THE IMPACT OF NATURAL GAS ON HEALTH CARE UTILIZATION

by
NILUFER CETIK

Submitted to the Graduate School of Social Sciences
in partial fulfilment of
the requirements for the degree of Master of Arts

Sabanci University
August 2020

THE IMPACT OF NATURAL GAS ON HEALTH CARE UTILIZATION

Approved by:

Asst. Prof. ERDAL AYDIN ....
(Thesis Supervisor)

Prof. ABDURRAHMAN AYDEMIR ....

Assoc. Prof. BURCAY ERUS ....

Date of Approval:

ABSTRACT

THE IMPACT OF NATURAL GAS ON HEALTH CARE UTILIZATION

NILUFER CETIK

ECONOMICS M.A. THESIS, AUGUST 2020

Thesis Supervisor: Asst. Prof. ERDAL AYDIN

Keywords: air pollution, natural gas use, morbidity, fixed effects model, propensity score matching, Turkey

This thesis explores a quasi-experiment of gradual expansion of natural gas across the country that substituted coal in heating and cooking purposes. The impact of natural gas on morbidity between 2012 and 2018 is analysed by using fixed effects regression and further by propensity score matching. The Family Physician Centers were effective after 2012 and they are easily available to the population, thus the time span of the thesis starts from 2012. The fixed effects results suggest a small decrease in the per capita visits to FPCs if a province gets access to natural gas pipelines and 16 percent decrease in the FPCs visits with a one-unit increase in natural gas utilization. That is if the number of subscribers per population increases by 0.1, visits to FPCs decrease by 1.6 percent. Since fixed effects method is highly constrained, the thesis also makes use of a weaker method that does not account for time variation: propensity score matching. The results suggest that the intense adoption of natural gas in provinces decreases per capita FPC visits by a 16 percent compared to the provinces where natural gas adoption is hardly ever existent. The thesis could be studied as an example for developing countries as it provides generalized evidence on the impact of switching to a modern and cleaner energy resource that requires a large-scale investment.

ÖZET

TÜRKIYE'DE DOĞAL GAZIN SAĞLIK HIZMETLERI KULLANIMINA ETKISI

NİLÜFER ÇETİK

EKONOMİ YÜKSEK LİSANS TEZİ, AĞUSTOS 2020

Tez Danışmanı: Dr. Öğr. Üyesi ERDAL AYDIN

Anahtar Kelimeler: hava kirliliği, doğal gaz kullanımı, morbidite, sabit etki modeli (fixed effects model), eğilim skoru eşleştirmesi (propensity score matching)

Bu tez, Türkiye'de doğal gazın kademeli olarak yaygınlaşmasının ve doğal gazın ısıtma ve pişirme amaçlı kullanılan kömürü ikame etmesinin oluşturduğu yarı-deneysel ortamda araştırma yapmaktadır. Sabit etkiler regresyonu (fixed effects regression) ve eğilim skoru eşleştirmesi (propensity score matching) kullanılarak doğal gazın 2012-2018 yılları arasında morbidite üzerindeki etkisi araştırılmaktadır. Sabit etkiler modeli sonuçlarına göre, bir il doğal gaz boru hatlarına erişim elde ettiği takdirde kişi başına düşen aile sağlığı merkezine başvurular yüzde 0,4 azalmakta ve doğal gaz kullanımındaki yüzde 0,1'lik artış kişi başına düşen aile sağlığı merkezine başvurularını yüzde 1,6 düşürmektedir. Eğilim skoru eşleştirmesi sonuçlarına göre, doğal gazın yoğun olarak kullanıldığı iller doğal gazın neredeyse hiç kullanılmadığı illerle karşılaştırıldığında, bu illerde kişi başı aile sağlığı merkezine başvuruların yüzde 16 daha az olduğu görülmektedir. Bu tez, gelişmekte olan ülkelere örnek teşkil edebilecek bir çalışma olduğundan oldukça önemlidir. Tez daha modern ve temiz bir enerji kaynağına geçmenin etkilerine genellenebilir kanıtlar sunmakta, bu değişim büyük ölçekli yatırım kararı gerektirdiğinden sonuçlarının geçerliliği önem arz etmektedir.

ACKNOWLEDGEMENTS

I would like to express my sincerest gratitude to my thesis supervisor Assistant Professor Erdal Aydın, for his valuable support, guidance and patience during the process. I cannot thank him enough for creating a stress-free environment for my first academic work, for instructing me all the time and for making thought-provoking comments on my work to reach its best version.

I would like to present special thanks to my thesis jury members, Professor Abdurrahman Aydemir and Associate Professor Burçay Erus for being in my dissertation committee, spending their precious time and sharing their valuable comments and criticisms with me. I have learned so much from their econometrics courses and admired their insightful approach that I have chosen my academic field accordingly.

Lastly, I wish to thank my friends who were there to support me anytime, all those years. I am grateful to have met them.

To my parents, my little sister,
and our cat Lotus

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

RESPIRE The Randomized Exposure Study of Pollution Indoors and Respiratory Effects 11

TurkStat Turkish Statistics Institute 24

WHO World Health Organization 13, 41

YLD Years Lived with Disability 14

1. INTRODUCTION

Noncommunicable diseases (NCDs) are one of the most serious concerns for public health today, causing the largest number of death and disease around the world[1] (WHO, 2018*a*) hence, undermining social and economic development of the countries. The risk factors of NCDs such as physical inactivity, unhealthy diet and the excessive consumption of alcohol and tobacco are well accepted yet environmental factors that are also one of the main contributors of NCDs are often overlooked.

The latest research on global burden of disease show that air pollution plays a significant role in NCDs related deaths (Moesgaard Iburg and Collaborators, 2016). In 2012, ambient and household air pollution were responsible for 2.8 and 3.7 million NCDs deaths respectively, from cardiovascular diseases (CVDs), chronic respiratory diseases and lung cancer and the numbers are ever rising. These risk factors of NCDs together with diabetes, accounted for over 80 percent of all premature NCDs deaths in 2018 which are actually preventable and avoidable[2] (WHO, 2018*a*). In addition, 3 billion people who cook and heat their homes with biomass fuels and coal are under serious risk because of decreased indoor air quality. In 2016 alone, 3.8 million premature deaths were attributable to indoor air pollution (WHO, 2017*b*). In 2018, both ambient and household air pollution were officially recognised as risk factors for NCDs (UN General Assembly, 2018).

In Europe, after climate change, the air pollution is ranked the second biggest environmental concern leading to severest health issues (European Commission and

[1] 41 million people die related to NCDs each year, equivalent to 71 percent of all deaths globally. 15 million of these deaths are listed premature and more than 85 percent of these premature deaths occur in low- and middle-income countries. Also, 91 percent of the world's population live in places where air quality levels exceed WHO limits.

[2] By 2018, CVDs were the number one cause of death globally with 31 percent of the total deaths (17.9 million lives annually). Cancer was the second leading cause of the global deaths with 16 percent (9.6 million lives) for which at least 30 percent of them could be prevented. The two most common chronic respiratory diseases were chronic obstructive pulmonary disease (COPD) and asthma. By 2018, 235 million people suffered from asthma and 3 million people died from COPD annually of which more than 90 percent of them occur in low- and middle- income countries. Moreover, air pollution was estimated to cause about 25 percent of ischaemic heart disease deaths and 24 percent of stroke deaths, 29 percent of lung cancer deaths and 43 percent of COPD deaths. Solely, ambient air pollution was responsible for about 17 percent of ischaemic heart disease and stroke deaths, 16 percent of the lung cancer deaths, 25 percent of COPD deaths, and about 26 percent of respiratory infection deaths. (WHO, 2018a)

European Parliament, 2018). The most dangerous pollutants to health are recognised as PM, NO_2 and ground-level O_3[3] (EEA, 2019). The total costs of ambient air pollution within OECD region was estimated as 1,280 dollars per capita for 2015 which is 5 percent of income in 2015. Furthermore, the non-market costs of ambient air pollution accounted for 94 percent of the total costs in 2015 (OECD, 2016).

Although the total burden of diseases was decreased by 4 percent from 2000 to 2013, the data from Turkey reflects no less of a burden of NCDs than the rest of the world. A recent study estimates that 88 percent of deaths in 2013 were related to NCDs, despite the fact that NCDs' share of mortality was declining. The same study estimates that 81 percent of the disease burden was related to NCDs and NCDs' share of morbidity was rising (Hacettepe University Faculty of Medicine, 2017). Additionally, 17 percent of the deaths from NCDs were premature deaths (WHO, 2017*a*). More recently, it is estimated that the deaths related to air pollution were about 30,000 in 2016 (OECD, 2019).

The total annual number of deaths from NCDs was projected to increase to 55 million by 2030 if no action is taken against risk factors and because individuals cannot have control over most of the sources of pollutants, policy-makers should step in by imposing effective public policies. They should collaborate and cooperate even with sectors outside health at national, regional and global levels (WHO, 2013).

Inspired by the above discussion, this thesis tries to estimate the impact of utilizing a cleaner energy resource, cleaner relative to the previously used ones, on the burden of disease that is caused by air pollution. The example demonstrated in this thesis is the gradual adoption of natural gas in Turkish provinces and the health outcomes is determined as per capita visits to family physician centers. The process sets a reliable example for developing countries who consider investing in clean energy either for economic incentives or for public health outcomes. Since such a large-scale investment decision requires countless analysis and planning, seeing the positive outcomes of such an investment not only in energy market but also in health indicators might help countries to make their decisions.

While accounting for endogeneity problems, the thesis provides a generalized evidence for developing countries who invest in the replacement of widely used energy source with a cleaner one. The results for regression analysis with fixed effect model suggests that variation in the timing of adoption of natural gas across provinces decreases the per capita visits to FPCs by a small percent, 0.04. Consistent with existence of natural gas pipelines, this effect is observed larger after adding addi-

[3] Responsible for 412 thousand, 71 thousand and 15,1 thousand premature deaths in 2016, respectively.

tional years to adoption of natural gas. One more year spent with available natural gas network reduced per capita FPC visits by 7.3 percent. Finally, a negative yet insignificant impact of 16 percent is captured with one unit increase in the natural gas utilization of population in a certain province, which translates to 1.6 percent decrease in the FPC visits if the number of subscribers per population incerases by 0.1. Supporting evidence is found by running robustness checks with a different definition for intensity measure. The results suggest that policy-makers not only should provide means of consumption for a healthier resource but also they should encourage adoptive behaviour to this new and cleaner energy resource. On the other hand, the results from caliper matching with replacement suggests that the intense adoption of natural gas in provinces decreases per capita FPC visits by 16 percent compared to the provinces natural gas adoption is hardly ever existent, confirming above discussion that a significant change in the health outcomes could be reached after a meaningful level of adoption.

The drawbacks of this study emerges from the data restrictions of health variables. We know that a change in the health status may reflect a different association of air pollution with the types of health care visits. Also, the severeness of the impact is different for distinct age groups, for distinct socioeconomic groups and for distinct illnesses. The data used in this study cannot capture these variations because the data consists of overall numbers of visits to family physician centers for each province. Additionally, the data spans a relatively short time which reduces variation between provinces.

The following chapter provides a thorough summary of related literature. In chapter "Data and Identification", the sources of data for dependent, independent and control variables, descriptive statistics and identification methodology for the quasi- experiment is presented. Chapter 4 "Methodology and Results" discusses the estimation methodology used in the analysis and results of those. The last chapter "Conclusion and Discussion" summarizes the aim and the findings of the thesis together with restrictions of the study.

2. RELATED LITERATURE

2.1 Mortality and Morbidity

The regulations on air quality depend mostly on the findings of epidemiological research when discussing the bio-mechanic relation of air pollution with mortality and morbidity.

One approach to study infant mortality and the deteriorations in infant health caused by air pollution is to use individual-level data gathered from birth and death certificates. The first example of this branch uses individual-level data and weekly pollution levels to investigate the infant mortality in California over the 1990s. Rather than examining quasi- experiments, it deals with endogeneity problem by using within zip code variation and finds that air pollution is an important determinant for infant health and CO is the most potent pollutant for infant mortality even at relatively low levels of air pollution (Currie and Neidell, 2005). A consecutive study takes infant health in New Jersey over the 1990s and finds that a one unit change in mean CO levels during the last trimester of pregnancy increases the risk of low birth weight by 8 percent, the same change during the first 2 weeks after birth increases the risk of infant mortality by 2.5 percent (Currie, Neidell, and Schmieder, 2009).

In the last decade the concerns for infant and children's health continue to draw attention. A European study in 2012 by Coneus and Spiess (Coneus and Spiess, 2012) takes a similar approach to that of Currie et al. (Currie, Neidell, and Schmieder, 2009). What they add to the previous study is that they track the children from their birth to 3 years of age by using representative data of 2002-2007 from the German Socio-Economic Panel (SOEP), hence they account for time-invariant and unobserved neighbourhood and maternal characteristics. The negative impact of

CO on infant health is enhanced with their results, suggesting a 289 g lower birth weight because of a high exposure to CO prior to birth. It is also found that high O_3 levels increases the prevalence of bronchitis and respiratory illnesses among children between ages 1-3. A subsequent study focuses on the commissioned desulfurization at power plants in Germany as natural experiment. The impact of decreased SO_2 pollution is estimated as a prevention of 850 to 1600 infant deaths from 1985 to 2003 (Luechinger, 2014). Another European research published in the same year uses data for multiple birth cohorts to gauge the impact of pollutants on respiratory health outcomes of children after age 1 and the previous results are strengthened such that marginal increases in CO and O_3 in the short term exposure are significantly associated with increases in respiratory treatments and CO exposure of previous year affects respiratory conditions more severely (Beatty and Shimshack, 2014).

A research on mortality from the developing countries takes Indonesian wildfires during the Fall of 1997 as the cause of extreme and unprecedented air pollution (Jayachandran, 2009). By employing 2000 Census data, the author finds that the fire-induced increase in air pollution is associated with a 1.2 percent decrease in cohort size which means 15,600 children or infants are missing. Moving a step forward, Arceo et al. (Arceo, Hanna, and Oliva, 2016) compares the impact of air pollution on infant mortality in developed countries with the results from developing countries, challenging the external validity of the results. They use the existence of thermal inversion as an instrument for air pollution in Mexico City and estimate a 0.40 percent and 0.33 percent increase in infant mortality with a 1 percent increase in PM_{10} and CO over a year, respectively. The results are significant for respiratory illnesses. Comparison of the results with the ones from the US suggests that in Mexico City, CO has a larger impact and therefore, using estimates from the US setting may understate the benefits of environmental regulations in developing countries.

The following two seminal epidemiological papers focus on the health outcomes of adults in the long term exposure to $PM_{2.5}$ and they provide mixed results for cause-specific mortality rates (Pope et al., 2004; Pope III et al., 2002). In the first study they use the data collected by American Cancer Society mortality survey of approximately 1.2 million adults with a long follow-up time from 1982 through 1998 (Pope III et al., 2002). The findings suggest that with a $10-\mu g/m^3$ (about 9 percent) increase in $PM_{2.5}$, cardiopulmonary mortality rate rises by 0.06 and lung cancer mortality rate rises by 0.08. The second study uses the same data with a more detailed cause-specific information on mortality. The study concludes that with a $10-\mu g/m^3$ increase in $PM_{2.5}$, mortality rate of all cardiovascular disease rises by 0.07 (Pope et al., 2004). However, contrary to the previous research, no statistically significant impact of long term exposure on respiratory diseases is found. Another epidemiological study investigates the relationship between the onset of serious health conditions and exposure to short- and long-tem PM_{10} and O_3 levels (Evans and Smith, 2005). Adding to Pope et al.'s (Pope et al., 2004; Pope III et al., 2002) findings for particulate matter, their results suggest that serious heart conditions are more likely to appear with an increased ambient air pollution both in current and long-term, when in fact, onset of chronic lung conditions are not associated with PM_{10} levels but rather they are associated with the long term exposure to high levels of O_3.

