660	0.624	0.624	0.655
(iii) Dependent variable: Very low birth weight (%)						
NutriPlus spending	0.001	0.018	-0.009	-0.079**	-0.161	-0.091*
	(0.019)	(0.018)	(0.021)	(0.040)	(0.106)	(0.047)
1% spending	0.000	0.000	0.000	0.000	-0.001	0.000
100% spending	0.000	0.006	-0.003	-0.027	-0.055	-0.031
R^2	0.481	0.512	0.510	0.466	0.462	0.498
(iv) Dependent variable: High birth weight (%)						
NutriPlus spending	-0.056**	-0.035	-0.039*	-0.083***	-0.039	-0.069***
	(0.022)	(0.031)	(0.022)	(0.020)	(0.042)	(0.022)
1% spending	0.000	0.000	0.000	0.000	0.000	0.000
100% spending	-0.019	-0.012	-0.013	-0.029	-0.013	-0.024
R^2	0.465	0.504	0.483	0.463	0.504	0.481
(v) Dependent variable: Normal birth weight (%)						
NutriPlus spending	-0.061	-0.081	-0.104	0.124	0.499	0.125

	(0.089)	(0.079)	(0.121)	(0.161)	(0.325)	(0.170)
1% spending	0.000	0.000	0.000	0.000	0.002	0.000
100% spending	-0.021	-0.028	-0.036	0.043	0.172	0.043
R^2	0.539	0.575	0.576	0.532	0.532	0.569
(vi) Dependent variable: Premature birth (%)						
NutriPlus spending	0.110	0.046	0.158**	0.128	-0.199	0.202
	(0.081)	(0.098)	(0.072)	(0.215)	(0.588)	(0.156)
1% spending	0.000	0.000	0.001	0.000	-0.001	0.001
100% spending	0.038	0.016	0.054	0.044	-0.069	0.070
R^2	0.682	0.711	0.714	0.681	0.707	0.713
Year fixed effects	No	Yes	No	No	Yes	No
County × Time trends	No	No	Yes	No	No	Yes
Observations	437	437	437	437	437	437

Notes: Columns (1)–(3) report the estimates from Eq. (3.1). Columns (4)–(6) report the estimates from Eq. (3.3). Instrumental variable is the average NutriPlus spending per capita in neighboring counties. Estimates are weighted using the number of births in the cell. Robust standard errors clustered by county are reported in parentheses. The age below 20, age 35 or above, college, male newborn, real GRDP per capita, and county fixed effects are included in panels (i) to (vi). The gestational age is included only in panels (i) to (v). See Tables C.1.12 to C.1.17 for details. 1% spending is calculated as coefficient × average NutriPlus spending per capita × 0.01. 1% impact is calculated as 1% spending / mean birth weight × 100%. 100% spending is calculated as coefficient × average NutriPlus spending per capita. 100% impact is calculated as 100% spending / mean birth weight × 100%. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

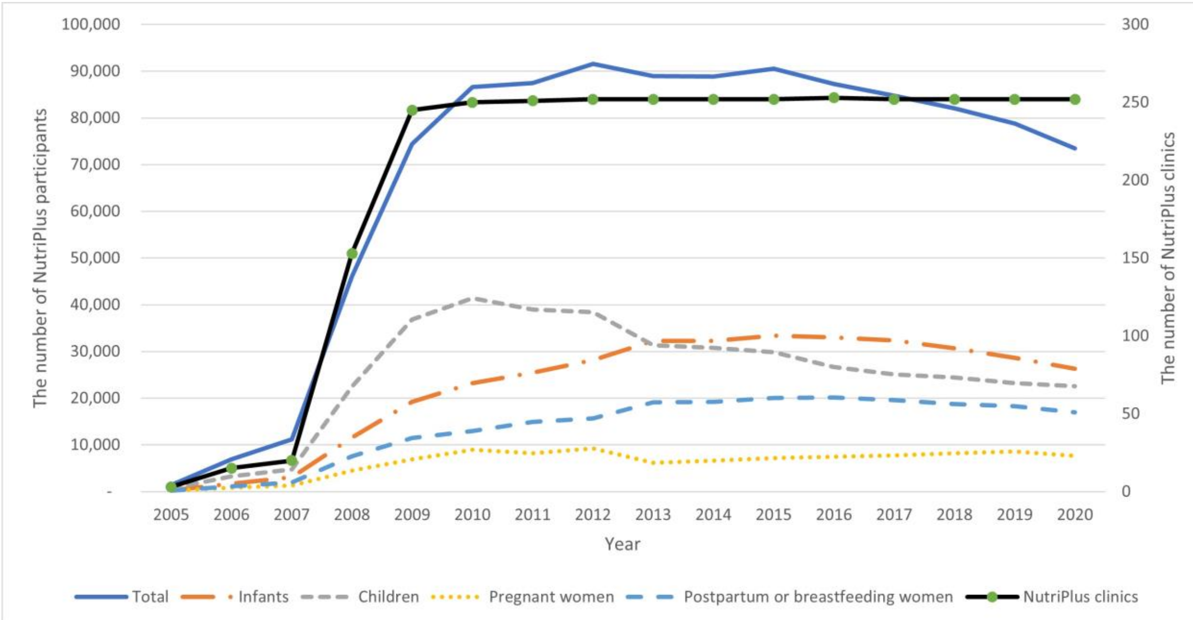


Figure 3.1. The trends of Nutriplus clinics and participants in South Korea

Notes: The left y-axis indicates the number of NutriPlus participants. The right y-axis indicates the number of NutriPlus clinics (i.e., public health centers providing NutriPlus services). This figure was created by authors using data from the Korea Health Promotion Institute.

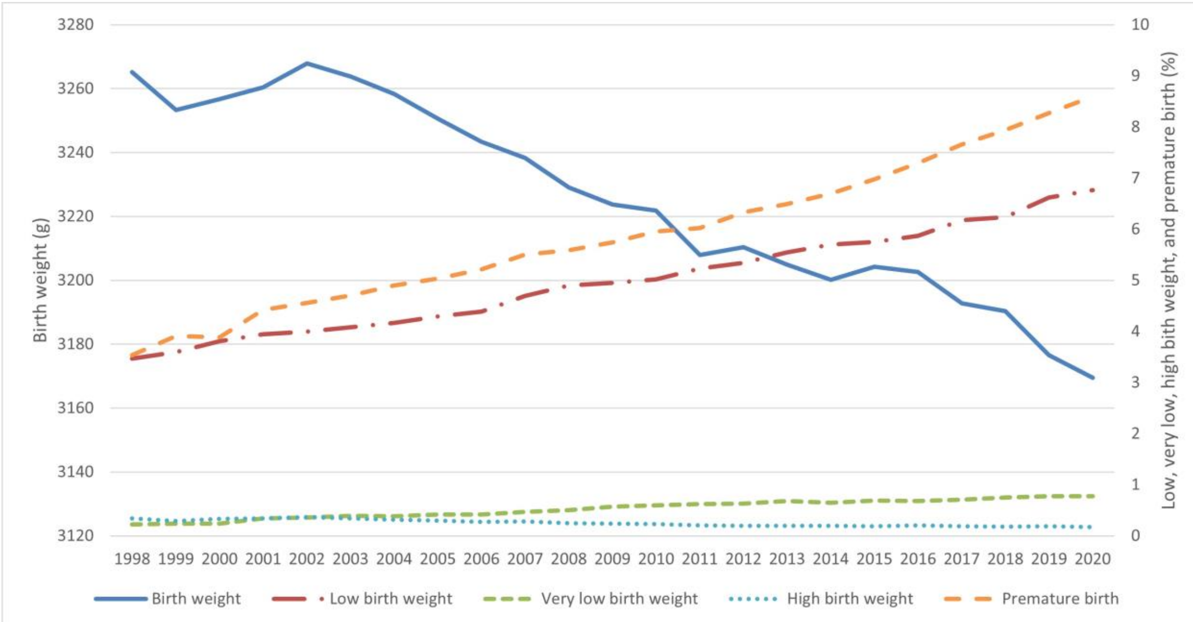


Figure 3.2. The trends of birth outcomes in South Korea

Notes: The left y-axis indicates the average birth weight in grams. The right y-axis indicates the fraction of low birth weight, very low birth weight, high birth weight, and premature birth. This figure was created by authors using the birth certificate data from the "microdata integrated service" provided by Statistics Korea.

REFERENCES

Adalja, Aaron, Erik Lichtenberg, and Elina T. Page. 2022. "Collective Investment in A Common Pool Resource: Grower Associations and Food Safety Guidelines." *American Journal of Agricultural Economics* 1–30. https://doi.org/10.1111/ajae.12291

Alberini, Anna, Erik Lichtenberg, Dominic Mancini, and Gregmar I. Galinato. 2008. "Was It Something I Ate? Implementation of the FDA Seafood HACCP Program." *American Journal of Agricultural Economics* 90(1): 28–41.

Almond, Douglas, Kenneth Y. Chay, and David S. Lee. 2005. "The Costs of Low Birth Weight." *The Quarterly Journal of Economics* 120(3): 1031–1083.

Angrist, Joshua D., and Jörn-Steffen Pischke. 2009. *Mostly harmless econometrics: An empiricist's companion*. Princeton university press.

Arnade, Carlos, Fred Kuchler, and Linda Calvin. 2016. "The Changing Role of Consumers and Suppliers in a Food Safety Event: the 2006 Food-Borne Illness Outbreak Linked to Spinach." *Applied Economics* 48(25): 2354–2366.

Bederson, Benjamin B., Ginger Z. Jin, Phillip Leslie, Alexander J. Quinn, and Ben Zou. 2018. "Incomplete Disclosure: Evidence of Signaling and Countersignaling." *American Economic Journal: Microeconomics* 10(1): 41–66.

Behrman, Jere R., and Mark R. Rosenzweig. 2004. "Returns to Birthweight." *The Review of Economics and Statistics* 86(2): 586–601.

Bellemare, Marc F., and Ngoc Nguyen. 2018. "Farmers Markets and Food-Borne Illness." *American Journal of Agricultural Economics* 100(3): 676–690.

Bilinski, Alyssa, and Laura A. Hatfield. 2020. "Nothing to see here? non-inferiority approaches to parallel trends and other model assumptions." Working Paper.

Black, Sandra E., Paul J. Devereux, and Kjell G. Salvanes. 2007. "From the Cradle to the Labor Market? The Effect of Birth Weight on Adult Outcomes." *The Quarterly Journal of Economics* 122(1): 409–439.

Bodea, Cristina, Masaaki Higashijima, and Raju J. Singh. 2016. "Oil and Civil Conflict: Can Public Spending Have a Mitigation Effect?" *World Development* 78: 1–12.

Brown, Brad, Marc Allard, Michael C. Bazaco, Joseph Blankenship, and Travis Minor. 2021. "An Economic Evaluation of the Whole Genome Sequencing Source Tracking Program in the U.S.." *PLOS ONE* 16(10): e0258262.

Calonico, Sebastian, Matias D. Cattaneo, and Rocío Titiunik. 2014. "Robust Data-Driven Inference in the Regression-Discontinuity Design." *The Stata Journal* 14(4): 909–946.

Calonico, Sebastian, Matias D. Cattaneo, Max H. Farrell, and Rocío Titiunik. 2017. "rdrobust: Software for regression-discontinuity designs." *The Stata Journal* 17(2): 372–404.

Cho, Hyunkuk, and Yong W. Lee. 2022. "Multiple births and low birth weight: Evidence from South Korea." *American Journal of Human Biology* 34(3): e23648.

Chorniy, Anna, Janet Currie, and Lyudmyla Sonchak. 2020. "Does Prenatal WIC Participation Improve Child Outcomes?" *American Journal of Health Economics* 6(2): 169–198.

Clarke, Damian, and Kathya Tapia-Schythe. 2021. "Implementing the Panel Event Study." *The*

Stata Journal 21(4): 853–884.

Collier, Paul, and Anke Hoeffler. 2004. "Greed and grievance in civil war." *Oxford Economic Papers* 56(4): 563–595.

Dai, Weijia, and Michael Luca. 2020. "Digitizing Disclosure: The Case of Restaurant Hygiene Scores." *American Economic Journal: Microeconomics* 12(2): 41–59.

Datta Gupta, Nabanita, Mette Deding, and Mette Lausten. 2013. "The effect of low birth weight on height, weight and behavioral outcomes in the medium-run." *Economics & Human Biology* 11(1): 42–55.

DeAngelo, Gregory, and Benjamin Hansen. 2014. "Life and Death in the Fast Lane: Police Enforcement and Traffic Fatalities." *American Economic Journal: Economic Policy* 6 (2): 231–57.

Fletcher, Jason M. 2011. "The medium term schooling and health effects of low birth weight: Evidence from siblings." *Economics of Education Review* 30(3): 517–527.

Gelman, Andrew, and Guido Imbens. 2019. "Why High-Order Polynomials Should Not Be Used in Regression Discontinuity Designs." *Journal of Business & Economic Statistics*, 37(3): 447–456.

Geng, Yong, Wei Liua, Kai Li, and Hanshu Chena. 2021. "Environmental regulation and corporate tax avoidance: A quasi-natural experiment based on the eleventh Five-Year Plan in China." *Energy Economics* 99: 105312.

Giambona, Erasmo, and Rafael P. Ribas. 2023. "Unveiling the Price of Obscenity: Evidence from Closing Prostitution Windows in Amsterdam." *Journal of Policy Analysis and*

Management https://doi.org/10.1002/pam.22459

Greene, William H. 2018. *Econometric analysis*. New York: Pearson.

Grembi, Veronica, Tommaso Nannicini, and Ugo Troiano. 2016. “Do Fiscal Rules Matter?” *American Economic Journal: Applied Economics* 8(3): 1–30.

Gujarati, Damodar. 2015. *Econometrics by example*. London: Palgrave.

Hansen, Benjamin, Keaton Miller, and Caroline Weber. 2020. “Federalism, partial prohibition, and cross-border sales: Evidence from recreational marijuana.” Journal of Public Economics 187: 104159.