Evidence on the impact of ambient air pollution from Latin America is consistent with the evidence from developed countries. In Santiago, Chile, it is found that a change equal to $10 - \mu g/m^3$ in daily PM_{10} averaged over three days is associated with a 1.1 percent increase in mortality, mostly from respiratory and cardiovascular conditions (Ostro et al., 1995). A more recent study uses data from nine Latin American cities and the results confirm that increased ambient concentrations increase risk of mortality. A statistically significant increase in mortality is associated with increased O_3 concentrations (Romieu et al., 2012).

2.2 Hospital Admissions

Earlier epidemiological studies that investigate the relationship between air pollution and hospital admissions commonly use time-series analysis, focus on the short-term effects and concentrate on a single city. Most of the studies on air pollution and health comes from England possibly because they experienced a severe air pollution episode in London, in 1952, so called the Great Smog of London. Although air quality in England has improved over the years, a report by the House of Commons (House of Commons, 2018) concludes that air pollution causes an estimated 40,000 premature deaths across the country each year, costing the UK an annual 20 billion pounds, making air pollution the second largest cause of mortality after smoking.

The two studies from London analyse the short term effects of air pollution on daily hospital admissions for respiratory disease for three distinct periods contained in 1987–1994 (Bremner et al., 1999; Ponce de Leon et al., 1996). They find that O_3 significantly increases daily admissions among all age groups; however, PM_{10} has little or no effect on hospital admissions. A third study from London takes a step forward by analysing the long-term impact of air pollution on the respiratory hospital admissions (Maddison, 2005). The contradictory results suggest a 0.14 percent decrease in the number respiratory admissions but not cardiovascular admissions after a 1 percent reduction in PM_{10} levels. Hence, the cost of CVDs should not be attributed to poor air quality yet the cost of respiratory diseases can be. The final research from England is conducted to gauge the hospital admissions between 1994 and 1996 in the West Midlands (Anderson et al., 2001). Neither respiratory admissions nor cardiovascular admissions were found associated with any air pollutant for all- ages. However, by taking only 0-14 age group into account, it is found that all pollutants are associated with hospital admissions.

Besides England, a handful of research is done for other places in the developed world. In Ontario, Canada, the number hospital admissions related to respiratory health between 1983 and 1988 are positively associated for all age groups (Burnett et al., 1994). The infants are the most harmed with having 15 percent of admissions associated with the O_3-sulfate pollution mix and elderly are the least harmed with having 4 percent of admissions associated. In Brisbane, Australia, similar results follow. Daily hospital admissions during the period 1987-1994 are associated with O_3 and particulate pollution for asthma and respiratory diseases. Sulphur dioxide is also associated with CVDs in addition to asthma and respiratory diseases (Petroeschevsky et al., 2001). However, a study in Perth, Western Australia concludes conversely and states that changes in O_3 concentrations were not significantly associated with any disease while particulate pollution affects respiratory diseases, CODP, pneumonia, asthma and CVD hospitalizations from 1992 to 1998 (PhD et al., 2006). In Rome, Italy, the results support the US context, suggesting that O_3 is strongly associated with acute respiratory infections for daily emergency admissions, but only for children. No effect was found for particulate matter and SO_2. Same-day level of NO_2 was associated with respiratory admissions and same-day level of CO was associated with asthma and COPD admissions for all ages (Fusco et al., 2001).

As an example from developing countries, in Hanoi, Vietnam, all ambient air pollutants except CO are found to have positive impact on the daily number of hospital admissions related to bronchitis and asthma between 2007 and 2014 among children aged 0–17. Including CO, all pollutants are positively associated with hospital admissions due to pneumonia. The positive associations are strongest for infants (Nhung et al., 2018). Furthermore, a 5 percent increase in the length of hospital stay among children aged 0 to 5 with ALRI is observed after an increase defined in O_3 or in PM_{10} levels (Nhung et al., 2019).

Several studies investigate the role of avoidance behaviour that could be observed after government's announcement on air quality of a particular day. If we assume people are informed about the effects of the ambient concentrations of different pollutants, then behaviour adoption is expected accordingly. Based on survey analysis, these studies reflect that people response to air quality briefings. More precisely, preexisting health conditions determine the attendance to outdoor activities depending on air quality (Skov et al., 1991) and existence of smog changes people's attitude towards avoidance (Bresnahan, Dickie, and Gerking, 1997). Not only air quality but also seasonal factors such as pollens or temperature play a role in observed avoidance behavior (Bickerstaff and Walker, 1999), children with asthma are exposed to restricted outdoor activities by their parents during poor air quality days (McDermott, Srivastava, and Croskell, 2006), and a smog alert decreases the attendance to outdoor activities by children and elderly because of higher benefits of avoidance and also by locals because of lower costs of avoidance (Neidell, 2009). For example, avoidance behaviour leads to 1 percent reduction in asthma hospitalizations on the days with smog alerts in California between 1992 to 1998. Additionally, the hospital admissions of children for asthma between ages 1–18 in California is positively associated with CO levels and the decreased level of pollution from 1992 to 1998 have prevented 5 to 14 percent increase in asthma admissions which translates to savings of an approximately 5.2 million dollar in hospital expenses (Neidell, 2004). The children's hospital admissions related to respiratory diseases in England from 2003 to 2007 increase by 0.1 percent with a 1 percent increase in NO_2 or O_3 concentrations (Janke, 2014). Avoidance behaviour is observed for the subset of hospital admissions for asthma, reflected as 8 percent decrease in hospital admissions after a 1 percent increase in NO_2 or O_3 concentrations. This is due to the avoidance behaviour that has a lower cost for this specific subset as attributable to received information on air quality that leads to either an adjustment for the dose of relieving medicine or carrying the inhaler if necessary. However, the avoidance behaviour that is observed only for the asthma patients does not create a statistically significant underestimation in the results for hospital admissions.

2.3 Instrumental Variables and Quasi- Experiments

Instrumenting air pollution in order to account for endogeneity problems that is caused by non-random levels of air pollution across locations or individual exposure choices has several examples in the literature. The previous studies mentioned in this chapter have dealt with endogeneity by using within zip code variation in individual level data (Currie and Neidell, 2005) or studying a change by policy decision (Luechinger, 2014) or by a natural disaster (Jayachandran, 2009). In this section, a number of studies that focus on dealing with endogeneity is presented.

Moretti and Neidell (Moretti and Neidell, 2011) construct their instrument as daily boat traffic at port of Los Angeles to estimate the short run impact of O_3 levels on respiratory hospitalizations while at the same time accounting for avoidance behaviour. They find that a 0.01 ppm increase in the five-day average O_3 results in an increase of 44 million dollars in annual costs related to respiratory hospitalizations. The cost of avoidance is 11 million dollars per year. The second study takes the variation in daily airplane taxi time in California as its instrument for air pollution and finds that a one standard deviation increase in daily pollution levels, particularly of CO, explains one third of average daily admissions for asthma which leads to an additional 540 thousand dollars of costs for respiratory and heart-related admissions (Schlenker and Walker, 2016). Even though it is confirmed that infants and elderly are the most sensitive to ambient air pollution, the aggregate effect is found to be larger for people between 20-64 ages. As another example from California, ambient air pollution is instrumented by using traffic congestion and its impact on infant mortality is found to increase by 0.2 percent as local traffic levels increase one standard deviation. The marginal effects are larger on weekly infant mortality rates, especially for premature or low birthweight infants (Knittel, Miller, and Sanders, 2016).

Accounting for endogeneity is also possible by identifying a quasi-experiment that cause variations in air pollution levels in the absence of randomized controlled trials. One of the most cited studies takes the 1981-1982 recession in the US as quasi-experiment due to reduced but varying air pollution level across cities (Chay and Greenstone, 2003). The researchers find that 1 percent reduction in TSPs results in a 0.35 percent decline in the infant mortality rate at the county level which translates to 2500 fewer infant deaths from 1980-1982. Another frequently cited paper in the literature studies the impact of long-run reduction in TSP pollution on the adult mortality after inducement of Clean Air Act Amendments of 1970 which mandated aggressive regulation of local polluters in polluted counties, hence creating a variation in regulation intensity (Chay, Dobkin, and Greenstone, 2003). The study finds that even though the regulation decreases TSPs pollution, a systematic association with regulation and adult mortality cannot be established.

The attention to quasi-experimental settings has risen in the last decade. Lleras-Muney (Lleras-Muney, 2010) uses compulsory relocation of the military members and their families to identify the causal impact of pollution by using individual-level data. Distinctively, it covers children from birth to the age of five. The author finds that only O_3 has a significant impact on military children's respiratory hospitalisations. More specifically, one standard deviation in O_3 increases the probability of a respiratory hospitalization of children by about 8 to 23 percent, but not of infants. Currie and Walker (Currie and Walker, 2011) investigate the variation in traffic congestion after the introduction of electronic toll collection in New Jersey and Philadelphia, which sharply reduced traffic-born air pollution near toll plazas. They find that E-ZPass reduced the incidence of prematurity and low birth weight in the area of toll plazas by 6.7-9.1 percent and 8.5-11.3 percent, respectively. These reductions are estimated to value 9.8-13.2 million dollars. Similarly, the policy change in Germany induces the existence of low emission zones to reduce air pollution stemming from traffic (Gehrsitz, 2017). The representative studies on the effectiveness of low emission zones on reducing the pollution conclude mixed results of positive or zero effects (Cyrys et al., 2014; Morfeld, Groneberg, and Spallek, 2014; Wolff, 2014). However, this study finds that in low emission zones the concentration of pollutants are reduced and further, highly restrictive zones enjoy notably less air pollution than the less restrictive zones. On infants' health outcomes, the author finds that neither average birth weight nor the prevalence of low-weight births appear to be significantly affected by the policy, but small reductions in the incidence of stillbirth are observed.

Moving to the developing country context, an early study explores the relationship between levels of particulate matter and daily mortality rates in Delhi, India between 1991 and 1994 (Cropper et al., 1997). In the study, the estimates suggest that the impact of particulate matter is one third of what is estimated in the US because only 30 percent of deaths occur before age 65 in the US whilst death rate before age 65 is 70 percent and the death rate before age 5 is 20 percent in Delhi. Hence, the elderly are more sensitive to air pollution while deaths occurring due to air pollution in Delhi cause more life-years to be lost. The second study from India investigates the impact of environmental regulations on pollution reduction and infant mortality, uses a wider data set from 1986 to 2007 (Greenstone and Hanna, 2014). Although the compliance with regulations is expected to be weak in India, it is found that the regulations reduced air pollution substantially but not water pollution. Moreover, the study finds a limited but insignificant decrease in infant mortality. However, in China, the environmental regulations which were introduced in 1988, resulted in a 20 percent decline in infant mortality rate that corresponds to 3.29 fewer infant deaths per 1000 live births and the greatest reduction in mortality occurred during the neonatal period, by 63 percent according to Tanaka (Tanaka, 2015).

2.4 Indoor Air Pollution and Randomised Controlled Trials

Burning biomass for cooking or heating is one of the most important factors that causes increased indoor air pollution, hence it may lead to severe health problems. The WHO states that 3 billion people cook using polluting open fires or simple stoves fuelled by kerosene, biomass and coal and annually close to 4 million people die prematurely from illness attributable to indoor air pollution. Household air pollution causes the four major NCDs and it is estimated that close to half of deaths due to pneumonia among children under 5 years of age are caused by particulate matter inhaled from household air pollution (WHO, 2018*b*).

In the early studies, indoor air pollution in the developing countries attracted attention due to traditional cooking and heating habits relied mostly on firewood or kerosene use. In rural Nepal, the prevalence of chronic bronchitis was found to increase with increasing exposure to domestic smoke from fireplace (Pandey, 1984). In rural Gambia, the exposure to $PM_{2.5}$ by parental smoking and smoke from cooking fires found to cause increased ALRI among children under five (Armstrong and Campbell, 1991). One of the later epidemiological studies (Sharma et al., 1998), focuses on the prevalence of ALRI among infants due to the use of wood or kerosene for cooking which alter the indoor air quality differently. The findings confirm that a higher incidence of ALRI occurs in kerosene using households and in high pollution area. In a similar study, researchers focus on the types of household fuel use for cooking and heating (Mitter et al., 2016). The analysis reveals that natural gas users had a reduced risks of all- mortality and cardiovascular mortality while kerosene or diesel users experienced increased risks of all- mortality and cardiovascular mortality.

The stoves that burn solid fuel are the prime causes of indoor air pollution especially in the developing world. Therefore, the three main randomised controlled trials for the replacement of traditional three-stone fire or firewood stoves with low-cost improved stoves are mentioned here.

The first one is RESPIRE study which replaces indoor open fires with wood burning chimney stoves in rural Guatemala. For the women aged between 15-50 years who are responsible for cooking, although the fuel burned is the same, the 0.61 decrease in CO exposure reduced the risk in all- respiratory symptoms by 0.42 and the odds of having respiratory symptoms is reduced by 0.7, yet no significant effect is found on lung function after 12 to 18 months of improved stove replacement (Smith-Sivertsen et al., 2009). Additionally, the odds of having sore eyes and headache were reduced by 0.18 and 0.63 respectively (Díaz et al., 2007). For the children under 18 months, it is found that 0.5 percent decrease in CO exposure significantly reduced physician-diagnosed pneumonia by 0.87 (Smith et al., 2011).

The second randomised controlled trial is conducted by J-Pal in India. After a four-year follow up, Hanna et al. (Hanna, Duflo, and Greenstone, 2016) find evidence of decreased smoke exposure in the first two years for the main cooks in the household. However, after a while this effect disappears due to neglected stove maintenance. The changes in a wide set of health outcomes are found insignificant.

In rural Senegal, although the positive impact of stove replacements on smoke exposure worked through behavioural adoptions, such as increased outdoor cooking and reduced cooking time, it is found that use of improved stoves reduced the prevalence rates for both self-reported respiratory disease symptoms and eye infections by almost 0.07 among the ones responsible for cooking. (Bensch and Peters, 2015).

All of the above randomised controlled trials adopt the stove replacements burning the same fuel. However, Beltramo and Levine (Beltramo and Levine, 2013) compare the traditional stove use with solar oven use in rural Senegal. No impact of solar ovens on reducing exposure to CO or self-reported respiratory symptoms was found. The authors suggest that the reason of having no evidence is the incapability of one-pot solar oven to cook big and complex meals as lunch. Subsequently, the drawbacks of these studies are noted to have a short follow-up time [1] which undermines the potential health outcomes in longer terms; high cost of stoves in RESPIRE study which makes the stoves inaccessible for most; trained fieldworker visits in RESPIRE which drives experiment closer to laboratory setting and undermines individual's true valuation; and high costs of maintenance of stoves in India which reduces the household compatibility in the long term.

At this point, one should also note that these studies are small-scale and they examine the switch to modern energy under very poor conditions and in underdeveloped countries. This thesis, on the other hand, while accounting for endogeneity problem of studying air pollution, provides a more generalized evidence for developing countries by switching to modern energy after a large-scale investment.

[1] 12 to 18 months for (Smith-Sivertsen et al., 2009), one year for (Bensch and Peters, 2015), and six months for (Beltramo and Levine, 2013).

3. BACKGROUND AND DATA

3.1 Turkish Setting

The current situation of Turkish environmental performance reflects deterioration in air quality primarily due to country's dependence on carbon-intensive energy supply in order to sustain economic development. According to a recent review, fossil fuels represented 88 percent of total primary energy supply in Turkey, while the OECD average was 80 percent (OECD, 2019). The coal use for energy supply creates emissions of PM, SO_2, NO_x, CO_2 and other pollutants, far more than any other fuels. Consequently, it was observed that Turkey's greenhouse gas emissions increased the most among OECD countries since 2008. In 2019, Turkey had 31 percent more concentrations than the average concentration level in Europe, leading about 90 percent of the population to suffer from air pollution. Although limit values were expected to abide by EU standards by 2019, in most of the regions, pollutant concentrations exceeding WHO guidelines (WHO, 2005) were recorded for more than a half of the year (TMMOB, 2019) and they are expected to more than double between 2015 and 2030 (OECD, 2019). Thus, air quality and climate change are major concerns, especially in large cities due to PM emissions from transportation, and in industrialised regions due to SO_2 emissions related to burning fossil fuels. Additionally, it is observed that regions with low income levels experience decreased air quality linked to coal use in heating and cooking purposes. This aspect originates from the government subsidies for poor families to use coal for heating, even though natural gas is increasingly replacing fossil fuels since 1990s.

The threat of air pollution on public health is at worrying levels. According to IMF, health expenditures linked to air pollution from burning fossil fuels amounted to 19,4 million dollars in 2015. Likewise, emissions from fossil fuel use had an estimated cost of 13,2 million dollars for climate change (HEAL, 2017). It is reported that with a reduction in the use of fossil fuels, it was possible to prevent 73.8 percent of the deaths from air pollution.

It is also crucial to inform public about the air quality to promote avoidance behaviour. However, it was observed that in 2019 alone, most of the measurement stations in Turkey failed to report concentration levels of pollutants (TMMOB, 2019) and the costless information available on the Ministry of Environment's website was not sufficiently large (OECD, 2019).

Hence, in the light of above discussion, the impact of reduced air pollution by substituting coal with a cleaner energy source on the public health outcomes is studied by exploring the variation created by the gradual expansion of natural gas use for heating and cooking purposes across country and over time. The expansion process of natural gas pipelines in Turkey started in late 1980s as a result of an attempt to diversify its energy portfolio to keep up with the country's rapid growth and structural changes without remaining vulnerable to any shocks. The expansion accelerated after 2000s and by 2019 all of the 81 provinces of Turkey acquired access to natural gas which eventually leads to a cleaner air (Cesur, Tekin, and Ulker, 2017).

The fundamental reason for the government to impose wide use of natural gas over time was due to its prospects of maintaining growth and development with diversified energy portfolio and the decision was not related to any possible health gain. Centrally located and industrialised provinces with relatively cold weather are amongst the first adopters of natural gas and the reason why is common sense: The cost of infrastructure for natural gas pipelines urges the investors to minimise the cost by choosing such provinces first and then connecting them to neighbouring provinces. Therefore, the expansion of natural gas can be treated as quasi- experiment to investigate the causal relationship between adoption of natural gas and health outcomes in order to overcome endogeneity problems that may arise while estimating the impact of air pollution on health outcomes.

Previously, two studies have been published on the impacts of natural gas use in Turkey on the variety of health outcomes (Cesur, Tekin, and Ulker, 2017, 2018). In these studies, it is found that one percent increase in the natural gas use intensity decreases infant mortality rate by 4 percent, overall mortality rate by 1.4 percent, the adult mortality rate by 1.9 percent, and the elderly mortality rate by 1.2 percent. Recently, more attention is paid on the morbidity rather than the mortality in both developed and developing countries because with an advancement of increased life-years, the DALY and YLD gained more importance to manage health-care systems efficiently and prevent any illnesses leading to mortality (WHO, 2010). Therefore, as an extension to previous two studies, in this thesis, it is studied that whether the replacement of coal with natural gas had any impact on the country's morbidity rate. The morbidity is captured by the data for the number of visits to family physician centers between 2012 and 2018.

3.2 Data

The thesis employs fixed effects model in order to test the impact of rising natural gas adoption on healthcare utilization. While several distinct independent variables are defined regarding natural gas in general, the independent variable of the regression analysis is annual per capita visits to family physician centers. A vector variables are included in the regression analysis as well. In this chapter, the characteristics of the data and the choice of controls are explained in three parts and descriptive statistics are provided in the final section of the chapter.

3.2.1 Natural Gas

The main independent variable in this study is the intensity of natural gas use in a given province. This variable is derived from the following equation:

$$Intensity_{it} = \frac{NGS_{it}}{Population_{it}}$$

where NGS_{it} stands for the number of natural gas subscribers in a given year t in province i. The data for natural gas subscribers comes from the bimonthly Natural Gas Journal [1] which is published since 1988, before natural gas use for cooking and heating purposes launched in the Turkish provinces. The data covers the number of natural gas subscribers to Natural Gas Companies across provinces for the years from 1996 to 2018. For several years, detailed information such as the building types in which the natural gas is used i.e. whether it is a residential building or community building, and boiler types in the buildings i.e. whether it is in central or individual

[1] The journal is published since 1988 and has an online archive starting from April 2007. http://www.dogalgaz.com.tr/ Last accessed to the online archive on May 15, 2020. The earlier issues are available in the journal's library on request.

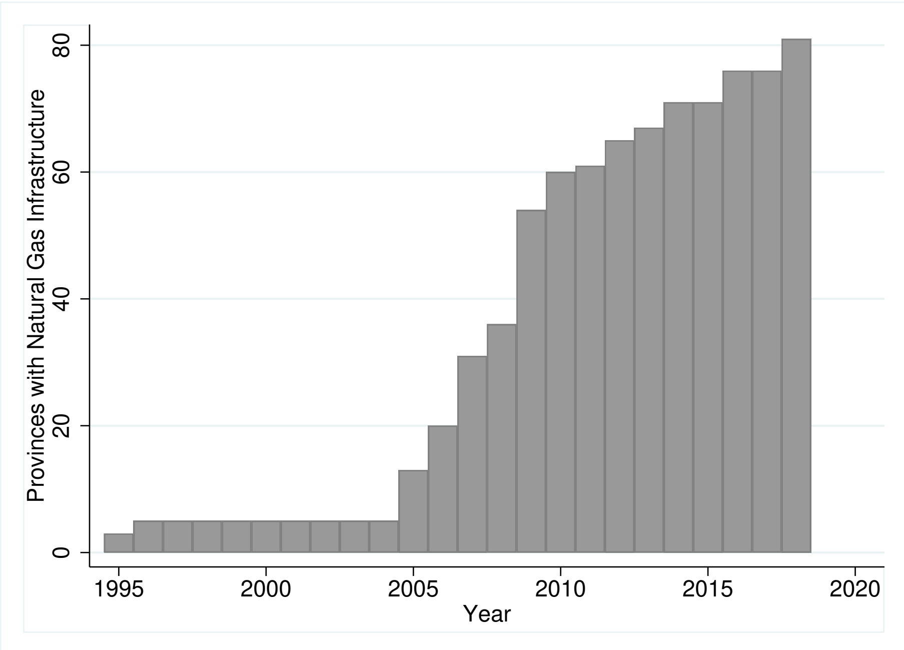
use, are provided.[2] With this detailed data set, the total number of subscribers is generated by:

$$NGS_{it} = RH_{itk} + CH_{itk}$$

where RH_{itk} is residential building heating, CH_{itk} is community and commercial building heating; k describes whether the heating system is central or individual heating as given in the data, which leads to different calculation methods for centrally (subscribed) heated buildings and individually (subscribed) heated flats.

The development of the natural gas infrastructure network over the country through the years can be seen in Figure 3.1. The number of provinces that had natural gas establishment rises from 31 to 61 from 2007 to 2011, and to 71 by 2014. By the year 2018, all of the 81 provinces had natural gas pipeline infrastructure.

Figure 3.1 Number of Provinces with Natural Gas Infrastructure

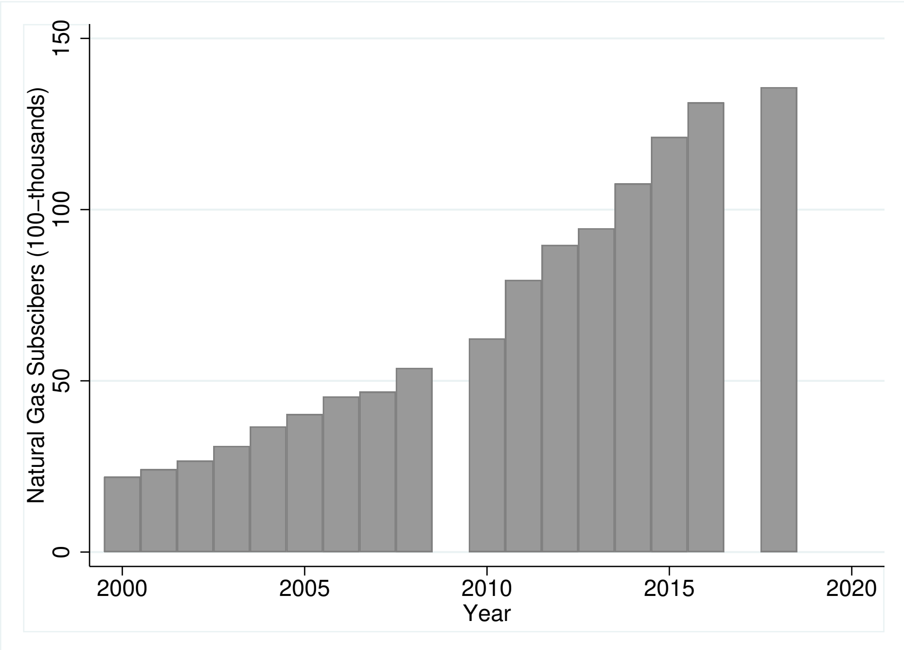


However, because it does not mean that an increase in the number of provinces with natural gas pipeline access directly increases the number of subscribers in the provinces, exploiting the variation in natural gas subscribers over the years and provinces will ameliorate the analysis. A brief look at the Figure 3.2 shows that even though the rise in the natural gas existence is steeper from 2006 to 2011,

[2] For the year 2010, only the number of natural gas users is available. The number of actual users amongst subscribers are usually less; however, comparing with the previous and preceding years' subscribers, the number of users is found to be in the number of subscribers' range. The same finding follows from 2013 to 2015 for which only the number of residential building subscribers are given. Also, in 2009 the four large city data (Ankara, Istanbul, Kocaeli, Sakarya) on natural gas subscribers are missing and in 2017, the data is not available. Therefore, the years 2010 and 2013 to 2015 for natural gas data are included in the analysis whereas the years 2009 and 2017 are not.

the rise in the natural gas subscribers is more between the years 2011 and 2016.[3] Furthermore, since the main variable is $Intensity_{it}$, one should consider the rise in the intensity level as well. In Figure 3.3, we see that the intensity of natural gas use in the provinces follow a similar trend to that of number of subscribers. It is also concluded that, similar to the differences between the rises in natural gas access and natural gas subscribers, the change in the intensity was larger from 2011 to 2016 than 2006 to 2011.[4]. Therefore, the analysis of this thesis which spans the years from 2012 to 2018 does not suffer from variation loss which might be a concern due to the extension of natural gas pipeline network all over the country.

Figure 3.2 Natural Gas Subscribers in 100-thousands

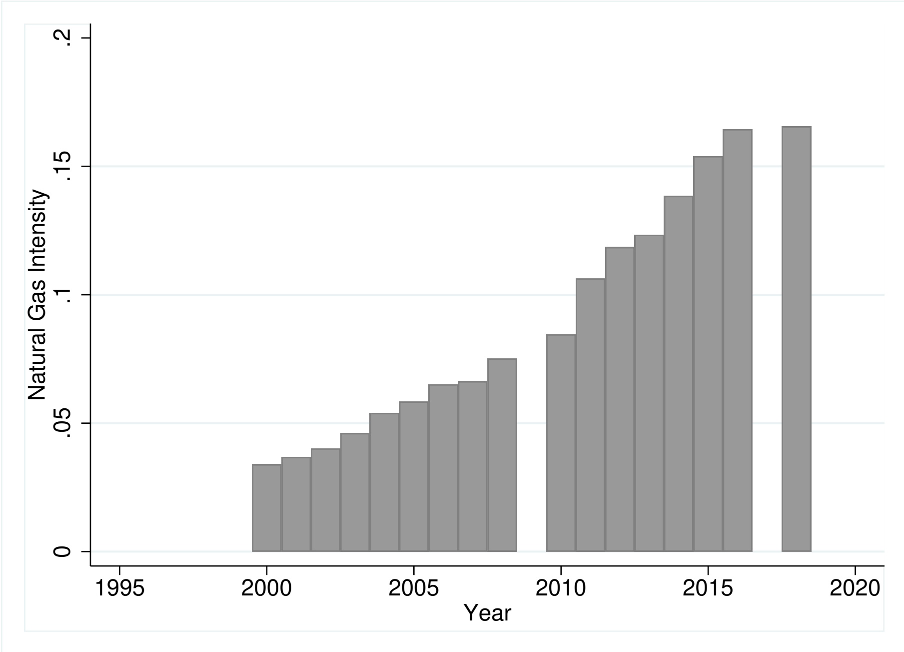


Notes: In 2009, the subscriber data for Ankara, Istanbul, Kocaeli and Sakarya are missing. Hence 2009 data is dropped. In 2017, the data is not available.

[3] 41 new provinces gained access to natural gas and 3,404,221 new subscriptions are made from 2006 to 2011 while 15 new provinces gained access to natural gas and 5,155,101 new subscriptions are made from 2011 to 2016

[4] 0.058% in the former period and 0.041% in the latter.

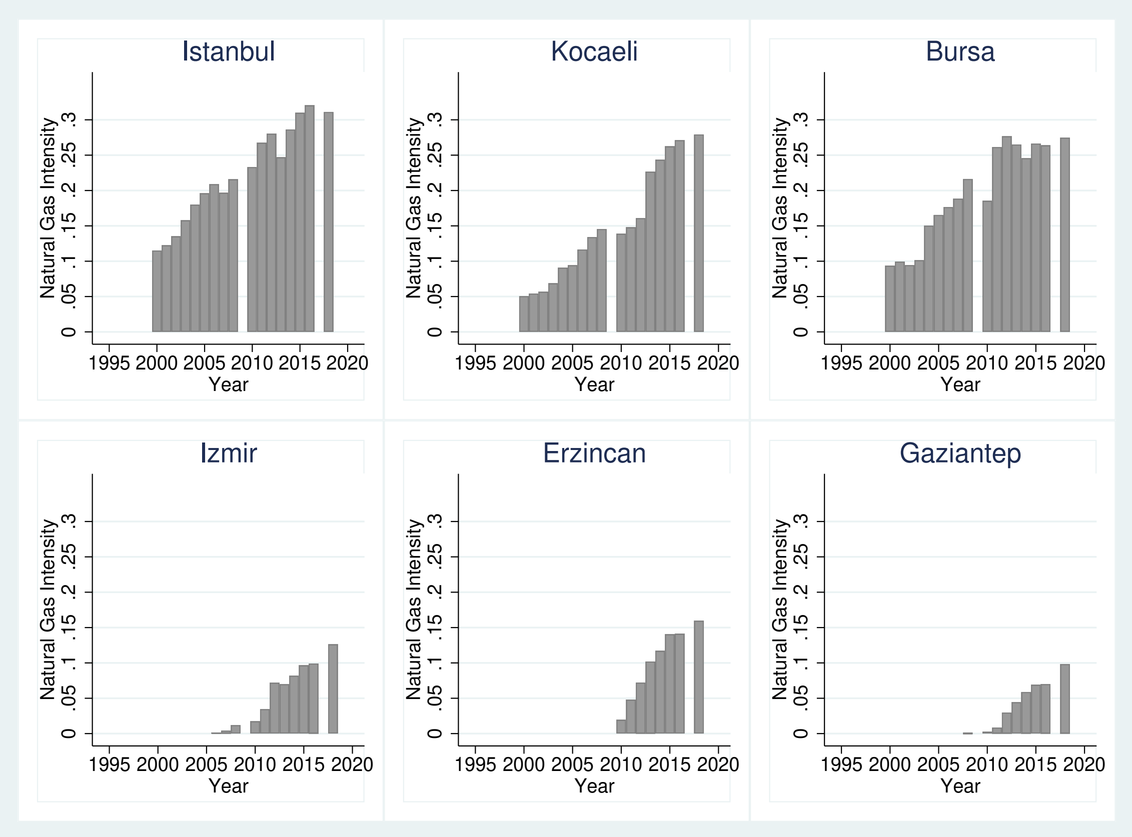
Figure 3.3 Natural Gas Intensity per capita



Notes: In 2009, the subscriber data for Ankara, Istanbul, Kocaeli and Sakarya are missing. Hence 2009 data is dropped. In 2017, the data is not available.