He, Guojun, Shaoda Wang, and Bing Zhang. 2020. “Watering Down Environmental Regulation in China.” *The Quarterly Journal of Economics* 135(4): 2135–2185.

Ho, Daniel E., Zoe C. Ashwood, and Cassandra Handan-Nader. 2019. “New Evidence on Information Disclosure through Restaurant Hygiene Grading.” *American Economic Journal: Economic Policy* 11(4): 404–428.

Hummer, Michael, Thomas Lehner, and Gerald Pruckner. 2014. “Low birth weight and health expenditures from birth to late adolescence.” *The European Journal of Health Economics* 15(3): 229–242.

Jin, Ginger Z., and Phillip Leslie. 2003. “The Effect of Information on Product Quality: Evidence from Restaurant Hygiene Grade Cards.” *The Quarterly Journal of Economics* 118(2): 409–451.

Jin, Ginger Z., and Phillip Leslie. 2019. “New Evidence on Information Disclosure through Restaurant Hygiene Grading: Reply.” *American Economic Journal: Economic Policy*

11(4): 429–443.

Johnsson, Inger W., Bengt Haglund, Fredrik Ahlsson, and Jan Gustafsson. 2015. “High birth weight and type 2 diabetes.” *Pediatric Obesity*, 10(2): 77-83.

Joyce, Ted, Diane Gibson, and Silvie Colman. 2005. “The changing association between prenatal participation in WIC and birth outcomes in New York City.” *Journal of Policy Analysis and Management* 24(4): 661–685.

Justino, Patricia, Bruno Martorano. 2018. “Welfare spending and political conflict in Latin America, 1970–2010.” *World Development* 107: 98–110.

Kahn-Lang, Ariella, and Kevin Lang. 2020. “The Promise and Pitfalls of Differences-in-Differences: Reflections on 16 and Pregnant and Other Applications.” *Journal of Business & Economic Statistics* 38(3): 613–620.

Koo, Y.-J., H.-M. Ryu, J.-H. Yang, J.-H. Lim, J.-E. Lee, M.-Y. Kim, and J.-H. Chung. 2012. “Pregnancy outcomes according to increasing maternal age.” *Taiwanese Journal of Obstetrics and Gynecology* 51(1): 60–65.

Krinsky, S., J. Errickson, and A.M. Ryan. 2022. “Grade pending: the effect of the New York City restaurant sanitary grades inspection program on Salmonellosis.” *Journal of Public Health* 30: 1171–1177.

Lee, J.W., and D.H. Baek. 2019. “Disclosure of Restaurant Hygiene Information and Food Poisoning: Focused to Restaurant Hygiene Grading in Seoul.” *Korean Society and Public Administration* 30(1): 255–276.

Liu, Cenjie, Jiayu Fang, and Rui Xie. 2021. “Energy policy and corporate financial performance:

Evidence from China's 11th five-year plan." *Energy Economics* 93: 105030.

Minor, Travis, and Matt Parrett. 2017. "The Economic Impact of the Food and Drug Administration's Final Juice HACCP Rule." *Food Policy* 68: 206–213.

Ollinger, Michael, and John Bovay. 2018. "Pass or Fail: Economic Incentives to Reduce Salmonella Contamination in Ground Beef Sold to the National School Lunch Program." *American Journal of Agricultural Economics* 100(2): 414–433.

Ollinger, Michael, and John Bovay. 2020. "Producer Response to Public Disclosure of Food-Safety Information." *American Journal of Agricultural Economics* 102(1): 186–201.

Pehkonen, Jaakko, Jutta Viinikainen, Jaana T. Kari, Petri Böckerman, Terho Lehtimäki, and Olli Raitakari. 2021. "Birth weight and adult income: An examination of mediation through adult height and body mass." *Health Economics* 30(10): 2383–2398.

Scharff, Robert L., John Besser, Donald J. Sharp, Timothy F. Jones, Gerner-Smidt Peter, and Craig W. Hedberg. 2016. "An Economic Evaluation of PulseNet." *American Journal of Preventive Medicine* 50(5S1): S66–S73.

Scharff, Robert L. 2018. "The Economic Burden of Foodborne Illness in the United States." In *Food Safety Economics*, edited by Tanya Roberts, 123–142. Dordrecht: Springer.

Sonchak, Lyudmyla. 2016. "The Impact of WIC on Birth Outcomes: New Evidence from South Carolina." *Maternal and Child Health Journal* 20(7): 1518–1525.

Wooldridge, Jeffrey M. 2019. *Introductory Econometrics: A Modern Approach*. Singapore: Cengage.

Wooldridge, Jeffrey M. 2021. "Two-Way Fixed Effects, the Two-Way Mundlak Regression, and

Difference-in-Differences Estimators." Available at SSRN: https://ssrn.com/abstract=3906345

World Health Organization. 2022. "Food safety." who.int/news-room/fact-sheets/detail/food-safety Accessed 8 July 2022

APPENDIX A

CHAPTER 1 APPENDIX

A.1. Supporting tables and figures

Table A.1.1. RD and differences-in-discontinuities estimates of the EWS on the number of outbreaks of food-borne illness per million (bandwidth: 24 months): OLS regression

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Total	Norovirus	E. coli	Salmonella	Vibrio	S. aureus	C. jejuni	Perfringens	B. cereus	Unknown
Panel A. RD estimates										
Model 1. linear running variable										
Post	-0.526*	-0.067	-0.322*	-0.020	-0.009	-0.084	0.031	0.009	0.040	-0.102
	(0.289)	(0.152)	(0.150)	(0.020)	(0.020)	(0.137)	(0.062)	(0.019)	(0.036)	(0.125)
% Change from Mean	-61%	-33%	-118%	-298%	-57%	-88%	257%	53%	528%	-45%
Obs.	686	686	686	686	686	686	686	686	686	686
Model 2. quadratic running variable										
Post	-0.137	0.216	-0.377*	-0.006	0.001	0.017	0.022	0.020	0.037	-0.061
	(0.291)	(0.178)	(0.205)	(0.006)	(0.039)	(0.125)	(0.045)	(0.105)	(0.051)	(0.133)
% Change from Mean	-16%	105%	-138%	-89%	6%	18%	182%	118%	489%	-27%
Obs.	686	686	686	686	686	686	686	686	686	686
Panel B. differences-in-discontinuities estimates										
Model 3. linear running variable										
Post × School	-0.493	-0.052	-0.329**	-0.020	-0.007	-0.092	0.032	0.011	0.035	-0.072
	(0.288)	(0.152)	(0.152)	(0.022)	(0.025)	(0.138)	(0.061)	(0.019)	(0.037)	(0.121)
% Change from Mean	-58%	-25%	-121%	-298%	-44%	-96%	265%	65%	462%	-32%
Obs.	2,744	2,744	2,744	2,744	2,744	2,744	2,744	2,744	2,744	2,744
Model 4. quadratic running variable										
Post × School	-0.137	0.212	-0.390*	-0.002	0.001	0.007	0.023	0.025	0.025	-0.031
	(0.294)	(0.182)	(0.207)	(0.010)	(0.044)	(0.128)	(0.045)	(0.105)	(0.052)	(0.128)
% Change from Mean	-16%	103%	-143%	-30%	6%	7%	190%	148%	330%	-14%
Obs.	2,744	2,744	2,744	2,744	2,744	2,744	2,744	2,744	2,744	2,744
Pre-Period Outcome Mean	0.857	0.205	0.273	0.007	0.016	0.095	0.012	0.017	0.008	0.225

Notes: Columns indicate different outcome variables. Panel A reports RD estimates from Eq. (1.1). Panel B reports differences-in-discontinuities estimates from Eq. (1.2). We do not have enough variation in Protozoa. Standard errors clustered at the region level are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.1.2. RD and differences-in-discontinuities estimates of the EWS on the number of cases of food-borne illness per million (bandwidth: 24 months): OLS regression

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Total	Norovirus	E. coli	Salmonella	Vibrio	S. aureus	C. jejuni	Perfringens	B. cereus	Unknown
Panel A. RD estimates										
Model 1. linear running variable										
Post	-1.183	15.181	-9.491	-0.159	-1.777	0.048	-6.963	3.843	2.909	-4.747
	(23.585)	(8.805)	(13.823)	(0.161)	(2.364)	(8.630)	(7.977)	(2.694)	(1.967)	(5.745)
% Change from Mean	-2%	96%	-34%	-296%	-139%	1%	-265%	308%	768%	-64%
Obs.	686	686	686	686	686	686	686	686	686	686
Model 2. quadratic running variable										
Post	40.580	41.320**	-5.303	-0.051	-1.864	1.856	-4.068	10.811	3.539	-5.541
	(30.074)	(15.660)	(16.237)	(0.052)	(3.781)	(9.815)	(4.104)	(16.840)	(2.267)	(4.757)
% Change from Mean	65%	261%	-19%	-95%	-146%	30%	-155%	867%	935%	-74%
Obs.	686	686	686	686	686	686	686	686	686	686
Panel B. differences-in-discontinuities estimates										
Model 3. linear running variable										
Post × School	-0.469	16.204*	-10.091	0.026	-1.857	0.110	-6.906	3.914	2.861	-4.717
	(23.143)	(8.776)	(13.886)	(0.216)	(2.401)	(8.719)	(7.941)	(2.673)	(1.956)	(5.813)
% Change from Mean	-1%	103%	-36%	48%	-145%	2%	-262%	314%	756%	-63%
Obs.	2,744	2,744	2,744	2,744	2,744	2,744	2,744	2,744	2,744	2,744
Model 4. quadratic running variable										
Post × School	39.876	41.184**	-5.814	0.138	-2.020	1.559	-3.806	10.916	3.440	-5.565
	(29.486)	(15.657)	(16.356)	(0.118)	(3.796)	(9.970)	(4.113)	(16.740)	(2.256)	(4.870)
% Change from Mean	63%	261%	-21%	257%	-158%	25%	-145%	876%	909%	-75%
Obs.	2,744	2,744	2,744	2,744	2,744	2,744	2,744	2,744	2,744	2,744
Pre-Period Outcome Mean	62.946	15.811	27.910	0.054	1.281	6.178	2.633	1.247	0.379	7.454

Notes: Columns indicate different outcome variables. Panel A reports RD estimates from Eq. (1.1). Panel B reports differences-in-discontinuities estimates from Eq. (1.2). We do not have enough variation in Protozoa. Standard errors clustered at the region level are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.1.3. RD and differences-in-discontinuities estimates of the EWS on the number of outbreaks of food-borne illness per million (bandwidth: 12 months): Poisson regression

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Total	Norovirus	E. coli	Salmonella	Vibrio	S. aureus	C. jejuni	Perfringens	B. cereus	Unknown
Panel A. RD estimates										
Model 1. linear running variable										
Post	-0.918	-0.931	-2.046			0.419	0.016			-2.322**
	(0.588)	(0.725)	(1.677)			(1.509)	(1.466)			(1.117)
Obs.	322	266	168			196	56			182
Model 2. quadratic running variable										
Post	-0.675	4.791***	-3.216			0.016	0.016			2.444
	(1.706)	(1.788)	(5.107)			(1.468)	(1.466)			(3.189)
Obs.	322	266	168			112	56			182
Panel B. differences-in-discontinuities estimates										
Model 3. linear running variable										
Post × School	0.277	0.772	0.325			-23.837***	0.512			-0.409
	(0.552)	(0.624)	(1.990)			(1.337)	(1.875)			(1.445)
Obs.	1372	1232	756			1400	938			1232
Model 4. quadratic running variable										
Post × School	-1.271	4.628*	1.319			-34.558	0.719			1.361
	(2.173)	(2.748)	(5.741)			(149.423)	(1.682)			(3.924)
Obs.	1372	1232	756			1400	784			1232

Notes: Columns indicate different outcome variables. Panel A reports RD estimates from Eq. (1.1). Panel B reports differences-in-discontinuities estimates from Eq. (1.2). We do not have enough variation in Protozoa. Standard errors clustered at the region level are in parentheses. Some estimates and standard errors are not calculated due to convergence issues. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.1.4. RD and differences-in-discontinuities estimates of the EWS on the number of cases of food-borne illness per million (bandwidth: 12 months): Poisson regression

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Total	Norovirus	E. coli	Salmonella	Vibrio	S. aureus	C. jejuni	Perfringens	B. cereus	Unknown
Panel A. RD estimates										
Model 1. linear running variable										
Post	-0.462	-0.507	-1.839			1.973	0.227			-1.997*
	(0.677)	(0.985)	(2.659)			(1.882)	(1.466)			(1.184)
Obs.	322	266	168			196	56			182
Model 2. quadratic running variable										
Post	-2.324	3.471	-8.083			0.113	0.227			6.559**
	(1.843)	(2.116)	(7.373)			(1.468)	(1.466)			(3.203)
Obs.	322	266	168			112	56			182
Panel B. differences-in-discontinuities estimates										
Model 3. linear running variable										
Post × School	0.501	1.819	2.849			3.537				-1.148
	(0.639)	(1.133)	(3.001)			(2.883)				(1.365)
Obs.	1372	1232	756			938				1232
Model 4. quadratic running variable										
Post × School	-4.871*	0.292	-4.815			8.445				2.987
	(2.637)	(1.995)	(8.580)			(9.048)				(4.278)
Obs.	1372	1232	756			784				1232

Notes: Columns indicate different outcome variables. Panel A reports RD estimates from Eq. (1.1). Panel B reports differences-in-discontinuities estimates from Eq. (1.2). We do not have enough variation in Protozoa. Standard errors clustered at the region level are in parentheses. Some estimates and standard errors are not calculated due to convergence issues. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.1.5. RD and differences-in-discontinuities estimates of the EWS on the number of outbreaks of food-borne illness per million (bandwidth: 24 months): Poisson regression