The variation in the natural gas use intensity over time and provinces is a crucial base for the analysis of this thesis. To capture the variation over time and especially over provinces, intensity levels for some of the provinces are presented separately in Figure 3.4. Selected provinces represent Turkish provinces with different characteristics. Istanbul, Kocaeli, Bursa, Izmir and Gaziantep are among the top ten provinces with respect to their population levels by 2018. Except Gaziantep, they are also in the first-category level with respect to their socio-economic development levels (Ministry of Industry and Technology, 2019). Gaziantep is placed in the third-category level and Erzincan is placed in the fourth-category level. Additionally, their geographical locations differ. The top row of the figure represents densely populated and industrialized Marmara Region while in the bottom row Izmir is located in the touristic Aegean Region with a typical Mediterranean climate. Erzincan is located in the least densely populated Eastern Anatolia Region which is known for its rugged mountains and severe winters with heavy snowfalls. Lastly, Gaziantep is situated in Southeastern Anatolia Region which has a broad plateau surface with rough winters in the mountainous area and quite warm and dry summers near coasts contrasting to Eastern Anatolia Region. As it can be seen from the Figure 3.4, the first adopters are the industrialized provinces while the natural gas intensity between them differ over years. Erzincan adopts natural gas later than any of the selected provinces due to its geographical location, yet it adopts natural gas more intensely in a short amount of time than Izmir and Gaziantep which experience warmer weather throughout the year. Hence, from the Figure 3.4, we can conclude that the transition of the energy use to natural gas in the provinces are not only far from complete but also the natural gas adoption degrees vary in between those. However, this aspect of expansion does not undermine the quasi-experimental nature of the energy transition. To see the the expansion being independent of province characteristics, one can refer to B

Figure 3.4 Natural Gas Intensity in Selected Provinces



Notes: y-axis has been scaled to be between 0 and 0.35 in order to ease the comprehension of different intensity levels and to carry on the continuity between the graphs for selected provinces. In 2009, the subscriber data for Ankara, Istanbul, Kocaeli and Sakarya are missing. Hence 2009 data is dropped. In 2017, the data is not available.

The final figure of this subsection is the map of Turkey by provinces. The different levels of natural gas utilization among provinces is illustrated for 2012 and 2018, separately. The darkest shade stands for the most intensely use of natural gas. Conversely, the lightest areas show the least intensely use of natural gas. There are several provinces in both maps that do not have any data on utilization, meaning that there were no subscriptions at the time. It is concluded by the Figure 3.5 and Figure 3.6 that variation among the provinces exist. The provinces that are difficult to reach at because of their geographical locations and shapes (i.e. Agri, Artvin, Giresun) suffer from less utilization due to the late adoption of natural gas. They are accompanied by the provinces with relatively warm weather which are located in the Southern part of the country (i.e. Mardin, Hatay, Antalya). Contrarily, the provinces with relatively long and severe winters that are centrally located or industrialized (i.e. Ankara, Istanbul, Eskisehir, Bursa) enjoy the highest utilization of natural gas. Additionally, there is a spillover effect of existence of natural gas such that neighbouring provinces enjoy utilization (i.e. Kirsehir, Kirikkale, Yalova). Also, from 2012 to 2018, the natural gas utilization is rising all over the country, demonstrated by the map going darker.

Figure 3.5 Natural Gas Intensity per capita in Provinces, 2012

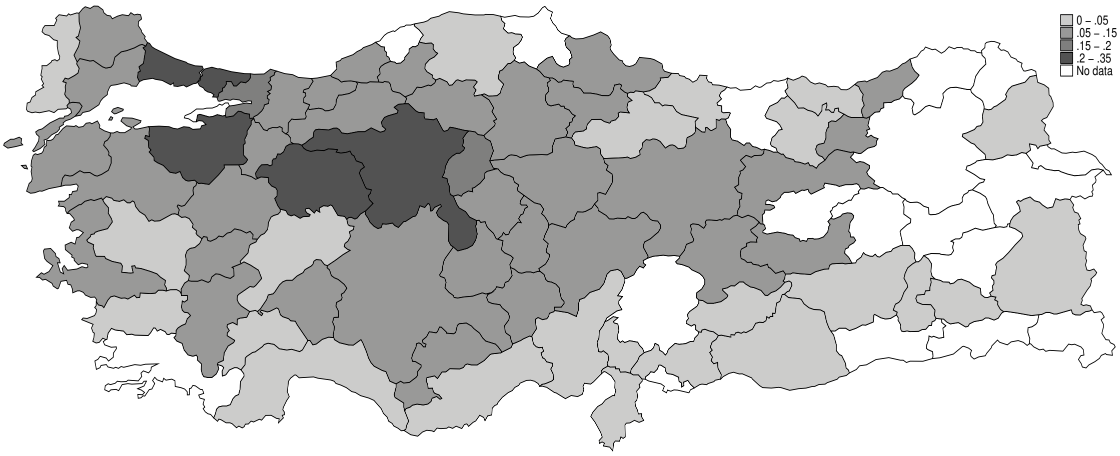
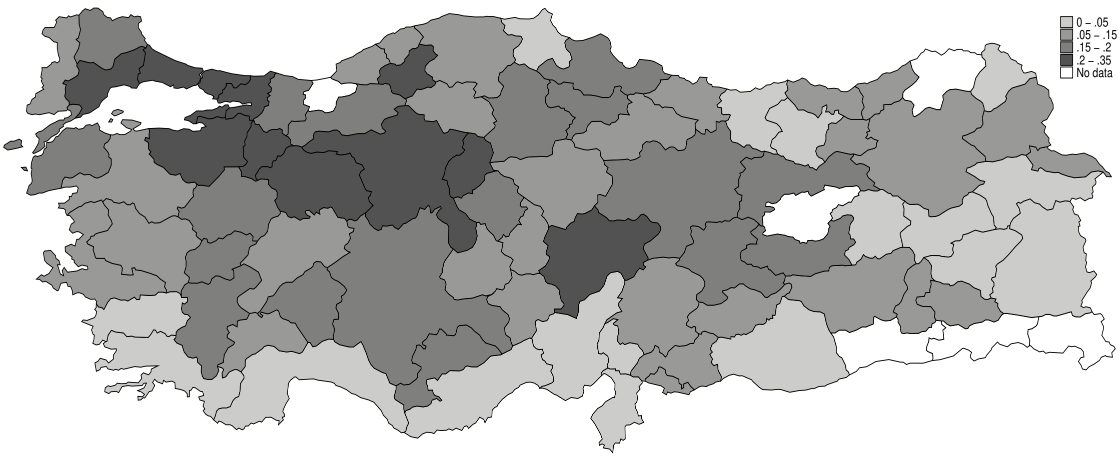


Figure 3.6 Natural Gas Intensity per capita in Provinces, 2018



3.2.2 Per capita Visits to Healthcare Facilities

Per capita visits to Family Physician Centers (FPC) is the dependent variable of this study. By taking per capita levels, the increasing size of the population due to refugees and births are controlled for. The aim is to gauge the variation in the incidence of respiratory health problems at province-year level as a result of the improvement in the type of energy in use that supposedly affects the indoor air quality through cooking and heating channels and the outdoor air quality through its utilization in the industries and in the commercial buildings. Hence, the annual visits to physicians because of a respiratory system malfunction is expected to be higher in the provinces in which the air quality is low. This thesis tries to gauge the impact of air quality on respiratory diseases by making use of natural gas adoption. The data for the annual number of visits to healthcare facilities of different levels comes from the Yearly Statistical Books of Ministry of Health and it covers the years between 2012 and 2018.

In Table 3.1, it is shown that the mean per capita visits to different level healthcare facilities in the provinces that experience above-limit PM_{10} levels is significantly different than the mean in the provinces that does not experience above-limit PM_{10} levels. The difference is larger for the logarithm of FPC visits[5].

Table 3.1 Differences in means between high- and low-pollution provinces

Variable	Low PM_{10}- Mean	High PM_{10}- Mean	difference	t	p-value
Log Visits to Healthcare Facilities	2.10	2.02	.082	2.21	0.029
Log Visits to FPCs	0.976	0.814	.161	2.18	0.031
Log Visits to Hospitals	1.70	1.64	.060	2.12	0.036
Observations	145	66			

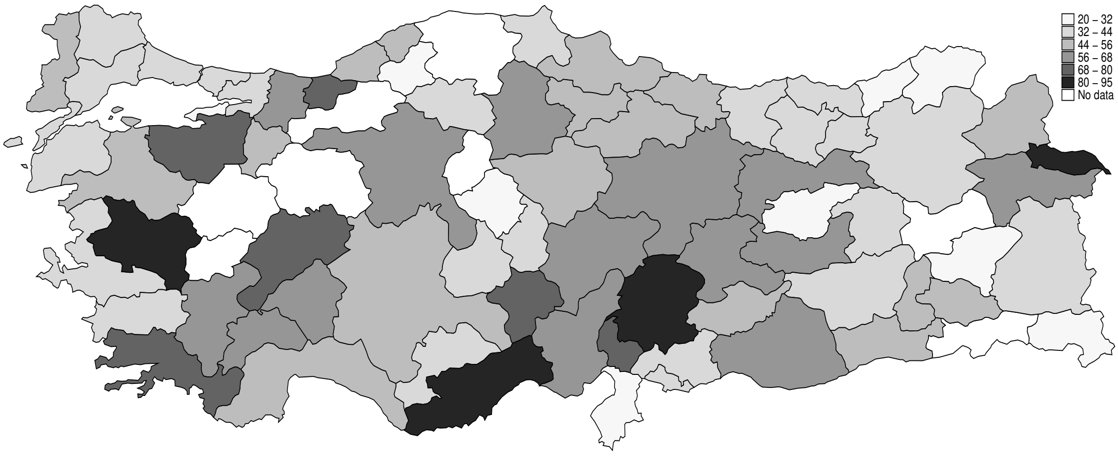
Daily air pollution data is taken from the database of Air Quality Monitoring Network, created by Ministry of Environment and Urbanization[6]. This website collects the data for air quality and pollution levels from several air quality monitoring stations all over the country and creates the daily and hourly air quality database at province and station level. In this study, the available daily pollution data for PM_{10} and SO_2 is gathered at province level. On the basis of WHO Air Quality Guidelines (WHO, 2005), the number of days in a year that the threshold was exceeded are calculated where the threshold levels for the pollution is set such that the 24-hour mean for PM_{10} should not exceed $50\mu g/m^3$. Then, this number was divided by the number of observations in that year in order to deal with the missing data on daily

[5] However, these differences might be driven by other socio-demographic characteristics of these cities that are potentially correlated with air pollution. Therefore, this simple comparison should not be taken as an evidence for the causal effect of air pollution.

[6] `https://sim.csb.gov.tr` -Last accessed to the online database on January 2, 2020.

pollution. Highly polluted provinces are the ones that exceed the mean for observed PM_{10} levels above the 24-hour limits by one standard deviation. On the other hand, the provinces in the group with lower levels of PM_{10} are the ones with one standard deviation below the mean for observed PM_{10} levels above the 24-hour limits. However, in Turkey, the data gathered from the air quality monitoring stations suffer severely from the inefficient calculations. In order for these stations to be used in the analysis, it is required that the calculations are made more than 75 percent of the year. Considering this, in Turkey, only 77 percent of the stations collected efficient data in 2018 and of these stations, 96.3 percent exceeded WHO PM_{10} guideline limits, 59.5 percent exceeded national limits (THH, 2019). In figure, the PM_{10} levels of each provinces are illustrated. It should be noted that the national limit for PM_{10} pollution was $44\mu g/m^3$ and WHO suggested not exceeding $20\mu g/m^3$ in 2018. There are 8 provinces such that in those provinces the measurements were not observed for at least 75 percent of the year.

Figure 3.7 PM_{10} levels, 2018



Source: (THH, 2019)

The following figures (Figure 3.8 and 3.9 demonstrate per capita visits to FPCs in 2012 and 2018, separately. It is observed that the more one moves towards East, the less people consult the healthcare facilities. This observation can be attributed to socio-economic development levels of the provinces and the regions. Also, in Aegean region, we see that consultations to the physicians have decreased. The visits to hospitals and all levels of healthcare facilities can be traced in Appendix A.1, A.2, A.3, A.4. The assumption is that if a province adopts natural gas densely, then it is expected to have less consultations to the healthcare facilities, specifically to FPCs. Although it seems as if there is no correlation between mean natural gas density and mean visits to healthcare facilities, it is expected that once accounted for the province fixed effects and time variant characteristics of provinces, the causation follows.

Figure 3.8 Visits to Family Physician Centers, 2012

Figure 3.9 Visits to Family Physician Centers, 2018

3.2.3 Control variables

The vector of time-variant variables is chosen in order to capture the differences of province characteristics which may be correlated with the use of health care services. The vector of controls include GDP per capita to account for province's economic development; number of hospitals, number of physicians and number of hospital beds to account for the advancement of healthcare provision, also the number of family physician centers control for the increased availability of family physician centers over time, number of automobiles, unemployment rate and population level as indicators of development, and finally, student per teacher in secondary education, the percentage of population with high school and college representing educational development levels. All of these control variables are provided at province level per year except the unemployment rate which is available at Nuts2 region level.[7]

The data for variables related to health; number of hospitals, number of physicians and number of hospital beds are taken from Yearly Statistical Books of Ministry of Health that starts from 2000. The number of Family Physican Centers are also taken from Yearly Statistical Books of Ministry of Health and the data starts from 2012 for the reason that the change in the healthcare provisions which establishes Family Physician Practice, came into force in 2010 throughout the country. The data for GDP per capita and number of automobiles per 1000 persons start from 2000, unemployment rate at Nuts2 level starts from 2004. The data exists until 2018, the final year which the analysis takes into account. All of the data is taken from TurkStat. Population data comes from TurkStat's Address Based Population Registration System for the years 2007 to 2018. The previous years' population level by province are only available in terms of population projections generated by TurkStat[8]. Lastly, the data for student per teacher in secondary education is taken from the Yearly Statistical Books of Ministry of Education from 2001 to 2018, and the percentage of population with high school and college degrees are taken from TurkStat from 2008 to 2018.

3.2.4 Descriptive Statistics

Table 3.2 shows the means and standard deviations of the province characteristics for 2012-2018 period. In the first column, the means and standard deviations are presented for the full sample. Examining columns 2 and 3 clarifies the differences in means between the provinces with and without natural gas establishment. From the two columns, it can be confirmed that almost every characteristics differ between the two groups. It should also be noted that provinces with natural gas access have higher demand for healthcare provision, higher income per capita levels, better schooling opportunities and higher number of automobiles. All these differences in characteristics could be evidence of a biased natural gas establishment procedure towards wealthier and more developed central urban regions.

[7] There are 26 regions in NUTS2 region level in Turkey, defined in 2002.

[8] Following Cesur et al.'s discussion (Cesur, Tekin, and Ulker, 2017) on accuracy of the projections, the population levels for 2001 to 2006 are simply taken from the population projections.

Moving to columns 4 and 5, although some of the significant differences disappear for the groups by intensity of natural gas utilization[9], in the provinces with high utilization of natural gas, we see on average, less per capita visits to healthcare facilities of any category together with higher income per capita and a lower number of physicians. These are the significant differences between the two groups. It is shown it the Table 3.2 that there is a higher mean per capita visit to any category of healthcare facilities in provinces with natural gas but a lower mean per capita visit in the provinces with higher intensity. The reason of this contradiction could be that the decision of natural gas network expansion is biased. Thus, the expansion of natural gas might be indicating an increase in the socio-economic development in a given province. The wealthier, more urbanized and central, more educated provinces has natural gas establishment, therefore they are more aware of their health and seek help if necessary. Whereas, rising utilization of natural gas might be indicating the outcomes of natural gas use in a longer term and might be concluding that a lower visits to healthcare facilities are conclusion of a better environment overall.

[9]The critical levels to be included in these groups are defined as being one standard deviation below and one standard deviation above the mean intensity for 2012-2018 period.

Table 3.2 Descriptive Statistics

	(1) All	(2) NG Existence =0	(3) NG Existence =1	(4) NG Intensity Lowest	(5) NG Intensity Highest
Per capita Visits to HCF	8.42	6.40	8.66 ***	8.57	7.70 ***
	(1.40)	(1.48)	(1.19)	(1.32)	(1.66)
Per capita Visits to FPCs	2.86	1.98	2.96 ***	3.09	2.53 ***
	(0.80)	(1.02)	(0.71)	(0.90)	(0.94)
Per capita Visits to Hospitals	5.56	4.42	5.70 ***	5.48	5.16 *
	(0.88)	(0.71)	(0.79)	(0.80)	(0.92)
Income per capita (in dollars)	8475.02	6708.50	8684.07 ***	7645.38	9645.24 ***
	(3026.04)	(2212.22)	(3042.74)	(2478.65)	(4415.61)
Hospitals per 1 Million	24.45	28.41	23.98 *	21.47	24.31
	(9.89)	(15.27)	(8.96)	(9.79)	(11.84)
Hospital Beds per 1 Million	2676.71	1896.21	2769.07 ***	2413.35	2336.26
	(857.83)	(448.53)	(847.87)	(570.64)	(820.27)
Physicians per 1 Million	593.32	602.48	592.24	606.01	562.76 **
	(98.04)	(145.09)	(91.03)	(91.08)	(115.08)
FPCs per capita	3337.91	3532.02	3314.94 ***	3448.19	3494.76
	(307.97)	(352.67)	(294.27)	(243.93)	(329.36)
Student per Teacher	13.18	15.00	12.97 **	13.82	13.74
	(2.91)	(4.50)	(2.59)	(3.03)	(3.54)
High School Graduate Rate	21.53	18.86	21.85 ***	20.47	22.18 **
	(3.79)	(4.72)	(3.53)	(3.56)	(4.96)
College Graduate Rate	11.84	9.54	12.11 ***	11.43	12.10
	(2.88)	(2.97)	(2.74)	(2.83)	(3.88)
Unemployment Rate by Region	9.30	9.83	9.24	10.18	10.04
	(4.72)	(6.01)	(4.55)	(4.70)	(4.93)
Automobiles per 1000	109.12	42.19	117.04 ***	102.20	93.91
	(53.07)	(42.56)	(48.43)	(55.35)	(66.57)
Population (1 million)	0.97	0.32	1.05 ***	1.04	1.51
	(1.75)	(0.20)	(1.83)	(0.75)	(3.12)
Observations	567	60	507	68	128

Notes. Standard deviations are in parentheses. *, ** and *** indicate that the mean is statistically different between the sample in columns (2) and (3) or columns (4) and (5) at the 10%, 5% and 1% levels respectively.

4. METHODOLOGY AND RESULTS

4.1 Regression Analysis

4.1.1 Empirical Strategy

The hypothesis of this thesis is that with the increased utilization of natural gas as an energy resource that is cleaner than the previous resources under use, the incidence and prevalence of respiratory diseases connected to the air pollution will decrease[1]. To see if the hypothesis is true, the following functional form is used in the regression analyses.

$$\begin{aligned} LogVisitsFPC_{it} = {} & \alpha + \beta_1 NG_{it} + \beta_2 \mathcal{D}_{it} + \beta_3 \mathcal{H}_{it} + \\ & \theta_t + \gamma_i + \beta_4 \tau_{it} + \beta_5 \tau_{it}^2 + \omega_{it} + \varepsilon_{it} \end{aligned} \tag{4.1}$$

where $LogVisitsFPC_{it}$ is the logarithm of per capita visits to family physician centers in a given year and province, NG_{it} is a generic variable for different determinants of natural gas in a given year and province. The analysis incorporates the following distinctive variables defined for natural gas and runs regressions for each of them separately: A dummy variable that represents that province i has natural gas infrastructure at time t if it equals 1, a continuous variable that specifies the num-

[1] In this section, only the results for per capita visits to family physician centers are included believing that highly inclusive and low cost healthcare provided by FPCs would derive accurate results that are close to real estimates. To see the results for per capita visits to hospitals or per capita visits to healthcare centers, one can refer to Appendix A.

ber of years after the installation of natural gas pipeline in province i, an intensity measure that represents the level of utilization of natural gas in province i at time t, defined as the natural gas subscribers over population. For robustness check of the model, 1 year lagged dummy variable has been generated in addition to altered versions of intensity measure that have been defined to test whether initial intensity measure is correct. The other components of the regression are: $\mathcal{D}_{it}$ which stands for the time-variant province characteristics for development indicators and $\mathcal{H}_{it}$ that represents the time-variant province characteristics for healthcare provision. θ_t is the set of time fixed effects, γ_i is the set of province fixed effects, τ_{it} and τ_{it}^2 stand for linear and quadratic time trends, respectively. Finally, ω_{it} represents region-by-year fixed effects. The ε_{it} is the error term.

The regression equations should incorporate time-varying province characteristics in order to overcome the omitted variable bias which could have an impact on both healthcare demand by people and natural gas provision. Additionally, the set of year fixed effects that would account for the prevailing unobserved factors between provinces, the set of province fixed effects that would overcome the permanent differences across provinces, linear and quadratic province-specific time trends in order to capture if there is still variation after controlling for the set of province fixed effects and finally, the set of region-by-year fixed effects that would control for any spill overs to neighbouring provinces should be included into regression equations. In this way, the bias of natural gas network expansion towards certain provinces with certain timings would be eliminated. Moreover, unobserved heterogeneity between the provinces would be accounted for.

The results of these regressions are presented in Table 4.1. Since the binary indicator for natural gas is independent of the number of subscribers to the Natural Gas Companies, the coefficient estimate for the dummy variable would measure the magnitude of a permanent change in the visits to FPCs as an impact of natural gas introduction to a given province in a given year. However, the introduction of natural gas does not lead to an immediate utilization by the households, but the utilization occurs over time which may be a result of the network that does not reach every neighbourhood at once or result of cost of subscription and necessary equipment. Therefore, it should be noted that the binary indicator as the independent variable would generate a biased estimate which would not be able to capture the true impact of the treatment accurately. One way to overcome this issue might be using a continuous variable instead of a dummy variable as the independent variable. The intensity for natural gas utilization in a given province at a given time would be an improvement for the analysis. Additionally, as a transitive variable, this thesis also incorporates the number of years after the introduction of natural gas as an independent variable whose coefficient estimate captures the gradual progress over years. If there is an advancement in the respiratory health conditions of the residents in a province due to reduced air pollution, it would be captured by the additional years of access to natural gas provision.

4.1.2 Results

A quick look at the Table 4.1 would give an idea of the results of the regressions. The complete results of the regressions that also include the coefficient estimates for the control variables are presented in Appendix A.

The column 1 of Table 4.1 reports the estimates of the three natural gas indicators when only the set of year fixed effects are included in the regression. It is observed that all the three coefficient estimates are positive and statistically significant, suggesting that the per capita visits to FPCs is 50 percent higher if there is NG access in the province. Also, with one-unit increase in the natural gas intensity measure, per capita visits to FPCs is 50 percent higher. This could be due to the instant positive impact of natural gas establishment on the socio-economic development of the provinces by means of investment or as discussed before, it could be due to the differences between the provinces that are correlated with per capita visits to FPCs. In both ways, it is acceptable for a society in the more developed provinces to care more for their health status and thus consult the professionals more. However, one should keep in mind that the results do not account for province- specific characteristics.

In columns 2 and 3, the inclusion of controls for development and controls for healthcare provision result in loss of significance for the estimate of dummy variable for natural gas existence. However, the significance is kept and the estimate has changed sign for the estimate of years with natural gas and for intensity measure. This means that without the time-variant observable characteristics, the results would be upward biased. Accounting for them suggests that increased utilisation of natural gas cuts the number of per capita visits to FPCs significantly. With one more year natural gas existence, an individual visits FPCs 1.3 percent less, and if the number of subscribers per population increases by 0.1, visits to FPCs decreases by around 6.4 percent. In column 4, we see that with the addition of province-fixed effects, one more year natural gas existence in a province, an individual decreases his visits to FPCs by 7.3 percent (column 4)[2]. The impact of intensity measure becomes positive.

Through the last column with the addition of province-specific linear and quadratic time trends and then controlling for region-by-year fixed effects, the coefficient estimates appear statistically insignificant. This may be due to restricted variation in the sample as a result of short term variation depicted both in FPC visits data and in our intensity measure. Additionally, small sample size in the analysis further restricts fixed effects analysis by raising the standard errors and thus causing insignificant results. However, the negative impact that is captured by the coefficient estimates are consistent with the notion that switching to natural gas from coal as an investment to modern and cleaner energy resource decreases per capita FPCs visits. The estimates show that not only the existence of natural gas network but also its utilization by the residents improve health status of individuals, hence they require less of a consultation to physicians. More precisely, introduction of natural gas in a province decreases the per capita visits by 0.4 percent. Note that the estimate of binary indicator for natural gas existence is likely to be upwards biased because of not accounting for dynamic nature of the treatment, which is gradual expansion of the utilization. Hence, we may expect a larger decline in terms of per capita visits to FPCs. Consistent with the hypothesis and the literature, we conclude that one-unit increase in the intensity, decreases the per capita visits to FPCs by 16 percent. This means that on average, if the number of subscribers per population incerases by 0.1, visits to FPC decreases by around 1.6 percent.

In addition, by referring to Table A.1, one can dig deeper into the analysis and compare the estimates for several development and healthcare provision controls. Despite most of the estimates of controls are not statistically significant in the last column, we observe that the higher number of physicians and the more FPCs are available in a province, the more people consult to FPCs, that is in line with supply sensitive health care theory (The Dartmouth Atlas of Health Care and Practice, 2020). Also, increased number of vehicles significantly increases the per capita FPC visits, by 67 percent. On the development controls, we can conclude that increased student-teacher ratio in primary schools increases per capita FPC visits, increased unemployment rate decreases per capita FPC visits, and increased income per capita increases per capita FPC visits, not surprisingly. Lastly, the estimates for covariates in the last column of Table A.2 are all insignificant and the magnitude and the signs of the estimates reflect immediate observable development characteristics of arrival of a new energy resource into the province.

[2]Because the independent variable is defined in terms of years, accounting for linear or quadratic linear time trends or region-by-year time trends would not be a correct way of doing analysis.

Table 4.1 Coefficient Estimates of Natural Gas on the Logarithm of per capita Visits to FPCs

Log Visits to FPCs							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
VARIABLES							
Dummy NG Existence	0.500***	0.070	0.068	0.079	-0.004	-0.004	-0.004
	(0.138)	(0.054)	(0.047)	(0.058)	(0.038)	(0.038)	(0.111)
Observations	567	567	567	567	567	567	567
R-squared	0.136	0.416	0.431	0.059	0.675	0.675	0.675
Years with NG	0.023***	-0.011**	-0.013***	-0.073***			
	(0.008)	(0.005)	(0.004)	(0.024)			
Observations	567	567	567	567			
R-squared	0.097	0.424	0.440	0.597			
Intensity per capita	0.531*	-0.677**	-0.640**	0.083*	-0.160	-0.160	-0.160
	(0.300)	(0.300)	(0.300)	(0.050)	(0.233)	(0.233)	(0.257)
Observations	413	413	413	413	413	413	413
R-squared	0.070	0.651	0.709	0.509	0.966	0.966	0.966
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls for Development	No	Yes	Yes	Yes	Yes	Yes	Yes
Controls for Healthcare Provision	No	No	Yes	Yes	Yes	Yes	Yes
Province fixed effects	No	No	No	Yes	Yes	Yes	Yes
Province-specific linear time trends	No	No	No	No	Yes	Yes	Yes
Province-specific quadratic time trends	No	No	No	No	No	Yes	Yes
Region-by-year fixed effects	No	No	No	No	No	No	Yes

Notes. Robust standard errors, clustered at the province level, are in parentheses. *, ** or *** indicates significance at the 95%, 99% or 99.9% levels respectively.

4.2 Robustness Check

The robustness check for the analysis by fixed effects model incorporates different measures for natural gas existence and intensity. As the binary indicator for natural gas existence in a province at a given time, one-year lagged dummy variable is 1 if it is past a year or more from the establishment of natural gas infrastructure in the province. By lagging the dummy variable, the dynamic nature of the expansion process shall be accounted for. The second and the third variables for the robustness analysis are different calculations for intensity levels. Because the number of subscribers are in terms of buildings or households, the initial intensity does not capture the actual number of people who use natural gas. However, by multiplying the subscribers with the average household size in a province at time t, the intensity measure reflects the percentage of people who are actually using natural gas[3]. It is derived from the following equation:

$$Intensity_{it} = \frac{NGS_{it} * Householdsize_{it}}{Population_{it}}$$

The second intensity measure is a binary variable defined in order to create a broader variation in the data. By this way, the comparison between the different levels of intensity become more reliable. This variable is defined as dividing intensity levels into two groups, high and low. The decision to assign provinces into one of these groups follow the same fashion with the previous analyses in this thesis. The provinces exceeding average intensity one standard deviation or more are assigned 1 and the ones with one standard deviation less than the average intensity are assigned 0. The intensity measure is defined by the initial equation presented in section 3.2.

Table 4.2 presents the coefficient estimates. It is observed that dummy variable for one-year lagged natural gas existence has a small in magnitude but positive impact on the per capita FPC visits. Intensity measured with household size demonstrate results somewhat close to the results from the main analysis. Accounting only for year fixed effects generate positive significant effect on the per capita FPC visits; however, this impact becomes negative and smaller in size after accounting for control variables, province fixed effects, time trends and region-by-year fixed effects. In can be said that one-unit increase in the intensity decreases per capita FPC visits by 4 percent. The second intensity variable of the robustness demonstrates conflicting results with the previous analysis. Initially, the estimates suggest a declining FPC visits on average for the provinces in high- intensity group. The estimate after province fixed effects are accounted for (column 4) suggests that one-unit increase in the intensity decreases per capita FPC visits by 1.6 percent, which is consistent with the hypothesis, and although less in magnitude, it is consistent with the previous results. The inclusion of time trends and region-by-year fixed effects result in an estimate with insignificant positive impact that is small in magnitude. However, these results are justifiable to some extent because intensity levels do not change significantly over-time for the provinces, thus the groups with high- and low-intensity include the same provinces mostly, hence limited variation can be observed.

---

[3] One drawback of this measure originates from the subscribers' data for that it counts buildings and flats other than household use, as well. Hence, for the industrialized and developed provinces, the intensity measure could exceed 1, because large commercial buildings or official buildings are considered as subscribers as in the case of Ankara and Istanbul.