	(1) Total	(2) Norovirus	(3) E. coli	(4) Salmonella	(5) Vibrio	(6) S. aureus	(7) C. jejuni	(8) Perfringens	(9) B. cereus	(10) Unknown
Panel A. RD estimates										
Model 1. linear running variable										
Post	-0.348	0.571	-1.151		-17.028***	-1.409	-1.241	8.043***	-7.826	-0.932
	(0.650)	(1.112)	(0.745)		(1.369)	(1.609)	(1.843)	(2.083)	(8.329)	(1.173)
Obs.	686	630	448		686	560	168	182	686	574
Model 2. quadratic running variable										
Post	0.217	2.579	-2.663		-104.085***	-0.479				-0.691
	(0.832)	(2.053)	(1.708)		(31.704)	(1.730)				(1.995)
Obs.	686	630	448		686	560				574
Panel B. differences-in-discontinuities estimates										
Model 3. linear running variable										
Post × School	0.078	1.461	-1.589	-26.580***	-13.765***	-1.878	0.428			-0.167
	(0.642)	(1.237)	(1.058)	(1.886)	(1.811)	(2.220)	(2.452)			(1.181)
Obs.	2744	2520	1960	2744	2744	2282	1008			2632
Model 4. quadratic running variable										
Post × School	0.146	2.216	-3.706	-0.767	-72.261***	-2.264				0.886
	(0.959)	(2.330)	(2.849)	(2.452)	(5.077)	(2.384)				(2.168)
Obs.	2744	2520	1960	2744	2744	2282				2632

Notes: Columns indicate different outcome variables. Panel A reports RD estimates from Eq. (1.1). Panel B reports differences-in-discontinuities estimates from Eq. (1.2). We do not have enough variation in Protozoa. Standard errors clustered at the region level are in parentheses. Some estimates and standard errors are not calculated due to convergence issues. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.1.6. RD and differences-in-discontinuities estimates of the EWS on the number of cases of food-borne illness per million (bandwidth: 24 months): Poisson regression

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Total	Norovirus	E. coli	Salmonella	Vibrio	S. aureus	C. jejuni	Perfringens	B. cereus	Unknown
Panel A. RD estimates										
Model 1. linear running variable										
Post	0.495	2.796**	-0.314			-0.436	-4.871**	10.217***		-0.732
	(0.627)	(1.176)	(0.862)			(1.663)	(2.179)	(2.427)		(1.292)
Obs.	686	630	448			560	168	182		574
Model 2. quadratic running variable										
Post	1.316	8.163**	-0.687			-1.843				-1.248
	(1.008)	(3.392)	(3.849)			(2.007)				(3.398)
Obs.	686	630	448			560				574
Panel B. differences-in-discontinuities estimates										
Model 3. linear running variable										
Post × School	0.909	5.818***	-2.363	-28.592***		-0.641	-3.669			-1.334
	(0.743)	(1.681)	(1.639)	(1.648)		(2.642)	(2.571)			(1.810)
Obs.	2744	2520	1960	2744		2282	1008			2632
Model 4. quadratic running variable										
Post × School	0.593	8.622***	5.637	-1.855		-4.046*				-0.995
	(1.278)	(3.294)	(5.456)	(16.091)		(2.189)				(4.221)
Obs.	2744	2520	1960	2744		2282				2632

Notes: Columns indicate different outcome variables. Panel A reports RD estimates from Eq. (1.1). Panel B reports differences-in-discontinuities estimates from Eq. (1.2). We do not have enough variation in Protozoa. Standard errors clustered at the region level are in parentheses. Some estimates and standard errors are not calculated due to convergence issues. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.1.7. Event study: The joint significance tests for parallel trends assumption in the pre-treatment period

	(1) Total	(2) Norovirus	(3) E. coli	(4) Salmonella	(5) Vibrio	(6) S. aureus	(7) C. jejuni	(8) Perfringens	(9) B. cereus	(10) Unknown
Panel A.										
Dependent variable: The number of outbreaks of food-borne illness per million										
P-value	0.000	0.000	0.000	0.000	0.050	0.191	0.662	0.174	0.599	0.000
Panel B.										
Dependent variable: The number of cases of food-borne illness per million										
P-value	0.000	0.000	0.000	0.000	0.135	0.191	0.234	0.683	0.531	0.000
Obs.	12,768	12,768	12,768	12,768	12,768	12,768	12,768	12,768	12,768	12,768

Notes: We formally tested that all the event study lead coefficients from Eq. (1.4) are simultaneously zero. We do not have enough variation in Protozoa from Eq. (1.4).

Table A.1.8. DID estimates of the EWS on the number of outbreaks of food-borne illness per million (without differential linear time trends): Poisson regression with Mundlak approach

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Total	Norovirus	E. coli	Salmonella	Vibrio	S. aureus	C. jejuni	Perfringens	B. cereus	Unknown
Panel A. Period 1 (Jan. 2002 – Oct. 2013)										
Post × School	0.247	0.215	0.505	-1.311		-0.635	-0.257	-0.082	0.804	-0.117
	(0.189)	(0.376)	(0.380)	(1.253)		(0.564)	(0.565)	(0.576)	(0.926)	(0.411)
Obs.	7560	4256	3640	3248		3192	1568	1400	1288	5264
Panel B. Period 2 (Jan. 2002 – Aug. 2014)										
Post× School	0.281	0.238	0.553	-0.759		-0.494	-0.517	0.061	0.701	-0.119
	(0.198)	(0.385)	(0.395)	(1.197)		(0.590)	(0.521)	(0.555)	(0.891)	(0.360)
Obs.	8120	4592	3920	3584		3248	1792	1848	1344	5656
Panel C. Period 3 (Jan. 2002 – Dec. 2020)										
Post × School	0.038	-0.280	0.298	-0.450		-0.307	-0.557	-0.335	0.542	-0.301
	(0.206)	(0.373)	(0.327)	(0.919)		(0.437)	(0.596)	(0.609)	(0.861)	(0.332)
Obs.	12264	7560	6328	5040		3808	3192	2744	1904	9072

Notes: Each panel reports DID estimates from Poisson regression with Mundlak approach. (see Wooldridge, 2021). Columns indicate different outcome variables. 20–44 years (%), 45–64 years (%), 65 years and over (%), male (%), middle school students (%), high school students (%), unemployment rate (%), time fixed effects, the average of the covariates over the time periods, and constant are included in all models. We do not have enough variation in Protozoa. Standard errors clustered at the region level are in parentheses. Some estimates and standard errors are not calculated due to convergence issues. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.1.9. DID estimates of the EWS on the number of cases of food-borne illness per million (without differential linear time trends): Poisson regression with Mundlak approach

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Total	Norovirus	E. coli	Salmonella	Vibrio	S. aureus	C. jejuni	Perfringens	B. cereus	Unknown
Panel A. Period 1 (Jan. 2002 – Oct. 2013)										
Post × School	-0.011	-0.593	0.637	-0.199		-0.814	0.126	-1.227	-0.174	-0.207
	(0.196)	(0.552)	(0.438)	(1.398)		(0.736)	(0.601)	(0.925)	(1.099)	(0.522)
Obs.	7560	4256	3640	3248		3192	1512	1400	1288	5264
Panel B. Period 2 (Jan. 2002 – Aug. 2014)										
Post × School	0.088	-0.553	0.772	1.234		-0.799	-0.091	-1.380	-0.186	-0.197
	(0.209)	(0.548)	(0.491)	(1.263)		(0.735)	(0.389)	(0.952)	(1.096)	(0.496)
Obs.	8120	4592	3920	3584		3248	1736	1848	1344	5656
Panel C. Period 3 (Jan. 2002 – Dec. 2020)										
Post × School	-0.211	-0.704*	0.304	-0.007		-0.632	-0.282	-1.728*	-0.196	-0.282
	(0.244)	(0.403)	(0.525)	(1.299)		(0.640)	(0.584)	(0.936)	(1.060)	(0.393)
Obs.	12264	7560	6328	5040		3808	3136	2744	1904	9072

Notes: Each panel reports DID estimates from Poisson regression with Mundlak approach. (see Wooldridge, 2021). Columns indicate different outcome variables. 20–44 years (%), 45–64 years (%), 65 years and over (%), male (%), middle school students (%), high school students (%), unemployment rate (%), time fixed effects, the average of the covariates over the time periods, and constant are included in all models. We do not have enough variation in Protozoa. Standard errors clustered at the region level are in parentheses. Some estimates and standard errors are not calculated due to convergence issues. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.1.10. DID estimates of the EWS on the number of outbreaks of food-borne illness per million (with differential linear time trends): Poisson regression with Mundlak approach

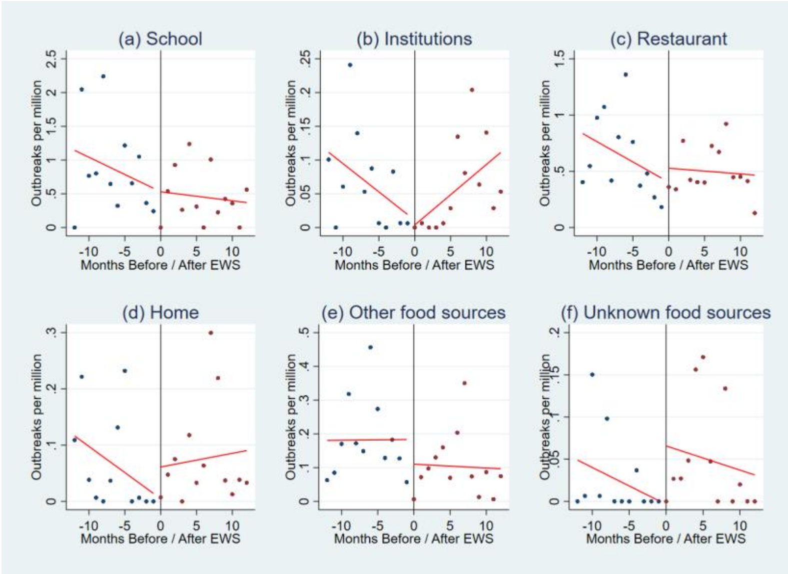
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Total	Norovirus	E. coli	Salmonella	Vibrio	S. aureus	C. jejuni	Perfringens	B. cereus	Unknown
Panel A. Period 1 (Jan. 2002 – Oct. 2013)										
Post × School	-0.001	0.705	-0.601	-1.621**		-0.601	-1.638	0.845	1.000	0.482
	(0.306)	(0.551)	(0.781)	(0.698)		(1.028)	(1.364)	(1.375)	(1.377)	(0.508)
Obs.	7560	4256	3640	3248		3192	1568	1400	1288	5264
Panel B. Period 2 (Jan. 2002 – Aug. 2014)										
Post × School	-0.004	0.629	-0.531	-1.662***		-0.668	-0.358	0.437	1.243	0.414
	(0.286)	(0.571)	(0.740)	(0.587)		(0.993)	(1.007)	(1.244)	(1.429)	(0.519)
Obs.	8120	4592	3920	3584		3248	1792	1848	1344	5656
Panel C. Period 3 (Jan. 2002 – Dec. 2020)										
Post × School	0.397*	0.902**	0.430	-1.151		-0.596	-0.083	0.973	0.885	0.113
	(0.234)	(0.386)	(0.488)	(0.768)		(0.936)	(0.480)	(0.846)	(1.031)	(0.484)
Obs.	12264	7560	6328	5040		3808	3192	2744	1904	9072

Notes: Each panel reports DID estimates from Poisson regression with Mundlak approach. (see Wooldridge, 2021). Columns indicate different outcome variables. 20–44 years (%), 45–64 years (%), 65 years and over (%), male (%), middle school students (%), high school students (%), unemployment rate (%), time fixed effects, the average of the covariates over the time periods, and constant are included in all models. We do not have enough variation in Protozoa. Standard errors clustered at the region level are in parentheses. Some estimates and standard errors are not calculated due to convergence issues. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

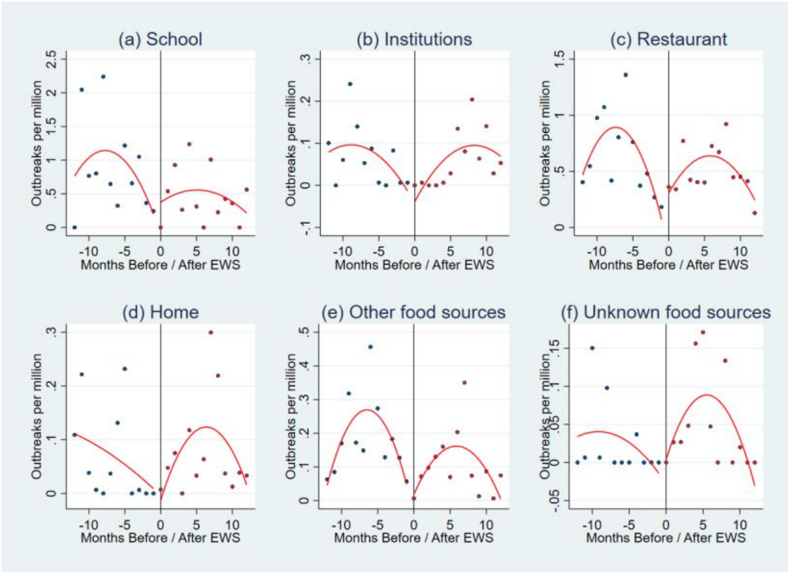
Table A.1.11. DID estimates of the EWS on the number of cases of food-borne illness per million (with differential linear time trends): Poisson regression with Mundlak approach

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Total	Norovirus	E. coli	Salmonella	Vibrio	S. aureus	C. jejuni	Perfringens	B. cereus	Unknown
Panel A. Period 1 (Jan. 2002 – Oct. 2013)										
Post × School	-0.312	1.400	-0.914	-1.181		-0.956	-5.763***	-0.181	2.694	-0.254
	(0.479)	(0.933)	(1.145)	(0.852)		(1.256)	(2.110)	(2.097)	(1.649)	(0.949)
Obs.	7560	4256	3640	3248		3192	1512	1400	1288	5264
Panel B. Period 2 (Jan. 2002 – Aug. 2014)										
Post × School	-0.401	1.149	-0.918	-2.484**		-0.980	-2.536	-0.411	2.708	-0.267
	(0.433)	(0.915)	(1.143)	(1.067)		(1.242)	(1.741)	(1.473)	(1.650)	(0.916)
Obs.	8120	4592	3920	3584		3248	1736	1848	1344	5656
Panel C. Period 3 (Jan. 2002 – Dec. 2020)										
Post × School	0.348	0.149	0.569	1.753		-1.088	0.188	0.134	0.856	-0.181
	(0.459)	(0.780)	(0.612)	(1.768)		(1.070)	(0.556)	(1.246)	(1.145)	(0.733)
Obs.	12264	7560	6328	5040		3808	3136	2744	1904	9072

Notes: Each panel reports DID estimates from Poisson regression with Mundlak approach. (see Wooldridge, 2021). Columns indicate different outcome variables. 20–44 years (%), 45–64 years (%), 65 years and over (%), male (%), middle school students (%), high school students (%), unemployment rate (%), time fixed effects, the average of the covariates over the time periods, and constant are included in all models. We do not have enough variation in Protozoa. Standard errors clustered at the region level are in parentheses. Some estimates and standard errors are not calculated due to convergence issues. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.



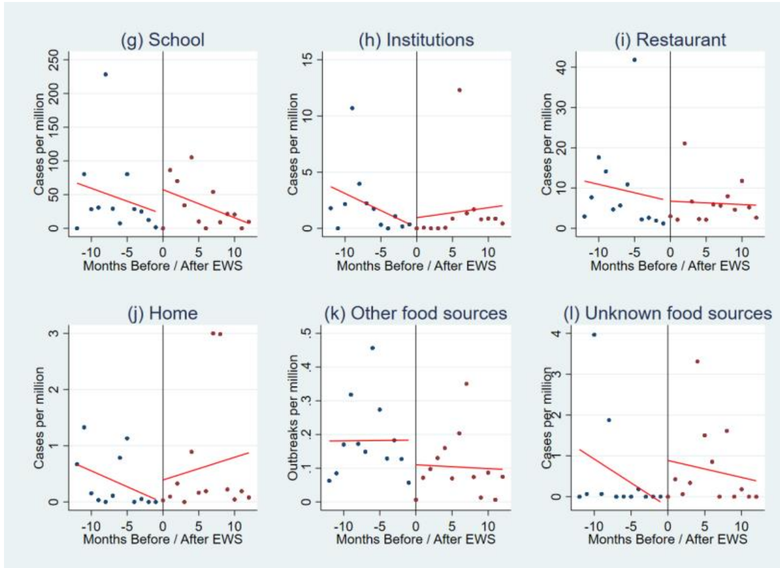
1. Linear fit



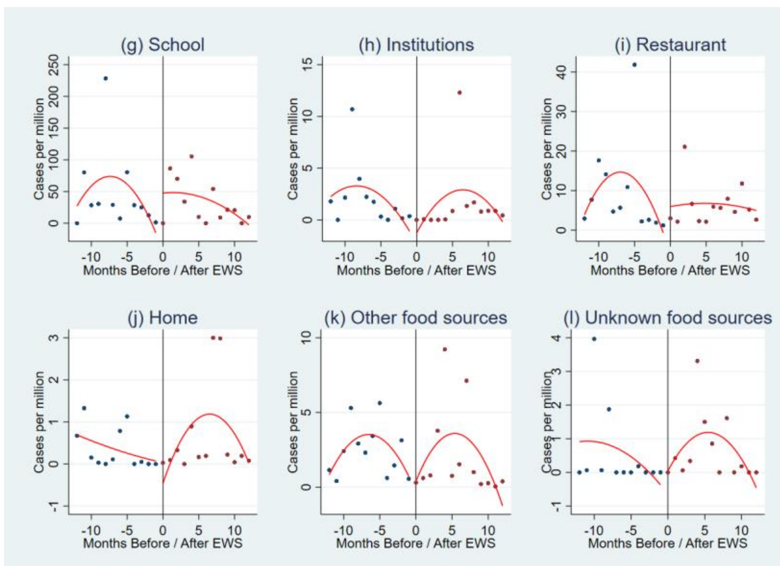
2. Quadratic fit

Figure A.1.1. Discontinuity of the outbreaks of food-borne illness (bandwidth is 12 months)

Notes: These figures illustrate trends in the average total outbreaks of food-borne illness per million. The horizontal axis is the month relative to the introduction of the EWS (i.e., Feb. 2008). The bandwidth is 12 months (i.e., from Feb. 2007 to Feb. 2009). Month fixed effects are not controlled in Figure A.1.1.



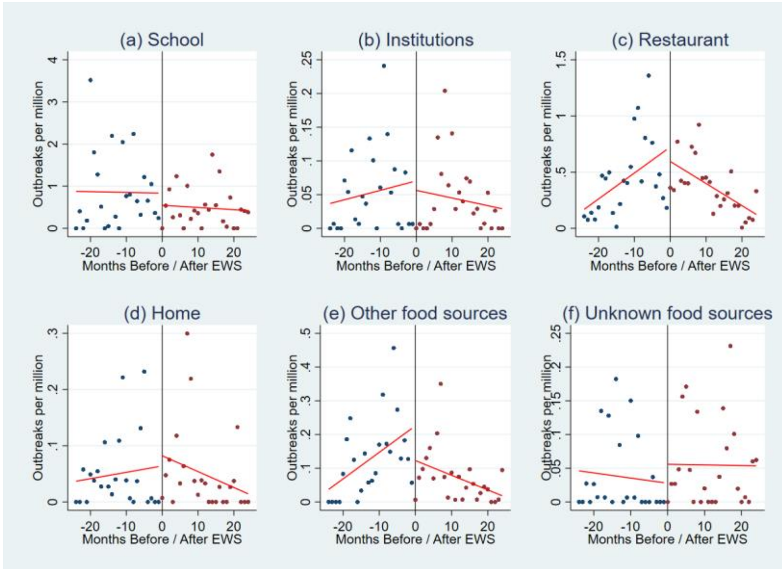
1. Linear fit



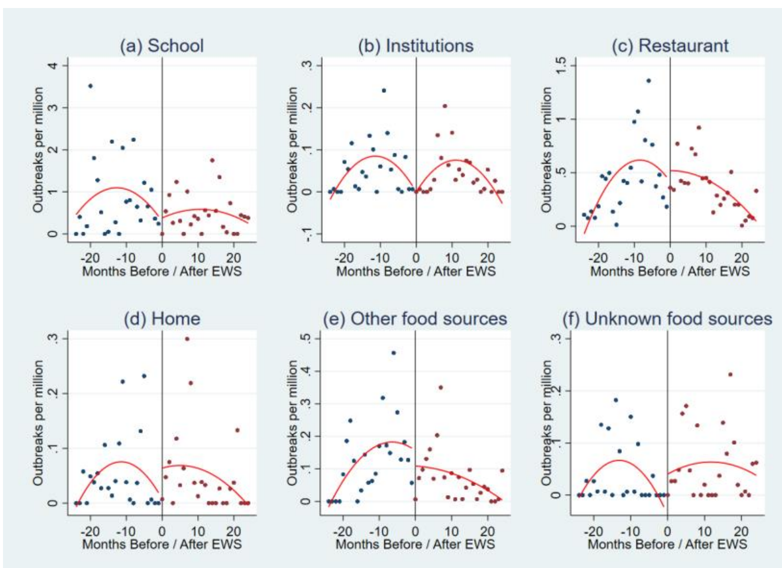
2. Quadratic fit

Figure A.1.2. Discontinuity of the cases of food-borne illness (bandwidth is 12 months)

Notes: These figures illustrate trends in the average total cases of food-borne illness per million. The horizontal axis is the month relative to the introduction of the EWS (i.e., Feb. 2008). The bandwidth is 12 months (i.e., from Feb. 2007 to Feb. 2009). Month fixed effects are not controlled in Figure A.1.2.



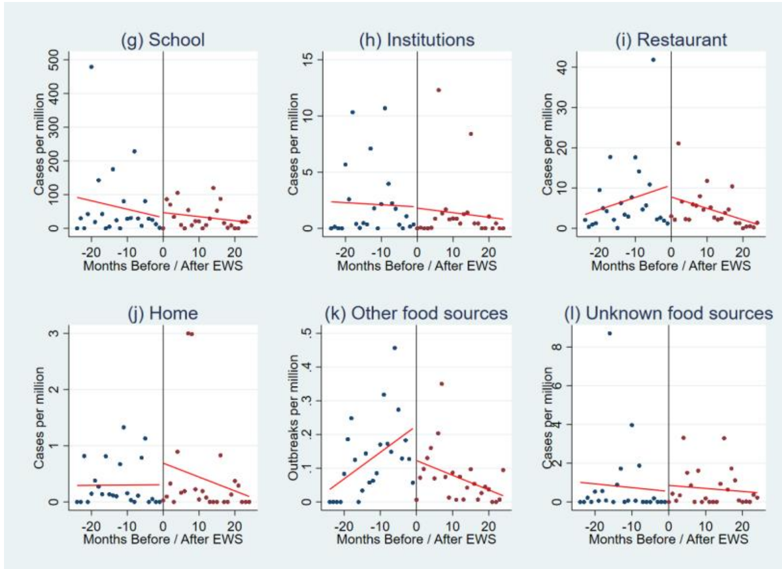
1. Linear fit



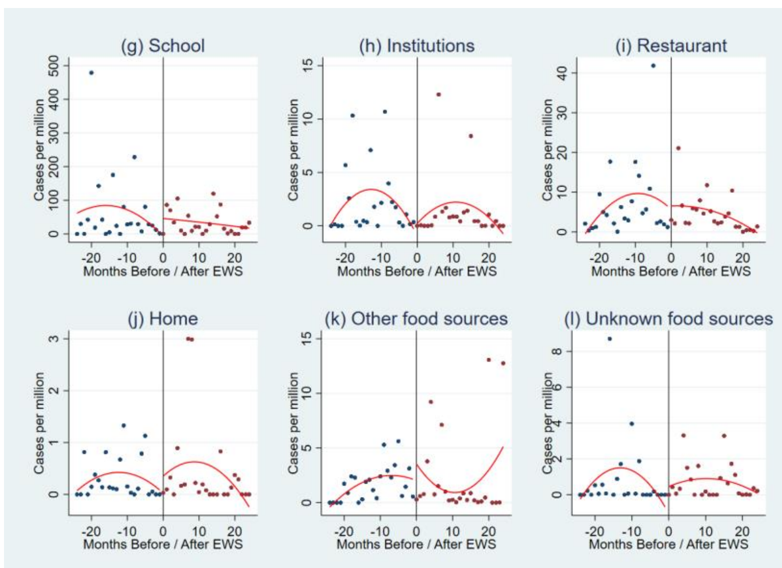
2. Quadratic fit

Figure A.1.3. Discontinuity of the outbreaks of food-borne illness (bandwidth is 24 months)

Notes: These figures illustrate trends in the average total outbreaks of food-borne illness per million. The horizontal axis is the month relative to the introduction of the EWS (i.e., Feb. 2008). The bandwidth is 24 months (i.e., from Feb. 2006 to Feb. 2010). Month fixed effects are not controlled in Figure A.1.3.



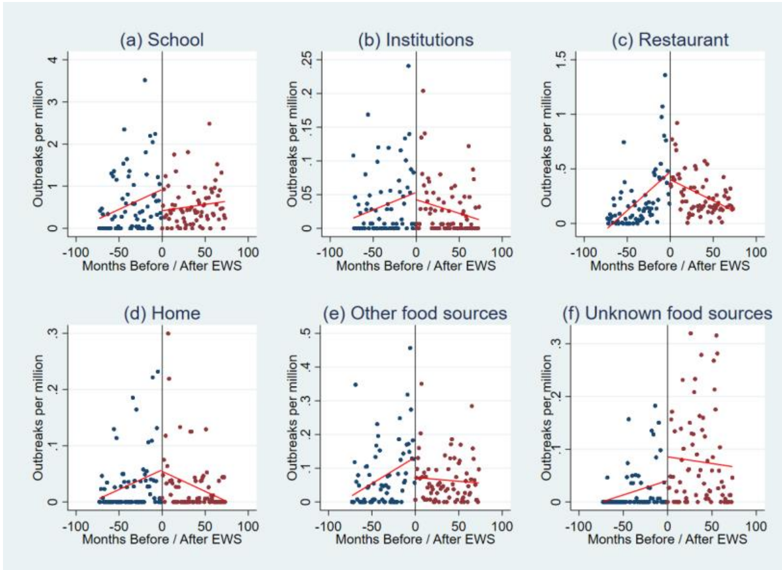
1. Linear fit



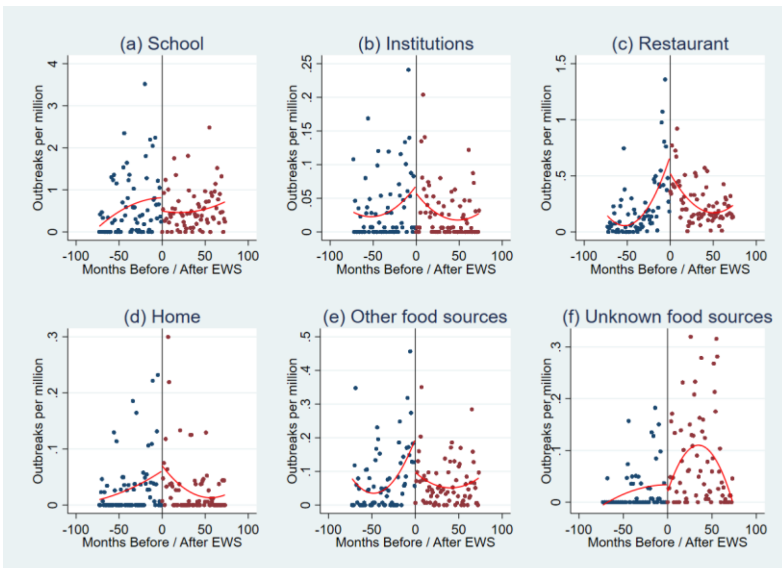
2. Quadratic fit

Figure A.1.4. Discontinuity of the cases of food-borne illness (bandwidth is 24 months)

Notes: These figures illustrate trends in the average total cases of food-borne illness per million. The horizontal axis is the month relative to the introduction of the EWS (i.e., Feb. 2008). The bandwidth is 24 months (i.e., from Feb. 2006 to Feb. 2010). Month fixed effects are not controlled in Figure A.1.4.