Table 4.2 Coefficient Estimates of Natural Gas on the Logarithm of per capita Visits to FPCs (Robustness)

Log Visits to FPCs							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
VARIABLES							
1 Year Lagged NG Existence	0.443***	0.018	0.017	-0.008	0.005	0.005	0.005
	(0.132)	(0.056)	(0.048)	(0.051)	(0.094)	(0.094)	(0.158)
Observations	486	486	486	486	486	486	486
R-squared	0.124	0.373	0.386	0.059	0.645	0.645	0.645
Intensity per Household	0.150*	-0.093	-0.079	0.013	-0.040	-0.040	-0.040
	(0.076)	(0.062)	(0.061)	(0.009)	(0.066)	(0.066)	(0.068)
Observations	413	413	413	413	413	413	413
R-squared	0.082	0.631	0.690	0.509	0.966	0.966	0.966
Intensity Level	-0.230**	-0.130***	-0.112**	-0.016	0.009	0.009	0.009
	(0.103)	(0.045)	(0.044)	(0.053)	(0.053)	(0.053)	(0.051)
Observations	196	196	196	196	196	196	194
R-squared	0.083	0.815	0.839	0.945	0.972	0.972	0.971
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls for Development	No	Yes	Yes	Yes	Yes	Yes	Yes
Controls for Healthcare Provision	No	No	Yes	Yes	Yes	Yes	Yes
Province fixed effects	No	No	No	Yes	Yes	Yes	Yes
Province-specific linear time trends	No	No	No	No	Yes	Yes	Yes
Province-specific quadratic time trends	No	No	No	No	No	Yes	Yes
Region-by-year fixed effects	No	No	No	No	No	No	Yes

Notes. Robust standard errors, clustered at the province level, are in parentheses. *, ** or *** indicates significance at the 95%, 99% or 99.9% levels respectively.

4.3 Propensity Score Matching

4.3.1 Empirical Strategy

This section of the thesis attempts to estimate the mean effect of utilizing natural gas. The treatment is defined such that if a province has a level of natural gas utilization above the mean by one standard deviation in a given year, then that province is in treatment group. Contrariwise, if a province has utilization levels of one standard deviation below the mean in a given year, then this province is untreated. Furthermore, the mean level of natural gas intensity and standard deviation that are used to define the treatment are calculated separately for each year. Since natural gas expansion over the country and consumers' choice to utilize it are not controlled experiments, to overcome the selection bias that would lead to calculation of a biased treatment effect, propensity score matching conditional on covariates is employed as the evaluation methodology.

Suppose that D_i represents the binary variable for treatment status. $D_i = 1$ if the province i is treated, say the percentage of subscribers to natural gas companies is high. Accordingly, $D_i = 0$ stands for if the province i is not treated, meaning that is has a low percentage of subscribers to natural gas companies. Denote the outcome conditional on being treated that is, the logarithm of per capite visits to FPCs, as Y_1 and the outcome conditional on not being treated as Y_0. Then the main question that investigates how much difference does natural gas use make on the outcome of treated group compared to outcome of theirs if they would have not been treated could be estimated by Average Treatment Effect on the Treated (ATET):

$$\begin{aligned} ATET &= E(Y_{i1} - Y_{i0} | D_i = 1) \\ & E[Y_{i1} | D_i = 1 - E[Y_{i0} | D = 1] \end{aligned} \tag{4.2}$$

Because we have data on $E[Y_{i1}|D_i = 1]$ but not on $E[Y_{i0}|D = 1]$, we shall follow Rubin's solution (Rubin, 1977) of conditioning the mean of potential outcome on the set of observable characteristics $\mathcal{X}$, assuming that after applying the procedure, outcomes are conditionally mean independent of the treatment. Then the equation for ATET becomes:

$$\begin{aligned} ATET &= E(Y_{i1} - Y_{i0}|D_i = 1, \mathcal{X}) \\ & E[Y_{i1}|D_i = 1, \mathcal{X}] - E[Y_{i0}|D = 1, \mathcal{X}] \end{aligned} \tag{4.3}$$

However, there still may be systematic differences with regard to the outcomes of the treated and untreated groups due to the selection process that might be affected by the set of unobserved characteristics (Heckman, Ichimura, and Todd, 1997). To deal with this problem, it is suggested to use conditional difference-in-differences matching estimator which compares the difference of the conditional outcomes of treated group before and after treatment with the difference of the conditional outcomes of untreated group.(Heckman, Ichimura, and Todd, 1997) Further, instead of conditioning on $\mathcal{X}$, conditioning on propensity scores is suggested to deal with the selection bias into the treatment due to the set of unobserved characteristics (Rosenbaum and Rubin, 1983). Propensity score $p(\mathcal{X})$ stands for the probability of receiving the treatment for a given province i with a set of observable characteristics $\mathcal{X} = x_i$ such that:

$$p(\mathcal{X}) \equiv Pr(D_i = 1|\mathcal{X} = x_i) \tag{4.4}$$

In this thesis, the propensity score matching is employed and propensity scores are estimated by using a logit model that assumes the assignments to the treatment is endogenous. The estimated propensity scores are used to construct the treatment and control groups for a caliper matching with replacement that creates a balanced sample to be used. Matching is made with replacement due to small number of observations in the untreated group in order to keep the bias low, even though matching with replacement generates a larger variance.

Table 4.3 demonstrates the bias in the means of the potential parameter estimates for the logit model. We see that the least biased are unemployment rate by region and logarithm of student per teacher, logarithm of college graduate rate, logarithm of FPCs per 1 million and logarithm of hospital beds per 1 million. By looking at the Table 4.3, we also cannot reject the null hypothesis that the means of these variables for the treatment and control group are statistically not different from each other, meaning the difference is zero. For rest of the variables, the bias gets higher in the magnitude and we conclude that the means are significantly different between each group for each variable.

In order to create a well balanced sample, it is important to decide on the parameter estimates for the logit model that generates the conditional propensity scores. The parameter estimates can be decided by checking the differences in the means, as well as differences in their medians. In the light of bias in the means (Table 4.3), balancing of the medians of potential parameter estimates before and after the adjustment are presented in Table 4.4. In the columns of unadjusted sample, we see that the ones for which the null hypothesis can be rejected are logarithm of income per capita, logarithm of physicians per 1 million and high school graduate rate. That is, we reject the null hypothesis that states the difference in their medians are equal to zero. Hence, we conclude that the three variables' medians significantly differ between the treated and untreated groups for having p-values of 0.011 and 0.001 and 0.006 respectively. Hence, they potentially affect the odds of receiving the treatment.

Table 4.3 Bias in the Means

	Mean			t-test		
Variable	Treated	Control	%bias	t	p>t	V(T)/V(C)
Log Income per capita	9.0729	8.8873	46.3	2.96	0.003	1.76*
Log Hospitals per 1 Million	3.0995	2.9717	30.2	2.03	0.043	0.88
Log Hospital Beds per 1 Million	7.7003	7.764	-22.6	-1.43	0.155	2.27*
Log Physicians per 1 Million	6.3133	6.3969	-49	-3.11	0.002	2.00*
Log FPCs per 1 Million	21.97	21.959	13.8	0.88	0.379	1.77*
Log Automobiles per 1000	4.1366	4.4248	-32.3	-2.04	0.043	2.16*
Log Student per Teacher	2.5888	2.6039	-6.5	-0.42	0.673	1.43*
High School Graduate Rate	22.182	20.475	39.6	2.51	0.013	1.94*
College Graduate Rate	12.098	11.429	19.7	1.26	0.211	1.87*
Unemployment Rate by Region	10.044	10.175	-2.7	-0.18	0.857	1.1

* if variance ratio outside [0.71; 1.42]

Ps	LR chi2	p>chi2	Mean Bias	Median Bias	B	R	%Var
0.301	76.21	0.000	26.3	26.4	144.0*	1.68	80

* if B>25%, R outside [0.5; 2]

In the light of above discussion, the logit model takes the 5-knotted restricted cubic splines of logarithm of income per capita, logarithm of physicians per 1 million and logarithm of high school graduate rate as the parameter estimates. In Table 4.4, it can be seen that the differences in medians of each variable are equal to zero in the columns for adjusted sample (caliper matching with caliper=0.25). Now, with 29 observations in untreated and 62 observations in treated group, the sample is well balanced to move on to caliper matching with replacement.

Table 4.4 Test for Differences in Medians

	Unadjusted			Adjusted (c=0.25)		
	(1)	(2)	(3)	(4)	(5)	(6)
Factor	Treatment = 0	Treatment = 1		Treatment = 0	Treatment = 1	
	Median	Median	p-value	Median	Median	p-value
Log Income per capita	8.9 (8.8, 9.1)	9.1 (8.7, 9.4)	0.011	8.9 (8.6, 9.2)	8.8 (8.6, 9.1)	0.76
Log Hospitals per 1 Million	2.9 (2.7, 3.3)	3.0 (2.8, 3.3)	0.053	3.0 (2.8, 3.3)	3.2 (2.9, 3.4)	0.17
Log Hospital Beds per 1 Million	7.8 (7.6, 7.9)	7.7 (7.4, 7.8)	0.096	7.8 (7.6, 7.8)	7.7 (7.4, 7.8)	0.31
Log Physicians per 1 Million	6.4 (6.3, 6.5)	6.3 (6.2, 6.4)	0.001	6.4 (6.3, 6.5)	6.4 (6.3, 6.5)	0.77
Log FPCs per 1 Million	22.0 (21.9, 22.0)	22.0 (21.9, 22.0)	0.36	22.0 (21.9, 22.0)	22.0 (21.9, 22.0)	0.34
Log Student per Teacher	2.6 (2.5, 2.7)	2.6 (2.4, 2.8)	0.77	2.6 (2.5, 2.8)	2.6 (2.5, 2.8)	0.34
High School Graduate Rate	21.3 (18.1, 23.2)	23.6 (17.3, 26.1)	0.006	21.5 (17.8, 23.2)	19.1 (16.6, 23.5)	0.35
College Graduate Rate	11.4 (9.4, 13.3)	11.8 (9.1, 14.7)	0.31	11.4 (8.2, 12.6)	10.6 (8.0, 12.0)	0.32
Log Automobiles per 1000	4.7 (3.9, 4.9)	4.7 (3.2, 5.0)	0.29	4.5 (3.4, 4.9)	4.2 (3.1, 4.9)	0.26
Unemployment Rate by Region	8.6 (6.8, 12.2)	9.0 (7.2, 10.7)	0.96	9.8 (6.8, 12.2)	7.7 (6.3, 9.6)	0.058
N	68	128		29	62	

4.3.2 Results

The results from initial matching procedure for several calipers[4] are presented in Table 4.5. We see that in the unmatched sample the treated group experience 24 percent less per capita visits to FPCs and it is statistically significant. However, the differences in the average treatment effect on the treated between the two groups is captured as approximately 11 percent but the significance is lost. Nevertheless, one should note that these results lack controlling for covarites which would enhance the

[4]The initial choice for caliper is 0.25 as a rule of thumb. The decreasing width of the caliper band is displayed as a robustness check. The increased width of the caliper band aims to catch one more observation on common support in treated group because the number of observations are small.

liability of the propensity score method.

Table 4.5 Matching and ATET

Log Visits to FPCs	(1)	(2)	(3)	(4)	(5)	(6)
c=0.15	Sample	Treated	Controls	Difference	S.E.	t-stat
	Unmatched	0.839	1.078	-0.239	0.064	-3.74
	ATET	0.830	0.940	-0.110	0.093	-1.18
	Treatment	Off Support	On Support	Total		
	Untreated	0	68	68		
	Treated	71	57	128		
	Total	71	125	196		
c=0.2	Sample	Treated	Controls	Difference	S.E.	t-stat
	Unmatched	0.839	1.078	-0.239	0.064	-3.74
	ATET	0.797	0.907	-0.110	0.098	-1.12
	Treatment	Off Support	On Support	Total		
	Untreated	0	68	68		
	Treated	67	61	128		
	Total	67	129	196		
c=0.25	Sample	Treated	Controls	Difference	S.E.	t-stat
	Unmatched	0.839	1.078	-0.239	0.064	-3.74
	ATET	0.788	0.911	-0.123	0.098	-1.26
	Treatment	Off Support	On Support	Total		
	Untreated	0	68	68		
	Treated	66	62	128		
	Total	66	130	196		
c=0.35	Sample	Treated	Controls	Difference	S.E.	t-stat
	Unmatched	0.839	1.078	-0.239	0.064	-3.74
	ATET	0.797	0.916	-0.119	0.098	-1.21
	Treatment	Off Support	On Support	Total		
	Untreated	0	68	68		
	Treated	65	63	128		
	Total	65	131	196		

Note: S.E. does not take into account that the propensity score is estimated.

The estimates of the caliper matching with replacement are presented in column 1 of Table 4.6. Although it may seem enough to replicate a randomized controlled experiment with observable data, potential bias may occur in the ATET estimates due to correlation between unobserved characteristics and either the treatment variable or the per capita visits to FPCs or both. In order to avoid such bias and prevent residual confounding or endogeneity, controls for healthcare provision and controls for development measures are included into analyses through columns 2 to 4. We see that without controlling for observable characteristics, the effect of the treatment on the treated is negative and significant except for caliper 0.15. The intense use of natural gas in a province reduces the per capita visits to FPCs by 18 percent, if the caliper is taken equal to 0.25. Through column 2 to 3, although the effect varies in the magnitude by addition of healthcare provision and income per capita controls, the sign of ATET estimates are still negative and they are significant. In the last column (column 4), the controls for development characteristics are included into the analysis and it is observed that the impact is much smaller compared to what was observed in columns 2 and 3, yet still significant and negative. Namely, if we take caliper equal to 0.25, if natural gas is used extensively in a province, a person living in this province would go to FPCs 16 percent less than the ones who live in the provinces where there is almost no natural gas use or it is utilised by a few.

Table 4.6 Coeffecient Estimates of Caliper Matching with Replacement

	(1)	(2)	(3)	(4)
VARIABLES	Log Visits to FPCs	Log Visits to FPCs	Log Visits to FPCs	Log Visits to FPCs
c=0.15				
Treatment	-0.138	-0.175*	-0.206***	-0.144**
	(0.092)	(0.092)	(0.054)	(0.057)
Observations	86	86	86	86
R-squared	0.026	0.438	0.823	0.918
c=0.20				
Treatment	-0.171*	-0.192**	-0.225***	-0.162***
	(0.098)	(0.091)	(0.054)	(0.057)
Observations	90	90	90	90
R-squared	0.034	0.521	0.838	0.924
c=0.25				
Treatment	-0.180*	-0.214**	-0.241***	-0.159***
	(0.098)	(0.091)	(0.055)	(0.058)
Observations	91	91	91	91
R-squared	0.037	0.514	0.832	0.922
c=0.35				
Treatment	-0.171*	-0.217**	-0.244***	-0.162***
	(0.098)	(0.094)	(0.057)	(0.059)
Observations	92	92	92	92
R-squared	0.033	0.485	0.822	0.920
Controls for Healthcare Provision	No	Yes	Yes	Yes
Control for GDP	No	No	Yes	Yes
Controls for Development	No	No	No	Yes

Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

5. CONCLUSION AND DISCUSSION

According to the WHO, 15 million of all deaths in the world are premature deaths related to NCDs, of these 15 million deaths, 85 percent are recorded in low- and middle-income countries (WHO, 2018*a*). In Turkey in 2013, it is reported that 88 percent of deaths are related to NCDs of which 17 percent were premature (Hacettepe University Faculty of Medicine, 2017; WHO, 2017*a*). The risk factors of NCDs consist of physical inactivity, unhealthy diet and the excessive consumption of alcohol and tobacco together with both ambient and household air pollution that are recognised as risk factors, recently (UN General Assembly, 2018).

The relation between air pollution and health outcomes has long been studied. It is found that CO is the most dangerous pollutant for infants, leading to low birth weight and mortality even at relatively low levels of air pollution (Coneus and Spiess, 2012; Currie and Neidell, 2005; Currie, Neidell, and Schmieder, 2009), both PM_{10} and SO_2 are also associated with infant mortality (Arceo, Hanna, and Oliva, 2016; Luechinger, 2014) and high O_3 levels increases the prevalence of respiratory illnesses (Beatty and Shimshack, 2014; Coneus and Spiess, 2012). Long term exposure to air pollution is as important as short term exposure. It is found that exposure to high PM, O_3 and CO levels increase both mortality and morbidity rates related to NCDs (Evans and Smith, 2005; Nhung et al., 2018; Romieu et al., 2012).