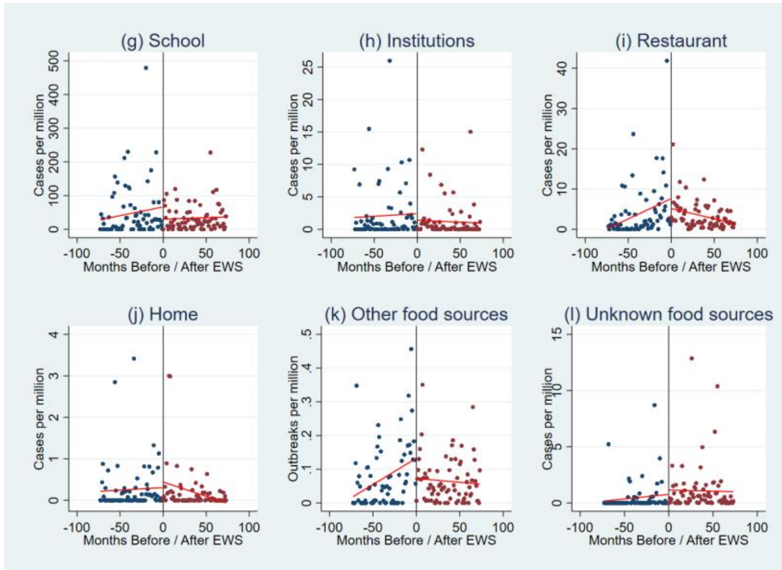
1. Linear fit



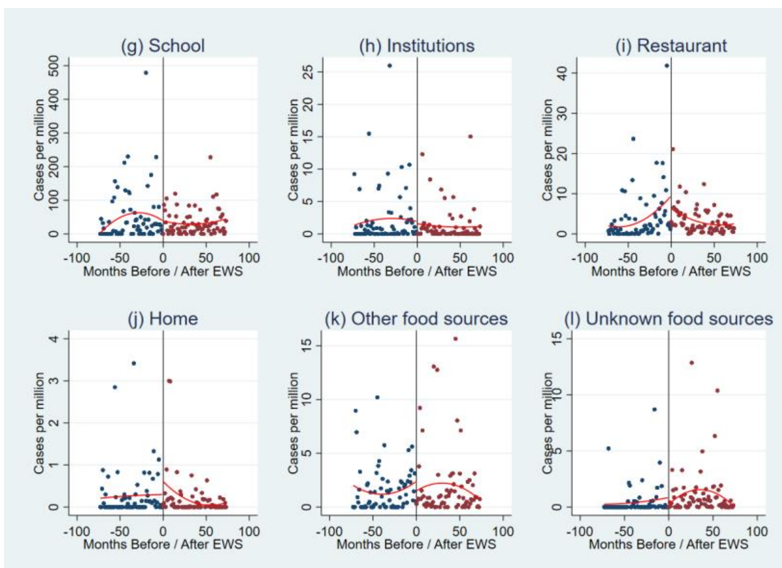
2. Quadratic fit

Figure A.1.5. Discontinuity of the outbreaks of food-borne illness (bandwidth is 73 months)

Notes: These figures illustrate trends in the average total outbreaks of food-borne illness per million. The horizontal axis is the month relative to the introduction of the EWS (i.e., Feb. 2008). The bandwidth is 73 months (i.e., from Jan. 2002 to Mar. 2014). Month fixed effects are not controlled in Figure A.1.5.



1. Linear fit



2. Quadratic fit

Figure A.1.6. Discontinuity of the cases of food-borne illness (bandwidth is 73 months)

Notes: These figures illustrate trends in the average total cases of food-borne illness per million. The horizontal axis is the month relative to the introduction of the EWS (i.e., Feb. 2008). The bandwidth is 73 months (i.e., from Jan. 2002 to Mar. 2014). Month fixed effects are not controlled in Figure A.1.6.

APPENDIX B

CHAPTER 2 APPENDIX

B.1. Supporting tables and figures

Table B.1.1. Event study: The joint significance tests for parallel trends assumption in the pre-treatment period (i.e., January 2002 – November 2009)

	Total	Norovirus	E. coli	Salmonella	Vibrio	S. aureus	C. jejuni	C. Perfringens	B. cereus	Other bacteria	Unknown
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Panel A: Dependent variable is outbreaks per million people											
P-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.825	0.347	0.593	0.000
Panel B: Dependent variable is cases per million people											
P-value	0.000	0.000	0.000	0.000	0.000	0.000	0.695	0.746	0.174	0.686	0.000

Notes: We formally tested that all the event study lead coefficients from Eq. (2.2) are simultaneously zero.

Table B.1.2. The effects of the Seoul City restaurant sanitary grades: Poisson regression

Dependent Variables	Total outbreaks				Total cases			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: DDD estimates without allowing group-specific time trends								
Seoul × Rest. × Post	-0.240	-0.240	-0.240	-0.240	0.132	0.132	0.132	0.132
	(0.209)	(0.216)	(0.213)	(0.215)	(0.577)	(0.580)	(0.577)	(0.577)
Obs.	7616	7616	7392	7392	7616	7616	7392	7392
Panel B: DDD estimates with allowing group-specific time trends								
Seoul × Rest. × Post	-1.314***	-1.011***	-0.320	-0.386	-0.500	-0.323	-0.235	-0.663
	(0.209)	(0.244)	(0.237)	(0.331)	(0.577)	(0.575)	(0.568)	(0.485)
Obs.	7616	7616	7392	7392	7616	7616	7392	7392
Seoul	Y	Y	N	N	Y	Y	N	N
Restaurant	Y	Y	N	N	Y	Y	N	N
Post	Y	Y	N	N	Y	Y	N	N
Covariates	N	Y	N	Y	N	Y	N	Y
Region FE	N	N	Y	Y	N	N	Y	Y
Source FE	N	N	Y	Y	N	N	Y	Y
Time FE	N	N	Y	Y	N	N	Y	Y

Notes: Column (4) of Panel A reports DDD estimates from Eq. (2.1). Column (8) of Panel B reports DDD estimates from Eq. (2.3). Seoul × Restaurant dummy, Seoul × Post dummy, Restaurant × Post dummy, and constant are included in all models. Seoul × Restaurant × Time is included in all models in Panel B. Standard errors in parentheses are clustered at region-source level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table B.1.3. Shortening the time window of sample: OLS regression

Dependent Variables	Dec.2008 - Nov. 2010 (2 years)		Dec.2007 - Nov. 2011 (4 years)		Dec.2006 - Nov. 2012 (6 years)	
	Outbreaks	Cases	Outbreaks	Cases	Outbreaks	Cases
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: DDD estimates without allowing group-specific time trends						
Seoul × Rest. × Post	-0.027	0.668	0.039	1.166	0.130**	2.860**
	(0.059)	(1.088)	(0.056)	(1.335)	(0.057)	(1.388)
Pre-Period Outcome Mean	0.242	3.297	0.308	3.885	0.338	4.619
% Change from Mean	-11%	20%	13%	30%	38%	62%
Obs.	1344	1344	2688	2688	4032	4032
Panel B: DDD estimates with allowing group-specific time trends						
Seoul × Rest. × Post	-0.073	1.426	0.064	1.125	0.148*	3.038
	(0.073)	(1.716)	(0.070)	(2.258)	(0.076)	(2.598)
Pre-Period Outcome Mean	0.242	3.297	0.308	3.885	0.338	4.619
% Change from Mean	-30%	43%	21%	29%	44%	66%
Obs.	1344	1344	2688	2688	4032	4032

Notes: Panel A reports DDD estimates from Eq. (2.1). Panel B reports DDD estimates from Eq. (2.3). Standard errors in parentheses are clustered at region-source level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table B.1.4. Heterogeneity analysis: Poisson regression

Dependent Variables	Outbreaks					
	Norovirus	E. coli	Salmonella	Vibrio	S. aureus	Unknown
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: DDD estimates without allowing group-specific time trends						
Seoul × Restaurant × Post	0.124	-3.072***	-0.646	0.102	18.898***	0.414
	(0.461)	(0.941)	(1.257)	(3.844)	(0.913)	(0.388)
Obs.	3752	3472	4592	2848	7616	6776
Panel B: DDD estimates with allowing group-specific time trends						
Seoul × Restaurant × Post	0.551	-2.677**	-0.456	1.527	15.940***	0.108
	(0.780)	(1.081)	(1.490)	(3.864)	(0.963)	(0.491)
Obs.	3752	3472	4592	2848	7616	6776

Notes: Panel A reports DDD estimates from Eq. (2.1). Panel B reports DDD estimates from Eq. (2.3). Standard errors in parentheses are clustered at region-source level. We do not have enough variation in C. jejuni, C. Perfringens, B. cereus, and other bacteria. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table B.1.5. Heterogeneity analysis with shortening the time window of sample (without allowing group-specific time trends.): OLS regression

Dependent Variables	Outbreaks per million people								
	Norovirus	E. coli	Salmonella	Vibrio	S. aureus	C. jejuni	C. Perfringens	B. cereus	Unknown
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Panel A: Dec.2008 – Nov. 2010 (2 years)									
Seoul × Restaurant × Post	0.007	-0.006	0.011	-0.012	-0.034***	-0.004	0.002	-0.017	0.032
	(0.011)	(0.011)	(0.021)	(0.012)	(0.008)	(0.006)	(0.007)	(0.017)	(0.037)
Pre-Period Outcome Mean	0.022	0.021	0.032	0.017	0.018	0.002	0.005	0.011	0.109
% Change from Mean	32%	-29%	35%	-69%	-190%	-215%	38%	-154%	29%
Obs.	1344	1344	1344	1344	1344	1344	1344	1344	1344
Panel B: Dec.2007 – Nov. 2011 (4 years)									
Seoul × Restaurant × Post	0.046**	-0.010	-0.001	0.006	-0.019**	-0.009**	-0.011**	-0.004	0.042
	(0.018)	(0.010)	(0.012)	(0.008)	(0.008)	(0.004)	(0.005)	(0.009)	(0.031)
Pre-Period Outcome Mean	0.038	0.025	0.025	0.021	0.019	0.006	0.008	0.014	0.149
% Change from Mean	120%	-40%	-4%	29%	-103%	-163%	-134%	-29%	28%
Obs.	2688	2688	2688	2688	2688	2688	2688	2688	2688
Panel C: Dec.2006 – Nov. 2012 (6 years)									
Seoul × Restaurant × Post	0.042**	0.007	-0.001	0.015*	-0.007	-0.003	-0.005	-0.006	0.087**
	(0.016)	(0.014)	(0.010)	(0.009)	(0.007)	(0.003)	(0.004)	(0.006)	(0.035)
Pre-Period Outcome Mean	0.043	0.029	0.028	0.028	0.021	0.005	0.007	0.011	0.161
% Change from Mean	97%	24%	-4%	54%	-33%	-63%	-71%	-54%	54%
Obs.	4032	4032	4032	4032	4032	4032	4032	4032	4032

Notes: The estimates report coefficients from Eq. (2.1). We do not have enough variation in other bacteria. Standard errors in parentheses are clustered at region-source level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table B.1.6. Heterogeneity analysis with shortening the time window of sample (with allowing group-specific time trends.): OLS regression

Dependent Variables	Outbreaks per million people								
	Norovirus	E. coli	Salmonella	Vibrio	S. aureus	C. jejuni	C. Perfringens	B. cereus	Unknown
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Panel A: Dec.2008 – Nov. 2010 (2 years)									
Seoul × Restaurant × Post	-0.027	-0.036***	0.014	0.010	-0.037***	-0.005	-0.005	-0.026**	0.042
	(0.019)	(0.008)	(0.027)	(0.013)	(0.007)	(0.007)	(0.009)	(0.012)	(0.052)
Pre-Period Outcome Mean	0.022	0.021	0.032	0.017	0.018	0.002	0.005	0.011	0.109
% Change from Mean	-122%	-175%	44%	57%	-207%	-269%	-96%	-236%	39%
Obs.	1344	1344	1344	1344	1344	1344	1344	1344	1344
Panel B: Dec.2007 – Nov. 2011 (4 years)									
Seoul × Restaurant × Post	0.043**	-0.004	0.020	0.005	-0.030***	-0.005	-0.011*	-0.008	0.066*
	(0.019)	(0.012)	(0.015)	(0.008)	(0.010)	(0.004)	(0.006)	(0.011)	(0.034)
Pre-Period Outcome Mean	0.038	0.025	0.025	0.021	0.019	0.006	0.008	0.014	0.149
% Change from Mean	112%	-16%	81%	24%	-162%	-90%	-134%	-58%	44%
Obs.	2688	2688	2688	2688	2688	2688	2688	2688	2688
Panel C: Dec.2006 – Nov. 2012 (6 years)									
Seoul × Restaurant × Post	0.075***	0.019	-0.002	0.013	-0.006	-0.008**	-0.009	-0.007	0.077*
	(0.021)	(0.019)	(0.013)	(0.010)	(0.008)	(0.004)	(0.005)	(0.008)	(0.045)
Pre-Period Outcome Mean	0.043	0.029	0.028	0.028	0.021	0.005	0.007	0.011	0.161
% Change from Mean	173%	65%	-7%	47%	-28%	-168%	-127%	-63%	48%
Obs.	4032	4032	4032	4032	4032	4032	4032	4032	4032

Notes: The estimates report coefficients from Eq. (2.3). We do not have enough variation in other bacteria. Standard errors in parentheses are clustered at region-source level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table B.1.7. Heterogeneity analysis: Poisson regression

Dependent Variables	Cases					
	Norovirus	E. coli	Salmonella	Vibrio	S. aureus	Unknown
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: DDD estimates without allowing group-specific time trends						
Seoul × Restaurant × Post	0.895	-2.559*	2.316*	2.038	15.352***	-0.491
	(0.674)	(1.525)	(1.401)	(3.538)	(1.015)	(0.506)
Obs.	3752	3472	4592	2848	7616	6776
Panel B: DDD estimates with allowing group-specific time trends						
Seoul × Restaurant × Post	0.304	-7.772***	2.332	1.504	16.315***	-0.981
	(1.254)	(2.078)	(1.555)	(3.936)	(1.018)	(0.798)
Obs.	3752	3472	4592	2848	7616	6776

Notes: Panel A reports DDD estimates from Eq. (2.1). Panel B reports DDD estimates from Eq. (2.3). Standard errors in parentheses are clustered at region-source level. We do not have enough variation in C. jejuni, C. Perfringens, B. cereus, and Other bacteria. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table B.1.8. Heterogeneity analysis with shortening the time window of sample (without allowing group-specific time trends.): OLS regression

Dependent Variables	Cases per million people								
	Norovirus	E. coli	Salmonella	Vibrio	S. aureus	C. jejuni	C. Perfringens	B. cereus	Unknown
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Panel A: Dec.2008 – Nov. 2010 (2 years)									
Seoul × Restaurant × Post	0.997**	-0.974***	0.493	-0.488	-0.471	-0.371	0.746	-0.068	0.825**
	(0.450)	(0.277)	(0.642)	(0.426)	(0.453)	(0.262)	(0.505)	(0.095)	(0.355)
Pre-Period Outcome Mean	0.308	0.306	0.806	0.329	0.237	0.052	0.254	0.045	0.944
% Change from Mean	323%	-318%	61%	-149%	-198%	-708%	293%	-152%	87%
Obs.	1344	1344	1344	1344	1344	1344	1344	1344	1344
Panel B: Dec.2007 – Nov. 2011 (4 years)									
Seoul × Restaurant × Post	1.515*	-1.035***	0.700*	0.016	-0.106	-0.201	0.068	0.020	0.218
	(0.778)	(0.338)	(0.414)	(0.226)	(0.243)	(0.147)	(0.286)	(0.063)	(0.226)
Pre-Period Outcome Mean	0.740	0.374	0.745	0.303	0.228	0.083	0.277	0.077	1.047
% Change from Mean	205%	-277%	94%	5%	-47%	-243%	25%	26%	21%
Obs.	2688	2688	2688	2688	2688	2688	2688	2688	2688
Panel C: Dec.2006 – Nov. 2012 (6 years)									
Seoul × Restaurant × Post	1.345**	0.100	0.640	0.100	-0.031	-0.013	0.060	0.007	0.654***
	(0.647)	(0.795)	(0.450)	(0.200)	(0.192)	(0.099)	(0.204)	(0.042)	(0.194)
Pre-Period Outcome Mean	0.762	0.715	1.038	0.402	0.241	0.102	0.196	0.063	1.077
% Change from Mean	176%	14%	62%	25%	-13%	-13%	31%	11%	61%
Obs.	4032	4032	4032	4032	4032	4032	4032	4032	4032

Notes: The estimates report coefficients from Eq. (2.1). We do not have enough variation in other bacteria. Standard errors in parentheses are clustered at region-source level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table B.1.9. Heterogeneity analysis with shortening the time window of sample (with allowing group-specific time trends.): OLS regression

Dependent Variables	Cases per million people								
	Norovirus	E. coli	Salmonella	Vibrio	S. aureus	C. jejuni	C. Perfringens	B. cereus	Unknown
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Panel A: Dec.2008 – Nov. 2010 (2 years)									
Seoul × Restaurant × Post	1.303*	-0.934*	1.133	-0.067	-0.012	-0.732	0.502	-0.141**	0.377
	(0.731)	(0.527)	(0.930)	(0.371)	(0.274)	(0.640)	(0.449)	(0.068)	(0.531)
Pre-Period Outcome Mean	0.308	0.306	0.806	0.329	0.237	0.052	0.254	0.045	0.944
% Change from Mean	423%	-305%	141%	-20%	-5%	-1397%	197%	-315%	40%
Obs.	1344	1344	1344	1344	1344	1344	1344	1344	1344
Panel B: Dec.2007 – Nov. 2011 (4 years)									
Seoul × Restaurant × Post	2.245	-1.403***	0.546	-0.111	-0.016	-0.289	0.125	0.020	0.139
	(1.394)	(0.414)	(0.655)	(0.256)	(0.221)	(0.209)	(0.358)	(0.099)	(0.455)
Pre-Period Outcome Mean	0.740	0.374	0.745	0.303	0.228	0.083	0.277	0.077	1.047
% Change from Mean	303%	-376%	73%	-37%	-7%	-349%	45%	26%	13%
Obs.	2688	2688	2688	2688	2688	2688	2688	2688	2688
Panel C: Dec.2006 – Nov. 2012 (6 years)									
Seoul × Restaurant × Post	1.826	0.849	0.368	0.029	0.064	-0.138	0.038	0.022	-0.032
	(1.157)	(1.684)	(0.548)	(0.227)	(0.187)	(0.105)	(0.296)	(0.068)	(0.286)
Pre-Period Outcome Mean	0.762	0.715	1.038	0.402	0.241	0.102	0.196	0.063	1.077
% Change from Mean	240%	119%	35%	7%	27%	-135%	19%	35%	-3%
Obs.	4032	4032	4032	4032	4032	4032	4032	4032	4032

Notes: The estimates report coefficients from Eq. (2.3). We do not have enough variation in other bacteria. Standard errors in parentheses are clustered at region-source level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

APPENDIX C

CHAPTER 3 APPENDIX

C.1. Supporting tables and figures

Table C.1.1. Information on NutriPlus spending

County	Year started Nutriplus	Year missed information	Year collected information
Jung-gu*	2008	2008-2011	2012-2020
Seongdong-gu*	2006	2006-2010	2011-2020
Gwangjin-gu*	2008	2008-2011	2012-2020
Dongdaemun-gu*	2008	2008-2010	2011-2020
Seongbuk-gu*	2005	2005-2008	2009-2020
Gangbuk-gu*	2007	2007-2008	2009-2020
Dobong-gu*	2008	2008-2010	2011-2020
Seodaemun-gu*	2008	2008-2011	2012-2020
Mapo-gu*	2008	2008-2010	2011-2020
Yangcheon-gu*	2008	2008-2013	2014-2020
Guro-gu*	2008	2008-2011	2012-2020
Geumcheon-gu*	2009	2009-2012	2013-2020
Yeongdeungpo-gu*	2008	2008-2010	2011-2020
Dongjak-gu*	2008	2008-2012	2013-2020
Gwanak-gu*	2008	2008-2009	2010-2020
Seocho-gu*	2008	2008-2009	2010-2020
Gangnam-gu*	2008	2008-2012, 2018	2013-2017, 2019, 2020
Songpa-gu*	2008	2008-2012, 2014	2013, 2015-2020
Gangdong-gu*	2006	2006-2014	2015-2020
Jongno-gu	2008	2008-2020	-
Yongsan-gu	2008	2008-2020	-
Jungnang-gu	2008	2008-2020	-
Nowon-gu	2008	2008-2020	-
Eunpyeong-gu	2008	2008-2020	-
Gangseo-gu	2008	2008-2020	-

Notes: Nineteen counties ultimately included in the main sample are denoted with a *.

Table C.1.2. Information on neighboring counties

County	Neighboring counties
Jung-gu*	Jongno-gu, Seongdong-gu*, Yongsan-gu, Mapo-gu*, and Seodaemun-gu*
Seongdong-gu*	Dongdaemun-gu*, Gwangjin-gu*, Yongsan-gu, and Jung-gu*
Gwangjin-gu*	Seongdong-gu*, Dongdaemun-gu*, and Jungnang-gu
Dongdaemun-gu*	Seongbuk-gu*, Jungnang-gu, Gwangjin-gu*, Seongdong-gu*, and Jongno-gu
Seongbuk-gu*	Gangbuk-gu*, Nowon-gu, Jungnang-gu, Dongdaemun-gu*, and Jongno-gu
Gangbuk-gu*	Dobong-gu*, Nowon-gu, and Seongbuk-gu*
Dobong-gu*	Gangbuk-gu* and Nowon-gu
Seodaemun-gu*	Eunpyeong-gu, Jongno-gu, Jung-gu*, and Mapo-gu*
Mapo-gu*	Seodaemun-gu*, Jung-gu*, and Yongsan-gu
Yangcheon-gu*	Gangseo-gu, Yeongdeungpo-gu*, and Guro-gu*
Guro-gu*	Yangcheon-gu*, Yeongdeungpo-gu*, Geumcheon-gu*, and Gwanak-gu**
Geumcheon-gu*	Guro-gu* and Gwanak-gu*
Yeongdeungpo-gu*	Gangseo-gu, Yangcheon-gu*, Guro-gu*, and Dongjak-gu*
Dongjak-gu*	Yeongdeungpo-gu*, Gwanak-gu*, and Seocho-gu*
Gwanak-gu*	Guro-gu*, Geumcheon-gu*, Dongjak-gu*, and Seocho-gu*
Seocho-gu*	Gwanak-gu*, Dongjak-gu*, and Gangnam-gu*
Gangnam-gu*	Seocho-gu* and Songpa-gu*
Songpa-gu*	Gangnam-gu* and Gangdong-gu*
Gangdong-gu*	Songpa-gu*
Jongno-gu	Seongbuk-gu*, Dongdaemun-gu*, Jung-gu*, Seodaemun-gu*, and Eunpyeong-gu
Yongsan-gu	Mapo-gu*, Jung-gu*, and Seongdong-gu*
Jungnang-gu	Nowon-gu, Seongbuk-gu*, Dongdaemun-gu*, and Gwangjin-gu*
Nowon-gu	Dobong-gu*, Gangbuk-gu*, Seongbuk-gu*, and Jungnang-gu
Eunpyeong-gu	Mapo-gu*, Seodaemun-gu*, and Jongno-gu
Gangseo-gu	Yangcheon-gu* and Yeongdeungpo-gu*

Notes: Nineteen counties ultimately included in the main sample are denoted with a *.

Table C.1.3. Unweighted descriptive statistics (N = 575)

Variable	Mean	Std. Dev.	Min	Max
NutriPlus spending per capita (1,000 KRW)	0.333	0.423	0.000	2.693
NutriPlus spending per capita in NBR (1,000 KRW)	0.360	0.413	0.000	2.006
Real GRDP per capita (1,000 KRW)	29325.000	9739.396	13402.000	45859.000
Mother's age (%)				
age $<$ 20	0.324	0.203	0.000	1.487
20 $\leq$ age $<$ 35	80.221	10.237	54.751	94.536
age $\geq$ 35	19.454	10.283	5.045	45.199
College (%)	69.950	17.355	24.261	96.924
Male newborn (%)	51.545	1.012	47.361	54.406
Gestation				
Gestational age (weeks)	38.976	0.253	38.426	39.535
Premature birth (%)	4.111	0.656	2.444	6.216
Birth weight (in grams)	3254.544	25.379	3188.528	3317.362
Birth weight (%)				
Very low birth weight (birth weight $<$ 1500g)	0.369	0.150	0.000	0.933
Low birth weight (birth weight $<$ 2500g)	3.411	0.527	2.005	5.146
Normal birth weight (2500g $\leq$ birth weight $\leq$ 4500g)	96.337	0.502	94.490	97.719
High birth weight (birth weight $>$ 4500g)	0.252	0.130	0.000	1.084

Notes: "NutriPlus spending per capita in NBR" indicates the average NutriPlus spending per capita in neighboring counties.

Table C.1.4. First-stage regression

	OLS					
	(1)	(2)	(3)	(4)	(5)	(6)
NBR (1,000 KRW)	0.618***	0.193*	0.656***	0.621***	0.202*	0.668***
	(0.089)	(0.099)	(0.092)	(0.093)	(0.097)	(0.101)
Age < 20 (%)	0.096	0.063	0.079	0.094	0.073	0.066
	(0.099)	(0.140)	(0.064)	(0.103)	(0.147)	(0.063)
Age ≥ 35 (%)	0.009*	0.042*	-0.002	0.009**	0.043*	-0.003
	(0.005)	(0.023)	(0.009)	(0.004)	(0.023)	(0.008)
College (%)	0.004	0.005	0.002	0.004	0.005	0.001
	(0.007)	(0.007)	(0.004)	(0.007)	(0.008)	(0.004)
Male newborn (%)	-0.002	0.001	0.001	-0.002	0.000	0.002
	(0.012)	(0.013)	(0.008)	(0.012)	(0.012)	(0.008)
Gestational age	-0.077	0.189	-0.246			
(weeks)	(0.340)	(0.276)	(0.244)			
Log(grdpc)	-0.206	-0.811	-0.420	-0.162	-0.988	-0.324
	(0.334)	(0.768)	(0.312)	(0.325)	(0.740)	(0.351)
County fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	No	Yes	No	No	Yes	No
County × Time trends	No	No	Yes	No	No	Yes
Observations	437	437	437	437	437	437
F-test on instrument	37.54	10.10	38.29	34.51	10.05	34.02

Notes: Table C.1.4 reports estimates from Eq. (3.2). The dependent variable is the NutriPlus spending per capita in a domestic county (1,000 KRW). Estimates are weighted using the number of births in the cell. Robust standard errors clustered by county are reported in parentheses. "NBR" indicates the average NutriPlus spending per capita in neighboring counties. "Log(grdpc)" indicates the log of real GRDP per capita. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table C.1.5. Effects of NutriPlus spending on birth weight

	OLS	OLS	OLS	IV	IV	IV
	(1)	(2)	(3)	(4)	(5)	(6)
NutriPlus spending	-8.116**	-1.130	-9.038***	-16.459***	-19.059	-13.134***
	(2.997)	(3.507)	(2.752)	(3.464)	(29.252)	(3.574)
Age < 20 (%)	-20.009***	-9.740***	-10.173***	-16.709***	-8.331**	-9.075***
	(3.457)	(3.222)	(2.694)	(3.707)	(3.826)	(2.478)
Age ≥ 35 (%)	-1.794***	-0.061	-0.704	-1.634***	0.820	-0.717
	(0.252)	(0.968)	(0.535)	(0.250)	(1.461)	(0.487)
College (%)	-0.735***	0.014	-1.377***	-0.608***	0.133	-1.320***
	(0.233)	(0.250)	(0.280)	(0.206)	(0.304)	(0.259)
Male newborn (%)	0.957	0.450	0.202	0.966	0.679	0.143
	(0.599)	(0.741)	(0.519)	(0.618)	(0.716)	(0.486)
Gestational age	51.385***	34.993*	65.720***	48.339***	40.327***	62.314***
(weeks)	(14.013)	(17.060)	(15.261)	(14.695)	(14.507)	(14.139)
Log(grdpc)	60.715***	-25.375	150.786***	55.292***	-41.239	144.685***
	(12.859)	(35.325)	(15.244)	(11.749)	(41.156)	(15.668)
County fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	No	Yes	No	No	Yes	No
County × Time trends	No	No	Yes	No	No	Yes
Constant	675.761	2116.412**	6358.993*	836.599	2035.010**	8557.766***
	(596.012)	(876.239)	(3385.535)	(631.810)	(814.873)	(3171.210)
Observations	358	358	358	354	354	354
R^2	0.841	0.888	0.884	0.837	0.878	0.884

Notes: Dependent variable is the birth weight (in grams). Columns (1)–(3) report the estimates from Eq. (3.1). Columns (4)–(6) report the estimates from Eq. (3.3). Instrumental variable is the average NutriPlus spending per capita in neighboring counties. Estimates are weighted using the number of births in the cell. "Log(grdpc)" indicates the log of real GRDP per capita. Robust standard errors clustered by county are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table C.1.6. Effects of NutriPlus spending on low birth weight (%).