According to OECD, fossil fuels represent 88 percent of total primary energy supply in Turkey (OECD, 2019). The coal use for energy supply creates emissions of PM, SO_2, NO_x, CO_2 and other pollutants which are listed as some of the most dangerous pollutants to health (EEA, 2019). Consequently, pollutant concentrations are recorded above the WHO guidelines (WHO, 2005) more than a half of the year in most of the regions and 90 percent of the air quality measurement stations failed to report measurements (TMMOB, 2019). Hence, gradual expansion of natural gas across country that started from late 1980 and completed in 2019 and substitutes coal in heating and cooking purposes is studied as the link between natural gas use and health outcomes are explored. Adding the results of two previous studies (Cesur, Tekin, and Ulker, 2017, 2018), it is found that a 0.4 percent decrease in the per capita visits to FPCs occur if a province gets access to natural gas pipelines and one-unit increase in the natural gas utilization results in a decrease of 16 percent in the per capita FPCs visits between 2012 and 2018. Although the magnitude of the estimate becomes smaller, the second intensity measure concludes consistent results such that one-unit increase in the intensity decreases per capita FPC visits by 4 percent. This means that 0.1 increase in the intensity measure decreases FPC visits by 1.6 or 0.4 percent, depending on which intensity measure you are using. On the other hand, propensity score matching suggest that the intense adoption of natural gas in provinces decreases per capita FPC visits by 16 percent compared to the provinces natural gas adoption is hardly ever existent.

Building an infrastructure of a cleaner energy source is necessary to acquire better health outcomes yet it is not sufficient. It should be available in each neighbourhood and available to every household from all socio-economic background. The more people benefit from the infrastructure, the healthier the society gets. Hence, regulatory bodies should embolden the adoption of this new resource either by subsidizing or informing potential users about its benefits. It is important that policy-makers impose effective policies to encourage people to use cleaner resource. They should collaborate and cooperate even with sectors outside health at national, regional and global levels (WHO, 2013).

This study sets a valid example for developing countries who are thinking of investing in a large scale energy improvement. Governments can consider not only the energy resource security and economic advantages that are stimulated from energy markets, but also positive health outcomes of such investments help mitigate cost of healthcare provision or lost workdays due to illnesses. Additionally, with increased quality of life, the government would better off for having benefits exceeding the cost of investment. Although the results of this study are applicable to other developing countries that invest in a cleaner energy source, a disadvantage of this thesis is the data restrictions of health variables. Because the data does not provide information on the reason for visiting the healthcare facilities, on the age and gender of the patients or on socioeconomic background of the patients we cannot conclude any impact of natural gas use on these groups of people. Also, the data in this study is relative small. Small sample may cause variation loss and thus concluding biased results. Therefore, further research would be beneficiary to understand the impact of expansion of natural gas on the specific health outcomes in the long run.

BIBLIOGRAPHY

Anderson, H R, S A Bremner, R W Atkinson, R M Harrison, and S Walters. 2001. "Particulate matter and daily mortality and hospital admissions in the west midlands conurbation of the United Kingdom: associations with fine and coarse particles, black smoke and sulphate." *Occupational and Environmental Medicine* 58(8): 504–510.

Arceo, Eva, Rema Hanna, and Paulina Oliva. 2016. "Does the Effect of Pollution on Infant Mortality Differ Between Developing and Developed Countries? Evidence from Mexico City." *The Economic Journal* 126(01): 257–280.

Armstrong, Joanna R M, and Harry Campbell. 1991. "Indoor AirPollution Exposure and Lower Respiratory Infections in Young Gambian Children." *International Journal of Epidemiology* 20(06): 424–429.

Beatty, Timothy K.M., and Jay P. Shimshack. 2014. "Air pollution and children's respiratory health: A cohort analysis." *Journal of Environmental Economics and Management* 67(1): 39 – 57.

Beltramo, Theresa, and David I. Levine. 2013. "The effect of solar ovens on fuel use, emissions and health: results from a randomised controlled trial." *Journal of Development Effectiveness* 5(2): 178–207.

Bensch, Gunther, and Jörg Peters. 2015. "The intensive margin of technology adoption - Experimental evidence on improved cooking stoves in rural Senegal." *Journal of Health Economics* 42: 44 – 63.

Bickerstaff, Karen, and Gordon Walker. 1999. "Clearing the smog? Public responses to air quality information." *Local Environment* 4(3): 279–294.

Bremner, Stephen, H Anderson, R Atkinson, A McMichael, D Strachan, J Bland, and J Bower. 1999. "Short term associations between outdoor air pollution and mortality in London 1992-4." *Occupational and environmental medicine* 56(05): 237–44.

Bresnahan, Brian W., Mark Dickie, and Shelby Gerking. 1997. "Averting Behavior and Urban Air Pollution." *Land Economics* 73(3): 340–357.

Burnett, R.T., R.E. Dales, M.E. Raizenne, D. Krewski, P.W. Summers, G.R. Roberts, M. Raadyoung, T. Dann, and J. Brook. 1994. "Effects of Low Ambient Levels of Ozone and Sulfates on the Frequency of Respiratory Admissions to Ontario Hospitals." *Environmental Research* 65(2): 172 – 194.

Cesur, Resul, Erdal Tekin, and Aydogan Ulker. 2017. "Air Pollution and Infant Mortality: Evidence from the Expansion of Natural Gas Infrastructure." *The Economic Journal* 127(600): 330–362.

Cesur, Resul, Erdal Tekin, and Aydogan Ulker. 2018. "Can natural gas save lives? Evidence from the deployment of a fuel delivery system in a developing country." *Journal of Health Economics* 59(04).

Chay, Kenneth, Carlos Dobkin, and Michael Greenstone. 2003. "The Clean Air Act of 1970 and Adult Mortality." *Journal of Risk and Uncertainty* 27(3): 279–300.

Chay, Kenneth Y., and Michael Greenstone. 2003. "The Impact of Air Pollution on Infant Mortality: Evidence from Geographic Variation in Pollution Shocks Induced by a Recession." *The Quarterly Journal of Economics* 118(3): 1121–1167.

Coneus, Katja, and C. Katharina Spiess. 2012. "Pollution exposure and child health: Evidence for infants and toddlers in Germany." *Journal of Health Economics* 31(1): 180 – 196.

Cropper, Maureen L., Nathalie B. Simon, Anna Alberini, Seema Arora, and P. K. Sharma. 1997. "The Health Benefits of Air Pollution Control in Delhi." *American Journal of Agricultural Economics* 79(5): 1625–1629.

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

Currie, Janet, and Reed Walker. 2011. "Traffic Congestion and Infant Health: Evidence from E-ZPass." *American Economic Journal: Applied Economics* 3(1): 65–90.

Currie, Janet, Matthew Neidell, and Johannes F. Schmieder. 2009. "Air pollution and infant health: Lessons from New Jersey." *Journal of Health Economics* 28(3): 688 – 703.

Cyrys, Josef, Annette Peters, Jens Soentgen, and H.-Erich Wichmann. 2014. "Low emission zones reduce PM10 mass concentrations and diesel soot in German cities." *Journal of the Air & Waste Management Association* 64(4): 481–487.

Díaz, Esperanza, Tone Smith-Sivertsen, Daniel Pope, Rolv Lie, Anaité Díaz Artiga, John McCracken, Byron Arana, Kirk Smith, and Nigel Bruce. 2007. "Eye discomfort, headache and back pain among Mayan Guatemalan women taking part in a randomised stove intervention trial." *Journal of epidemiology and community health* 61(02): 74–9.

EEA. 2019. "Air quality in Europe — 2019 report." Luxembourg: European Union.

European Commission and European Parliament. 2018. "Eurobarometer 88.1 (2017); Parlemeter 2017, Cultural Heritage, Future of Europe, Attitudes of European citizens towards the environment.".

Evans, Mary F., and V. Kerry Smith. 2005. "Do new health conditions support mortality–air pollution effects?" *Journal of Environmental Economics and Management* 50(3): 496 – 518.

Fusco, D., F. Forastiere, P. Michelozzi, T. Spadea, B. Ostro, M. Arcà, and C.A. Perucci. 2001. "Air pollution and hospital admissions for respiratory conditions in Rome, Italy." *European Respiratory Journal* 17(6): 1143–1150.

Gehrsitz, Markus. 2017. "The effect of low emission zones on air pollution and infant health." *Journal of Environmental Economics and Management* 83: 121 – 144.

Greenstone, Michael, and Rema Hanna. 2014. "Environmental Regulations, Air and Water Pollution, and Infant Mortality in India." *The American Economic Review* 104(10): 3038–3072.

Hacettepe University Faculty of Medicine. 2017. "Ulusal Hastalık Yuku Calismasi Sonuclari ve Cozum Onerileri.".

Hanna, Rema, Esther Duflo, and Michael Greenstone. 2016. "Up in Smoke: The Influence of Household Behavior on the Long-Run Impact of Improved Cooking Stoves." *American Economic Journal: Economic Policy* 8(1): 80–114.

HEAL. 2017. "Hidden Costs.".

Heckman, James, Hidehiko Ichimura, and Petra E. Todd. 1997. "Matching As An Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Programme." *Review of Economic Studies* 64(4): 605–654.

House of Commons. 2018. "Improving air quality." Environment, Food and Rural Affairs, Environmental Audit, Health and Social Care, and Transport Committees.

Janke, Katharina. 2014. "Air pollution, avoidance behaviour and children's respiratory health: Evidence from England." *Journal of Health Economics* 38: 23 – 42.

Jayachandran, Seema. 2009. "Air Quality and Early-Life Mortality: Evidence from Indonesia's Wildfires." *The Journal of Human Resources* 44(4): 916–954.

Knittel, Christopher R., Douglas L. Miller, and Nicholas J. Sanders. 2016. "Caution, drivers! Children present: traffic, pollution, and infant health.(Author abstract)." 98(2): 350.

Lleras-Muney, Adriana. 2010. "The Needs of the Army: Using Compulsory Relocation in the Military to Estimate the Effect of Air Pollutants on Children's Health." *The Journal of Human Resources* 45(3): 549–590.

Luechinger, Simon. 2014. "Air pollution and infant mortality: A natural experiment from power plant desulfurization." *Journal of Health Economics* 37: 219 – 231.

Maddison, David. 2005. "Air pollution and hospital admissions: an ARMAX modelling approach." *Journal of Environmental Economics and Management* 49(1): 116 – 131.

McDermott, Marc, Rajendu Srivastava, and Sarah Croskell. 2006. "Awareness of and Compliance with Air Pollution Advisories: A Comparison of Parents of Asthmatics with Other Parents." *Journal of Asthma* 43(3): 235–239.

Ministry of Industry and Technology. 2019. "Illerin ve Bolgelerin Sosyo-ekonomik Gelismislik Siralamasi Arastirmasi: SEGE 2017.".

Mitter, Sumeet S., Rajesh Vedanthan, Farhad Islami, Akram Pourshams, Hooman Khademi, Farin Kamangar, Christian C. Abnet, Sanford M. Dawsey, Paul D. Pharoah, Valentin Fuster, Paolo Boffetta, and Reza Malekzadeh. 2016. "Household fuel use and cardiovascular disease mortality: Golestan cohort study.(Report)." 133(24): 2360–10.

Moesgaard Iburg, Kim, and GBD 2015 Risk Factors Collaborators. 2016. "Global, regional, and national comparative risk assessment of 79 behavioral, environmental and occupational, and metabolic risks or clusters of risks, 1990–2015: a systematic analysis for the Global Burden of Disease Study 2015." *Lancet* 388(10053): 1659.

Moretti, Enrico, and Matthew Neidell. 2011. "Pollution, Health, and Avoidance Behavior: Evidence from the Ports of Los Angeles." *The Journal of Human Resources* 46(1): 154–175.

Morfeld, Peter, David Groneberg, and Michael Spallek. 2014. "Effectiveness of Low Emission Zones: Large Scale Analysis of Changes in Environmental NO2, NO and NOx Concentrations in 17 German Cities." *PloS one* 9(08): e102999.

Neidell, Matthew. 2009. "Information, Avoidance Behavior, and Health: The Effect of Ozone on Asthma Hospitalizations." *The Journal of Human Resources* 44(2): 450–478.

Neidell, Matthew J. 2004. "Air pollution, health, and socio-economic status: the effect of outdoor air quality on childhood asthma." *Journal of Health Economics* 23(6): 1209 – 1236.

Nhung, Nguyen Thi Trang, Christian Schindler, Tran Minh Dien, Nicole Probst-Hensch, and Nino KÃŒnzli. 2019. "Association of ambient air pollution with lengths of hospital stay for hanoi children with acute lower-respiratory infection, 2007â2016." *Environmental Pollution* 247: 752 – 762.

Nhung, Nguyen Thi Trang, Christian Schindler, Tran Minh Dien, Nicole Probst-Hensch, Laura Perez, and Nino KÃŒnzli. 2018. "Acute effects of ambient air pollution on lower respiratory infections in Hanoi children: An eight-year time series study." *Environment International* 110: 139 – 148.

OECD. 2016. "The economic consequences of outdoor air pollution, Organisation or Economic Co-operation and Development." Paris.

OECD. 2019. "OECD Environmental Performance Reviews: Turkey 2019.".

Ostro, Bart, José Sánchez, Carlos Aranda, and Gunnar Eskeland. 1995. "Air pollution and mortality : results from Santiago, Chile." *Expos. Anal. Environ. Epidemiol.* 6(07).

Pandey, Mrigendra. 1984. "Domestic smoke pollution and chronic bronchitis in a rural community of the Hill Region of Nepal." *Thorax* 39(06): 337–9.

Petroeschevsky, Anna, Rod Simpson, Lukman Thalib, and Shannon Rutherford. 2001. "Associations between Outdoor Air Pollution and Hospital Admissions in Brisbane, Australia." *Archives of environmental health* 56(01): 37–52.

PhD, A. L. Hinwood, N. De Klerk, C. Rodriguez, P. Jacoby, T. Runnion, P. Rye, L. Landau, F. Murray, M. Feldwick, and J. Spickett. 2006. "The relationship between changes in daily air pollution and hospitalizations in Perth, Australia 1992â1998: A case-crossover study." *International Journal of Environmental Health Research* 16(1): 27–46.

Ponce de Leon, A, H R Anderson, J M Bland, D P Strachan, and J Bower. 1996. "Effects of air pollution on daily hospital admissions for respiratory disease in London between 1987-88 and 1991-92." *Journal of Epidemiology & Community Health* 50(Suppl 1): s63–s70.

Pope, C., Richard Burnett, George Thurston, Michael Thun, Eugenia Calle, Daniel Krewski, and John Godleski. 2004. "Cardiovascular Mortality and Long-Term Exposure to Particulate Air Pollution Epidemiological Evidence of General Pathophysiological Pathways of Disease." *Circulation* 109(02): 71–7.

Pope III, C. Arden, Richard T. Burnett, Michael J. Thun, Eugenia E. Calle, Daniel Krewski, Kazuhiko Ito, and George D. Thurston. 2002. "Lung Cancer, Cardiopulmonary Mortality, and Long-term Exposure to Fine Particulate Air Pollution." *JAMA* 287(03): 1132–1141.

Romieu, Isabelle, Nelson Gouveia, Luis Cifuentes, Antonio De Leon, Washington Junger, Jeanette Vera, Valentina Strappa, Hurtado-Díaz Magali, Taha Anwar Taha, Leonora Rojas-Bracho, Luz Carbajal-Arroyo, and Guadalupe Tzintzun-Cervantes. 2012. "Multicity study of air pollution and mortality in Latin America (the ESCALA Study)." *Research report (Health Effects Institute)* 171(10): 5–86.

Rosenbaum, Paul R., and Donald B. Rubin. 1983. "The central role of the propensity score in observational studies for causal effects." *Biometrika* 70(04): 41–55.