	OLS (1)	OLS (2)	OLS (3)	IV (4)	IV (5)	IV (6)
NutriPlus spending	0.114	0.162	0.156	0.091	-0.157	0.069
	(0.081)	(0.117)	(0.113)	(0.140)	(0.607)	(0.152)
Age < 20 (%)	0.179	0.054	0.086	0.172	0.072	0.111
	(0.145)	(0.157)	(0.186)	(0.145)	(0.136)	(0.175)
Age ≥ 35 (%)	-0.002	-0.017	-0.008	-0.002	0.003	-0.004
	(0.005)	(0.031)	(0.019)	(0.004)	(0.040)	(0.018)
College (%)	-0.003	-0.011*	0.006	-0.003	-0.009	0.008
	(0.005)	(0.006)	(0.010)	(0.004)	(0.006)	(0.009)
Male newborn (%)	-0.002	-0.010	0.004	-0.013	-0.017	-0.007
	(0.029)	(0.027)	(0.029)	(0.028)	(0.028)	(0.028)
Gestational age	-1.127***	-1.058**	-1.346***	-1.143***	-0.981***	-1.501***
(weeks)	(0.366)	(0.368)	(0.437)	(0.356)	(0.361)	(0.415)
Log(grdpc)	0.358	1.054	-0.635	0.377	0.668	-0.624
	(0.348)	(1.030)	(0.688)	(0.312)	(1.098)	(0.652)
County fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	No	Yes	No	No	Yes	No
County × Time trends	No	No	Yes	No	No	Yes
Constant	43.945**	35.494*	43.226	45.158***	36.444*	36.905
	(16.925)	(20.381)	(124.557)	(16.164)	(19.280)	(117.432)
Observations	358	358	358	354	354	354
R^2	0.635	0.675	0.667	0.642	0.670	0.674

Notes: Dependent variable is the fraction of low birth weight (%). Columns (1)–(3) report the estimates from Eq. (3.1). Columns (4)–(6) report the estimates from Eq. (3.3). Instrumental variable is the average NutriPlus spending per capita in neighboring counties. Estimates are weighted using the number of births in the cell. "Log(grdpc)" indicates the log of real GRDP per capita. Robust standard errors clustered by county are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table C.1.7. Effects of NutriPlus spending on very low birth weight (%)

	OLS	OLS	OLS	IV	IV	IV
	(1)	(2)	(3)	(4)	(5)	(6)
NutriPlus spending	-0.009	0.029	-0.023	-0.044	-0.074	-0.050
	(0.019)	(0.029)	(0.023)	(0.027)	(0.127)	(0.031)
Age < 20 (%)	-0.001	-0.017	0.032	0.012	-0.009	0.042
	(0.028)	(0.035)	(0.051)	(0.026)	(0.035)	(0.044)
Age ≥ 35 (%)	-0.002	0.000	-0.004	-0.002	0.004	-0.004
	(0.002)	(0.007)	(0.006)	(0.002)	(0.008)	(0.005)
College (%)	0.000	0.000	-0.001	0.001	0.001	-0.001
	(0.001)	(0.001)	(0.003)	(0.001)	(0.002)	(0.003)
Male newborn (%)	-0.003	-0.003	-0.003	-0.006	-0.004	-0.005
	(0.006)	(0.007)	(0.006)	(0.005)	(0.006)	(0.006)
Gestational age	-0.290**	-0.334**	-0.482***	-0.310**	-0.310**	-0.529***
(weeks)	(0.131)	(0.136)	(0.145)	(0.121)	(0.128)	(0.126)
Log(grdpc)	0.104	-0.013	0.096	0.086	-0.068	0.068
	(0.117)	(0.259)	(0.166)	(0.110)	(0.267)	(0.173)
County fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	No	Yes	No	No	Yes	No
County × Time trends	No	No	Yes	No	No	Yes
Constant	10.867*	13.672*	17.332	11.886**	13.224**	22.494
	(6.040)	(7.260)	(34.263)	(5.557)	(6.732)	(31.951)
Observations	358	358	358	354	354	354
R^2	0.512	0.550	0.546	0.514	0.541	0.550

Notes: Dependent variable is the fraction of very low birth weight (%). Columns (1)–(3) report the estimates from Eq. (3.1). Columns (4)–(6) report the estimates from Eq. (3.3). Instrumental variable is the average NutriPlus spending per capita in neighboring counties. Estimates are weighted using the number of births in the cell. "Log(grdpc)" indicates the log of real GRDP per capita. Robust standard errors clustered by county are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table C.1.8. Effects of NutriPlus spending on high birth weight (%)

	OLS (1)	OLS (2)	OLS (3)	IV (4)	IV (5)	IV (6)
NutriPlus spending	-0.061**	-0.034	-0.041*	-0.081***	-0.085	-0.064**
	(0.022)	(0.030)	(0.024)	(0.023)	(0.113)	(0.025)
Age < 20 (%)	-0.046*	-0.047	-0.070**	-0.039*	-0.043	-0.063**
	(0.023)	(0.040)	(0.029)	(0.023)	(0.035)	(0.029)
Age ≥ 35 (%)	-0.006***	-0.005	-0.002	-0.005***	-0.003	-0.002
	(0.002)	(0.005)	(0.005)	(0.002)	(0.006)	(0.004)
College (%)	-0.006***	-0.005*	-0.005**	-0.005***	-0.005**	-0.004**
	(0.002)	(0.003)	(0.002)	(0.001)	(0.002)	(0.002)
Male newborn (%)	0.001	-0.003	0.000	-0.000	-0.003	-0.001
	(0.004)	(0.005)	(0.004)	(0.004)	(0.004)	(0.004)
Gestational age (weeks)	-0.027	-0.073	0.010	-0.034	-0.059	-0.013
	(0.071)	(0.070)	(0.115)	(0.070)	(0.072)	(0.101)
Log(grdpc)	0.224	0.120	0.278	0.214*	0.074	0.258*
	(0.131)	(0.222)	(0.165)	(0.117)	(0.205)	(0.141)
County fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	No	Yes	No	No	Yes	No
County × Time trends	No	No	Yes	No	No	Yes
Constant	-0.497	2.482	17.922	-0.149	2.313	14.868
	(3.606)	(4.025)	(28.979)	(3.451)	(3.805)	(26.120)
Observations	358	358	358	354	354	354
R^2	0.487	0.531	0.508	0.485	0.524	0.508

Notes: Dependent variable is the fraction of high birth weight (%). Columns (1)–(3) report the estimates from Eq. (3.1). Columns (4)–(6) report the estimates from Eq. (3.3). Instrumental variable is the average NutriPlus spending per capita in neighboring counties. Estimates are weighted using the number of births in the cell. “Log(grdpc)” indicates the log of real GRDP per capita. Robust standard errors clustered by county are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table C.1.9. Effects of NutriPlus spending on normal birth weight (%)

	OLS	OLS	OLS	IV	IV	IV
	(1)	(2)	(3)	(4)	(5)	(6)
NutriPlus spending	-0.053	-0.128	-0.114	-0.010	0.242	-0.005
	(0.092)	(0.127)	(0.127)	(0.149)	(0.656)	(0.162)
Age < 20 (%)	-0.133	-0.008	-0.016	-0.133	-0.029	-0.048
	(0.145)	(0.168)	(0.185)	(0.144)	(0.138)	(0.173)
Age ≥ 35 (%)	0.008	0.022	0.011	0.008*	-0.000	0.006
	(0.006)	(0.033)	(0.019)	(0.004)	(0.042)	(0.018)
College (%)	0.009	0.016*	-0.001	0.009*	0.014**	-0.003
	(0.006)	(0.008)	(0.011)	(0.005)	(0.007)	(0.010)
Male newborn (%)	0.001	0.013	-0.004	0.013	0.020	0.009
	(0.029)	(0.027)	(0.029)	(0.028)	(0.029)	(0.028)
Gestational age	1.154***	1.131***	1.337**	1.177***	1.039***	1.514***
(weeks)	(0.374)	(0.372)	(0.479)	(0.369)	(0.367)	(0.448)
Log(grdpc)	-0.582	-1.173	0.357	-0.591	-0.742	0.366
	(0.424)	(1.110)	(0.744)	(0.381)	(1.171)	(0.691)
County fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	No	Yes	No	No	Yes	No
County × Time trends	No	No	Yes	No	No	Yes
Constant	56.552***	62.023***	38.852	54.991***	61.243***	48.226
	(17.755)	(20.856)	(129.926)	(17.045)	(19.941)	(120.416)
Observations	358	358	358	354	354	354
R^2	0.537	0.587	0.577	0.545	0.577	0.585

Notes: Dependent variable is the fraction of normal birth weight (%). Columns (1)–(3) report the estimates from Eq. (3.1). Columns (4)–(6) report the estimates from Eq. (3.3). Instrumental variable is the average NutriPlus spending per capita in neighboring counties. Estimates are weighted using the number of births in the cell. “Log(grdpc)” indicates the log of real GRDP per capita. Robust standard errors clustered by county are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table C.1.10. Effects of NutriPlus spending on premature birth (%)

	OLS	OLS	OLS	IV	IV	IV
	(1)	(2)	(3)	(4)	(5)	(6)
NutriPlus spending	0.118	0.096	0.152	0.106	-0.467	0.157
	(0.092)	(0.159)	(0.090)	(0.174)	(0.981)	(0.134)
Age < 20 (%)	-0.021	-0.162	-0.018	-0.029	-0.112	-0.027
	(0.119)	(0.094)	(0.144)	(0.109)	(0.143)	(0.137)
Age ≥ 35 (%)	0.001	-0.019	-0.032*	0.001	0.009	-0.029**
	(0.009)	(0.026)	(0.016)	(0.009)	(0.042)	(0.015)
College (%)	-0.003	-0.004	-0.018	-0.002	-0.001	-0.017
	(0.007)	(0.006)	(0.012)	(0.007)	(0.009)	(0.011)
Male newborn (%)	0.013	0.014	0.026	0.007	0.012	0.020
	(0.019)	(0.018)	(0.020)	(0.018)	(0.019)	(0.018)
Log(grdpc)	1.373***	2.013**	1.146	1.385***	1.368	1.211
	(0.377)	(0.762)	(0.887)	(0.367)	(1.082)	(0.814)
County fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	No	Yes	No	No	Yes	No
County × Time trends	No	No	Yes	No	No	Yes
Constant	-10.298**	-16.125**	-155.269	-10.225***	-10.457	-262.238***
	(3.763)	(7.179)	(101.860)	(3.601)	(10.071)	(91.338)
Observations	358	358	358	354	354	354
R^2	0.697	0.729	0.726	0.700	0.714	0.728

Notes: Dependent variable is the fraction of premature birth (%). Columns (1)–(3) report the estimates from Eq. (3.1). Columns (4)–(6) report the estimates from Eq. (3.3). Instrumental variable is the average NutriPlus spending per capita in neighboring counties. Estimates are weighted using the number of births in the cell. “Log(grdpc)” indicates the log of real GRDP per capita. Robust standard errors clustered by county are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table C.1.11. First-stage regression using linear interpolation

	OLS					
	(1)	(2)	(3)	(4)	(5)	(6)
NBR (1,000 KRW)	0.513***	0.289***	0.547***	0.514***	0.289***	0.559***
	(0.084)	(0.091)	(0.088)	(0.088)	(0.091)	(0.096)
Age < 20 (%)	0.137	0.077	0.107	0.135	0.078	0.099
	(0.090)	(0.132)	(0.065)	(0.094)	(0.138)	(0.066)
Age ≥ 35 (%)	0.012**	0.051**	0.000	0.012**	0.051**	0.000
	(0.005)	(0.018)	(0.007)	(0.004)	(0.018)	(0.007)
College (%)	0.007	0.007	0.006	0.007	0.007	0.005
	(0.006)	(0.007)	(0.004)	(0.006)	(0.007)	(0.004)
Male newborn (%)	-0.002	-0.001	0.001	-0.002	-0.001	0.001
	(0.008)	(0.008)	(0.006)	(0.008)	(0.008)	(0.006)
Gestational age	-0.042	0.030	-0.227			
(weeks)	(0.316)	(0.300)	(0.214)			
Log(grdpc)	-0.295	-1.286**	-0.588**	-0.271	-1.312**	-0.481
	(0.308)	(0.611)	(0.271)	(0.319)	(0.610)	(0.330)
County fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	No	Yes	No	No	Yes	No
County × Time trends	No	No	Yes	No	No	Yes
Observations	437	437	437	437	437	437
F-test on instrument	37.54	10.10	38.29	34.51	10.05	34.02

Notes: Table C.1.11 reports estimates from Eq. (3.2). The dependent variable is the NutriPlus spending per capita in a domestic county (1,000 KRW). Estimates are weighted using the number of births in the cell. Robust standard errors clustered by county are reported in parentheses. "NBR" indicates the average NutriPlus spending per capita in neighboring counties. "Log(grdpc)" indicates the log of real GRDP per capita. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table C.1.12. Effects of NutriPlus spending on birth weight using linear interpolation

	OLS (1)	OLS (2)	OLS (3)	IV (4)	IV (5)	IV (6)
NutriPlus spending	-7.122**	1.449	-7.933***	-18.399***	1.426	-16.007***
	(2.867)	(2.895)	(2.440)	(4.941)	(13.434)	(4.262)
Age < 20 (%)	-17.073***	-4.617	-6.613	-13.052***	-4.615*	-4.394
	(3.530)	(3.714)	(5.091)	(3.791)	(2.786)	(4.744)
Age ≥ 35 (%)	-1.397***	0.350	-0.191	-1.188***	0.352	-0.250
	(0.287)	(0.854)	(0.523)	(0.291)	(1.049)	(0.504)
College (%)	-0.705**	0.148	-1.410***	-0.558*	0.148	-1.329***
	(0.330)	(0.313)	(0.320)	(0.318)	(0.316)	(0.298)
Male newborn (%)	1.134*	0.625	0.447	1.152*	0.625	0.474
	(0.637)	(0.664)	(0.530)	(0.623)	(0.611)	(0.499)
Gestational age (weeks)	65.967***	43.095**	81.617***	61.957***	43.096***	75.156***
	(14.394)	(15.661)	(14.496)	(15.431)	(14.289)	(13.814)
Log(grdpc)	57.260***	-35.414	162.085***	51.134***	-35.449	151.324***
	(17.599)	(31.680)	(15.186)	(17.168)	(36.918)	(14.447)
County fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	No	Yes	No	No	Yes	No
County × Time trends	No	No	Yes	No	No	Yes
Constant	124.078	1873.356**	7107.571**	320.233	1872.447**	8965.570***
	(616.237)	(761.246)	(3017.064)	(672.499)	(782.194)	(3060.024)
Observations	437	437	437	437	437	437
R^2	0.816	0.877	0.867	0.807	0.877	0.864

Notes: Dependent variable is the fraction of birth weight (in grams). Columns (1)–(3) report the estimates from Eq. (3.1). Columns (4)–(6) report the estimates from Eq. (3.3). Instrumental variable is the average NutriPlus spending per capita in neighboring counties. Estimates are weighted using the number of births in the cell. "Log(grdpc)" indicates the log of real GRDP per capita. Robust standard errors clustered by county are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table C.1.13. Effects of NutriPlus spending on low birth weight (%) using linear interpolation

	OLS (1)	OLS (2)	OLS (3)	IV (4)	IV (5)	IV (6)
NutriPlus spending	0.117	0.116	0.143	-0.041	-0.460	-0.056
	(0.078)	(0.071)	(0.106)	(0.157)	(0.321)	(0.161)
Age < 20 (%)	0.159	0.036	0.063	0.216	0.090	0.117
	(0.148)	(0.165)	(0.175)	(0.160)	(0.142)	(0.180)
Age ≥ 35 (%)	-0.006	-0.021	-0.013	-0.003	0.016	-0.015
	(0.005)	(0.025)	(0.014)	(0.004)	(0.031)	(0.013)
College (%)	-0.002	-0.011*	0.007	-0.000	-0.006	0.009
	(0.005)	(0.006)	(0.009)	(0.005)	(0.006)	(0.008)
Male newborn (%)	-0.001	0.001	0.007	-0.001	0.003	0.007
	(0.024)	(0.024)	(0.022)	(0.023)	(0.024)	(0.021)
Gestational age	-1.331***	-1.165***	-1.557***	-1.387***	-1.152***	-1.716***
(weeks)	(0.290)	(0.312)	(0.321)	(0.285)	(0.349)	(0.310)
Log(grdpc)	0.274	1.087	-0.828	0.189	0.192	-1.093*
	(0.336)	(0.835)	(0.650)	(0.300)	(1.003)	(0.586)
County fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	No	Yes	No	No	Yes	No
County × Time trends	No	No	Yes	No	No	Yes
Constant	52.757***	38.936**	31.212	55.838***	46.538**	-31.409
	(13.709)	(16.897)	(102.544)	(13.135)	(19.420)	(93.391)
Observations	437	437	437	437	437	437
R^2	0.629	0.661	0.660	0.624	0.624	0.655

Notes: Dependent variable is the fraction of low birth weight (%). Columns (1)–(3) report the estimates from Eq. (3.1). Columns (4)–(6) report the estimates from Eq. (3.3). Instrumental variable is the average NutriPlus spending per capita in neighboring counties. Estimates are weighted using the number of births in the cell. "Log(grdpc)" indicates the log of real GRDP per capita. Robust standard errors clustered by county are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table C.1.14. Effects of NutriPlus spending on very low birth weight (%) using linear interpolation

	OLS	OLS	OLS	IV	IV	IV
	(1)	(2)	(3)	(4)	(5)	(6)
NutriPlus spending	0.001	0.018	-0.009	-0.079**	-0.161	-0.091*
	(0.019)	(0.018)	(0.021)	(0.040)	(0.106)	(0.047)
Age < 20 (%)	-0.005	-0.001	0.029	0.024	0.016	0.052
	(0.026)	(0.033)	(0.041)	(0.026)	(0.039)	(0.036)
Age ≥ 35 (%)	-0.002	0.004	-0.000	-0.001	0.016*	-0.001
	(0.002)	(0.007)	(0.005)	(0.002)	(0.009)	(0.005)
College (%)	0.001	0.001	0.000	0.002	0.003	0.001
	(0.001)	(0.001)	(0.003)	(0.001)	(0.002)	(0.002)
Male newborn (%)	0.000	0.001	0.000	0.001	0.001	0.001
	(0.006)	(0.006)	(0.006)	(0.005)	(0.005)	(0.006)
Gestational age	-0.298**	-0.349**	-0.459***	-0.326***	-0.345***	-0.525***
(weeks)	(0.125)	(0.131)	(0.126)	(0.116)	(0.126)	(0.112)
Log(grdpc)	0.067	-0.159	0.129	0.024	-0.437	0.019
	(0.103)	(0.252)	(0.142)	(0.103)	(0.298)	(0.146)
County fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	No	Yes	No	No	Yes	No
County × Time trends	No	No	Yes	No	No	Yes
Constant	11.347*	15.436**	34.646	12.728**	17.689***	32.224
	(5.688)	(6.850)	(32.452)	(5.271)	(6.653)	(29.660)
Observations	437	437	437	437	437	437
R^2	0.481	0.512	0.510	0.466	0.462	0.498

Notes: Dependent variable is the fraction of very low birth weight (%). Columns (1)–(3) report the estimates from Eq. (3.1). Columns (4)–(6) report the estimates from Eq. (3.3). Instrumental variable is the average NutriPlus spending per capita in neighboring counties. Estimates are weighted using the number of births in the cell. "Log(grdpc)" indicates the log of real GRDP per capita. Robust standard errors clustered by county are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table C.1.15. Effects of NutriPlus spending on high birth weight (%) using linear interpolation

	OLS	OLS	OLS	IV	IV	IV
	(1)	(2)	(3)	(4)	(5)	(6)
NutriPlus spending	-0.056**	-0.035	-0.039*	-0.083***	-0.039	-0.069***
	(0.022)	(0.031)	(0.022)	(0.020)	(0.042)	(0.022)
Age < 20 (%)	-0.039	-0.028	-0.052	-0.030	-0.028	-0.044
	(0.027)	(0.032)	(0.033)	(0.024)	(0.029)	(0.031)
Age ≥ 35 (%)	-0.005***	-0.005	-0.003	-0.004***	-0.005	-0.003
	(0.001)	(0.004)	(0.004)	(0.001)	(0.004)	(0.003)
College (%)	-0.006***	-0.005**	-0.006**	-0.005***	-0.005**	-0.006***
	(0.001)	(0.002)	(0.002)	(0.001)	(0.002)	(0.002)
Male newborn (%)	0.002	-0.000	0.001	0.002	-0.000	0.001
	(0.004)	(0.004)	(0.004)	(0.003)	(0.004)	(0.004)
Gestational age	0.021	-0.034	0.086	0.011	-0.034	0.062
(weeks)	(0.063)	(0.063)	(0.091)	(0.063)	(0.059)	(0.083)
Log(grdpc)	0.224**	0.132	0.312**	0.210**	0.126	0.273**
	(0.101)	(0.156)	(0.131)	(0.087)	(0.143)	(0.109)
County fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	No	Yes	No	No	Yes	No
County × Time trends	No	No	Yes	No	No	Yes
Constant	-2.454	0.678	5.242	-2.017	0.687	-0.773
	(2.999)	(3.061)	(21.080)	(2.930)	(2.697)	(18.222)
Observations	437	437	437	437	437	437
R^2	0.465	0.504	0.483	0.463	0.504	0.481

Notes: Dependent variable is the fraction of high birth weight (%). Columns (1)–(3) report the estimates from Eq. (3.1). Columns (4)–(6) report the estimates from Eq. (3.3). Instrumental variable is the average NutriPlus spending per capita in neighboring counties. Estimates are weighted using the number of births in the cell. "Log(grdpc)" indicates the log of real GRDP per capita. Robust standard errors clustered by county are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table C.1.16. Effects of NutriPlus spending on normal birth weight (%) using linear interpolation

	OLS	OLS	OLS	IV	IV	IV
	(1)	(2)	(3)	(4)	(5)	(6)
NutriPlus spending	-0.061	-0.081	-0.104	0.124	0.499	0.125
	(0.089)	(0.079)	(0.121)	(0.161)	(0.325)	(0.170)
Age < 20 (%)	-0.120	-0.007	-0.011	-0.186	-0.063	-0.073
	(0.137)	(0.158)	(0.157)	(0.151)	(0.135)	(0.167)
Age ≥ 35 (%)	0.011*	0.026	0.016	0.007	-0.011	0.018
	(0.005)	(0.025)	(0.014)	(0.005)	(0.031)	(0.013)
College (%)	0.008	0.015**	-0.001	0.005	0.010	-0.003
	(0.006)	(0.007)	(0.009)	(0.005)	(0.006)	(0.008)
Male newborn (%)	-0.001	-0.001	-0.008	-0.002	-0.002	-0.009
	(0.024)	(0.024)	(0.023)	(0.024)	(0.024)	(0.022)
Gestational age	1.310***	1.199***	1.471***	1.375***	1.186***	1.654***
(weeks)	(0.292)	(0.300)	(0.358)	(0.294)	(0.348)	(0.342)
Log(grdpc)	-0.499	-1.220	0.515	-0.398	-0.318	0.820
	(0.386)	(0.874)	(0.684)	(0.335)	(1.028)	(0.598)
County fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	No	Yes	No	No	Yes	No
County × Time trends	No	No	Yes	No	No	Yes
Constant	49.697***	60.385***	63.546	46.179***	52.775***	132.182
	(14.127)	(16.631)	(106.481)	(13.708)	(19.369)	(94.834)
Observations	437	437	437	437	437	437
R^2	0.539	0.575	0.576	0.532	0.532	0.569

Notes: Dependent variable is the fraction of normal birth weight (%). Columns (1)–(3) report the estimates from Eq. (3.1). Columns (4)–(6) report the estimates from Eq. (3.3). Instrumental variable is the average NutriPlus spending per capita in neighboring counties. Estimates are weighted using the number of births in the cell. “Log(grdpc)” indicates the log of real GRDP per capita. Robust standard errors clustered by county are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table C.1.17. Effects of NutriPlus spending on premature birth (%) using linear interpolation

	OLS	OLS	OLS	IV	IV	IV
	(1)	(2)	(3)	(4)	(5)	(6)
NutriPlus spending	0.110	0.046	0.158**	0.128	-0.199	0.202
	(0.081)	(0.098)	(0.072)	(0.215)	(0.588)	(0.156)
Age < 20 (%)	0.026	-0.125	0.014	0.020	-0.102	0.003
	(0.101)	(0.098)	(0.120)	(0.099)	(0.101)	(0.115)
Age ≥ 35 (%)	0.000	-0.021	-0.031**	-0.000	-0.005	-0.031**
	(0.008)	(0.021)	(0.014)	(0.009)	(0.039)	(0.013)
College (%)	0.000	-0.002	-0.015	-0.000	-0.000	-0.015
	(0.007)	(0.006)	(0.011)	(0.008)	(0.009)	(0.010)
Male newborn (%)	0.016	0.024	0.027	0.016	0.025*	0.027
	(0.018)	(0.016)	(0.018)	(0.017)	(0.015)	(0.016)
Log(grdpc)	1.322***	2.002***	1.100	1.329***	1.616	1.144
	(0.398)	(0.676)	(0.869)	(0.399)	(1.105)	(0.807)
County fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	No	Yes	No	No	Yes	No
County × Time trends	No	No	Yes	No	No	Yes
Constant	-10.133**	-16.667**	-137.259	-10.161***	-13.300	-272.404***
	(3.943)	(6.330)	(80.873)	(3.903)	(10.217)	(75.010)
Observations	437	437	437	437	437	437
R^2	0.682	0.711	0.714	0.681	0.707	0.713

Notes: Dependent variable is the fraction of premature birth (%). Columns (1)–(3) report the estimates from Eq. (3.1). Columns (4)–(6) report the estimates from Eq. (3.3). Instrumental variable is the average NutriPlus spending per capita in neighboring counties. Estimates are weighted using the number of births in the cell. "Log(grdpc)" indicates the log of real GRDP per capita. Robust standard errors clustered by county are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.