Rubin, Donald B. 1977. "Assignment to Treatment Group on the Basis of a Covariate." *Journal of Educational Statistics* 2(1): 1–26.

Schlenker, Wolfram, and W. Reed Walker. 2016. "Airports, Air Pollution, and Contemporaneous Health." *The Review of Economic Studies* 83(2 (295)): 768–809.

Sharma, Sangeeta, Gulshan Rai Sethi, Ashish Rohtagi, Anil Chaudhary, Ravi Shankar, Jawahar Singh Bapna, Veena Joshi, and Debarati Ghua Sapir. 1998. "Indoor Air Quality and Acute Lower Respiratory Infection in Indian Urban Slums." *Environmental Health Perspectives* 106(5): 291–297.

Skov, Torsten, Torben Cordtz, Lilli Kirkeskov Jensen, Peter Saugman, Kirsten Schmidt, and Peter Theilade. 1991. "Modifications of health behaviour in response to air pollution notifications in Copenhagen." *Social Science Medicine* 33(5): 621 – 626.

Smith, Kirk R, John P McCracken, Martin W Weber, Alan Hubbard, Alisa Jenny, Lisa M Thompson, John Balmes, AnaitÃ© Diaz, Byron Arana, and Nigel Bruce. 2011. "Effect of reduction in household air pollution on childhood pneumonia in Guatemala (RESPIRE): a randomised controlled trial." *The Lancet* 378(9804): 1717 – 1726.

Smith-Sivertsen, Tone, Esperanza Díaz, Dan Pope, Rolv T. Lie, Anaite Díaz, John McCracken, Per Bakke, Byron Arana, Kirk R. Smith, and Nigel Bruce. 2009. "Effect of Reducing Indoor Air Pollution on Women's Respiratory Symptoms and Lung Function: The RESPIRE Randomized Trial, Guatemala." *American Journal of Epidemiology* 170(05): 211–220.

Tanaka, Shinsuke. 2015. "Environmental regulations on air pollution in China and their impact on infant mortality." *Journal of Health Economics* 42: 90 – 103.

The Dartmouth Atlas of Health Care, The Dartmouth Institute for Health Policy, and Clinical Practice. 2020. "General Frequently Asked Questions." https://www.dartmouthatlas.org/faq/.

THH. 2019. "Hava Kirliligi ve Saglik Etkileri: Kara Rapor.".

TMMOB. 2019. "Hava Kirliligi Raporu 2019.".

UN General Assembly. 2018. "Political declaration of the third high-level meeting of the General Assembly on the prevention and control of non-communicable diseases. Time to deliver: Accelerating our response to address noncommunicable diseases for the health and well-being of present and future generations.".

WHO. 2005. "WHO Air quality guidelines for particulate matter, ozone, nitrogen dioxide and sulfur dioxide Global update 2005 Summary of risk assessment.".

WHO. 2010. "Global status report on noncommunicable diseases 2010.".

WHO. 2013. "Global action plan for the prevention and control of NCDs 2013-2020, World Health Organization.".

WHO. 2017*a*. "Noncommunicable diseases progress monitor 2017." Geneva: World Health Organization.

WHO. 2017*b*. "Preventing noncommunicable diseases by reducing environmental risk factors." Geneva: World Health Organization; 2017 (WHO/FWC/EPE/17.1). Licence: CC BY-NC-SA 3.0 IGO.

WHO. 2018*a*. "Ambient Air Pollution." taken from https://www.who.int/en/news-room/fact-sheets/detail/ambient-(outdoor)-air-quality-and-health on 4th of July.

WHO. 2018a. "Burden of disease from the joint effects of household and ambient air pollution for 2016, World Health Organization.".

WHO. 2018*b*. "Household Air Pollution." taken from https://www.who.int/en/news-room/fact-sheets/detail/household-air-pollution-and-health on 4th of July.

Wolff, Hendrik. 2014. "Keep Your Clunker in the Suburb: Low-emission Zones and Adoption of Green Vehicles." *The Economic Journal* 124(578): F481–F512.

APPENDIX A

Table A.1 The Impact of Natural Gas Intensity per capita on the Logarithm of per capita Visits to FPCs

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
VARIABLES	Log Visits to FPCs	Log Visits to FPCs	Log Visits to FPCs	Log Visits to FPCs	Log Visits to FPCs	Log Visits to FPCs	Log Visits to FPCs
Intensity per capita	0.531*	-0.677**	-0.640**	0.083*	-0.160	-0.160	-0.160
	(0.300)	(0.300)	(0.300)	(0.050)	(0.233)	(0.233)	(0.257)
Log Income per capita		0.223***	0.271***	-0.036	0.038	0.038	0.038
		(0.078)	(0.077)	(0.176)	(0.261)	(0.261)	(0.153)
Log Automobiles per 1000		0.313***	0.272***	0.024	0.671	0.672	0.672**
		(0.045)	(0.045)	(0.208)	(0.463)	(0.463)	(0.280)
Log Student per Teacher		-0.020	-0.358***	0.006	0.122	0.123	0.123
		(0.127)	(0.097)	(0.068)	(0.094)	(0.094)	(0.082)
High School Graduate Rate		0.004	0.008	-0.012	0.011	0.011	0.011
		(0.007)	(0.007)	(0.009)	(0.013)	(0.013)	(0.013)
College Graduate Rate		-0.023*	-0.037***	-0.026	0.006	0.006	0.006
		(0.012)	(0.012)	(0.020)	(0.032)	(0.032)	(0.025)
Unemployment Rate by Region		-0.006	-0.007**	-0.004	-0.005	-0.005	-0.005*
		(0.003)	(0.003)	(0.005)	(0.007)	(0.007)	(0.003)
Log Hospitals per 1 Million			-0.227***	-0.155**	0.002	0.002	0.002
			(0.058)	(0.076)	(0.075)	(0.075)	(0.066)
Log Hospital Beds per 1 Million			-0.004	0.003	0.034	0.034	0.034
			(0.056)	(0.062)	(0.100)	(0.100)	(0.067)
Log Physicians per 1 Million			0.078	0.213**	0.101	0.101	0.101
			(0.124)	(0.091)	(0.136)	(0.136)	(0.073)
Log FPCs per 1 Million			0.141	0.025	0.106	0.107	0.107
			(0.241)	(0.200)	(0.258)	(0.258)	(0.193)
Constant	1.069***	-2.061***	-4.204	0.393	-5.881	-42.887**	-34.156
	(0.048)	(0.682)	(5.627)	(5.073)	(8.326)	(18.674)	(21.567)
Observations	413	413	413	413	413	413	413
R-squared	0.070	0.651	0.709	0.509	0.966	0.966	0.966
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls for Development	No	Yes	Yes	Yes	Yes	Yes	Yes
Controls for Healthcare Provision	No	No	Yes	Yes	Yes	Yes	Yes
Province fixed effects	No	No	No	Yes	Yes	Yes	Yes
Province-specific linear time trends	No	No	No	No	Yes	Yes	Yes
Province-specific quadratic time trends	No	No	No	No	No	Yes	Yes
Region-by-year fixed effects	No	No	No	No	No	No	Yes

Notes. Robust standard errors, clustered at the province level, are in parentheses. *, ** or *** indicates significance at the 95%, 99% or 99.9% levels respectively.

Table A.2 The Impact of Natural Gas Existence on the Logarithm of per capita Visits to FPCs

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
VARIABLES	Log Visits to FPCs	Log Visits to FPCs	Log Visits to FPCs	Log Visits to FPCs	Log Visits to FPCs	Log Visits to FPCs	Log Visits to FPCs
Dummy for NG existence	0.500***	0.070	0.068	0.079	-0.004	-0.004	-0.004
	(0.138)	(0.054)	(0.047)	(0.058)	(0.038)	(0.038)	(0.111)
Log Income per capita		0.275**	0.348***	-0.027	0.187	0.186	0.186
		(0.121)	(0.126)	(0.388)	(0.278)	(0.278)	(0.633)
Log Automobiles per 1000		0.306***	0.270***	0.354	0.400	0.400	0.400
		(0.049)	(0.047)	(0.552)	(0.383)	(0.383)	(0.952)
Log Student per Teacher		-0.246	-0.360**	-0.211**	-0.022	-0.022	-0.022
		(0.166)	(0.152)	(0.097)	(0.095)	(0.095)	(0.296)
High School Graduate Rate		-0.017	-0.017	-0.059	0.025	0.025	0.025
		(0.013)	(0.014)	(0.055)	(0.023)	(0.023)	(0.045)
College Graduate Rate		-0.018	-0.021	-0.065	-0.085	-0.085	-0.085
		(0.016)	(0.017)	(0.083)	(0.147)	(0.147)	(0.087)
Unemployment Rate by Region		0.001	-0.000	-0.001	-0.002	-0.002	-0.002
		(0.004)	(0.003)	(0.004)	(0.006)	(0.006)	(0.010)
Log Hospitals per 1 Million			-0.170**	-0.087	-0.090	-0.090	-0.090
			(0.065)	(0.083)	(0.151)	(0.151)	(0.262)
Log Hospital Beds per 1 Million			0.003	0.056	0.310	0.310	0.310
			(0.068)	(0.158)	(0.244)	(0.244)	(0.274)
Log Physicians per 1 Million			0.227*	0.017	-0.142	-0.142	-0.142
			(0.119)	(0.119)	(0.213)	(0.213)	(0.252)
Log FPCs per 1 Million			-0.215	-0.137	-0.122	-0.122	-0.122
			(0.226)	(0.134)	(0.157)	(0.157)	(0.263)
Constant	0.613***	-1.675*	1.957	4.770	-0.407	-0.426	17.409
	(0.135)	(0.907)	(4.568)	(5.044)	(4.628)	(29.235)	(63.560)
Observations	567	567	567	567	567	567	567
R-squared	0.136	0.416	0.431	0.059	0.675	0.675	0.675
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls for Development	No	Yes	Yes	Yes	Yes	Yes	Yes
Controls for Healthcare Provision	No	No	Yes	Yes	Yes	Yes	Yes
Province fixed effects	No	No	No	Yes	Yes	Yes	Yes
Province-specific linear time trends	No	No	No	No	Yes	Yes	Yes
Province-specific quadratic time trends	No	No	No	No	No	Yes	Yes
Region-by-year fixed effects	No	No	No	No	No	No	Yes

Notes. Robust standard errors, clustered at the province level, are in parentheses. *, ** or *** indicates significance at the 95%, 99% or 99.9% levels respectively.

Table A.3 The Impact of Years with Natural Gas on the Logarithm of per capita Visits to FPCs

VARIABLES	(1) Log Visits to FPCs	(2) Log Visits to FPCs	(3) Log Visits to FPCs	(4) Log Visits to FPCs
Years with NG	0.023***	-0.011**	-0.013***	-0.073***
	(0.008)	(0.005)	(0.004)	(0.024)
Log Income per capita		0.311***	0.409***	-0.008
		(0.109)	(0.121)	(0.414)
Log Automobiles per 1000		0.368***	0.320***	0.284
		(0.049)	(0.046)	(0.574)
Log Student per Teacher		-0.092	-0.264*	-0.182*
		(0.160)	(0.144)	(0.097)
High School Graduate Rate		-0.017	-0.016	-0.057
		(0.012)	(0.014)	(0.059)
College Graduate Rate		-0.013	-0.019	-0.061
		(0.013)	(0.014)	(0.089)
Unemployment Rate by Region		0.002	0.001	-0.001
		(0.004)	(0.004)	(0.005)
Log Hospitals per 1 Million			-0.197***	-0.086
			(0.062)	(0.091)
Log Hospital Beds per 1 Million			0.063	0.048
			(0.071)	(0.167)
Log Physicians per 1 Million			0.114	0.057
			(0.125)	(0.117)
Log FPCs per 1 Million			-0.256	-0.158
			(0.218)	(0.164)
Constant	0.802***	-2.589***	2.309	6.178
	(0.104)	(0.836)	(4.158)	(5.805)
Observations	567	567	567	567
R-squared	0.097	0.424	0.440	0.597
Year fixed effects	Yes	Yes	Yes	Yes
Controls for Development	No	Yes	Yes	Yes
Controls for Healthcare Provision	No	No	Yes	Yes
Province fixed effects	No	No	No	Yes
Province-specific linear time trends	No	No	No	No
Province-specific quadratic time trends	No	No	No	No
Region-by-year fixed effects	No	No	No	No

Notes. Robust standard errors, clustered at the province level, are in parentheses. *, ** or *** indicates significance at the 95%, 99% or 99.9% levels respectively.

Table A.4 Coefficient Estimates of Natural Gas on the Logarithm of per capita Visits to Hospitals

Log Visits to Hospitals							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
VARIABLES							
Dummy NG Existence	0.224***	0.086***	0.041	0.044***	-0.011	-0.011	-0.011
	(0.044)	(0.031)	(0.029)	(0.014)	(0.018)	(0.018)	(0.018)
Observations	567	567	567	567	567	567	567
R-squared	0.326	0.501	0.607	0.648	0.940	0.940	0.940
Intensity per capita	0.263*	-0.064	-0.009	-0.002	-0.259**	-0.259**	-0.259
	(0.139)	(0.149)	(0.130)	(0.047)	(0.111)	(0.111)	(0.214)
Observations	413	413	413	413	413	413	413
R-squared	0.169	0.343	0.493	0.633	0.928	0.928	0.928
Years with NG	0.010***	0.001	-0.001	-0.026**			
	(0.003)	(0.003)	(0.002)	(0.012)			
Observations	567	567	567	567			
R-squared	0.267	0.486	0.604	0.903			
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls for Development	No	Yes	Yes	Yes	Yes	Yes	Yes
Controls for Healthcare Provision	No	No	Yes	Yes	Yes	Yes	Yes
Province fixed effects	No	No	No	Yes	Yes	Yes	Yes
Province-specific linear time trends	No	No	No	No	Yes	Yes	Yes
Province-specific quadratic time trends	No	No	No	No	No	Yes	Yes
Region-by-year fixed effects	No	No	No	No	No	No	Yes

Notes. Robust standard errors, clustered at the province level, are in parentheses. *, ** or *** indicates significance at the 95%, 99% or 99.9% levels respectively.

Table A.5 Coefficient Estimates of Natural Gas on the Logarithm of per capita Visits to Healthcare Facilities

Log Visits to HCFs							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
VARIABLES							
Dummy NG Existence	0.298***	0.067***	0.039*	0.043**	-0.012	-0.012	-0.012
	(0.060)	(0.024)	(0.022)	(0.016)	(0.015)	(0.015)	(0.016)
Observations	567	567	567	567	567	567	567
R-squared	0.325	0.716	0.770	0.592	0.961	0.961	0.961
Intensity per capita	0.324**	-0.273*	-0.224	0.025	-0.217*	-0.217*	-0.217
	(0.153)	(0.155)	(0.146)	(0.031)	(0.114)	(0.114)	(0.165)
Observations	413	413	413	413	413	413	413
R-squared	0.129	0.565	0.662	0.667	0.956	0.956	0.956
Years with NG	0.013***	-0.003	-0.006***	-0.040***			
	(0.004)	(0.002)	(0.002)	(0.011)			
Observations	567	567	567	567			
R-squared	0.241	0.714	0.779	0.935			
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls for Development	No	Yes	Yes	Yes	Yes	Yes	Yes
Controls for Healthcare Provision	No	No	Yes	Yes	Yes	Yes	Yes
Province fixed effects	No	No	No	Yes	Yes	Yes	Yes
Province-specific linear time trends	No	No	No	No	Yes	Yes	Yes
Province-specific quadratic time trends	No	No	No	No	No	Yes	Yes
Region-by-year fixed effects	No	No	No	No	No	No	Yes

Notes. Robust standard errors, clustered at the province level, are in parentheses. *, ** or *** indicates significance at the 95%, 99% or 99.9% levels respectively.

Figure A.1 Visits to Hospitals, 2012

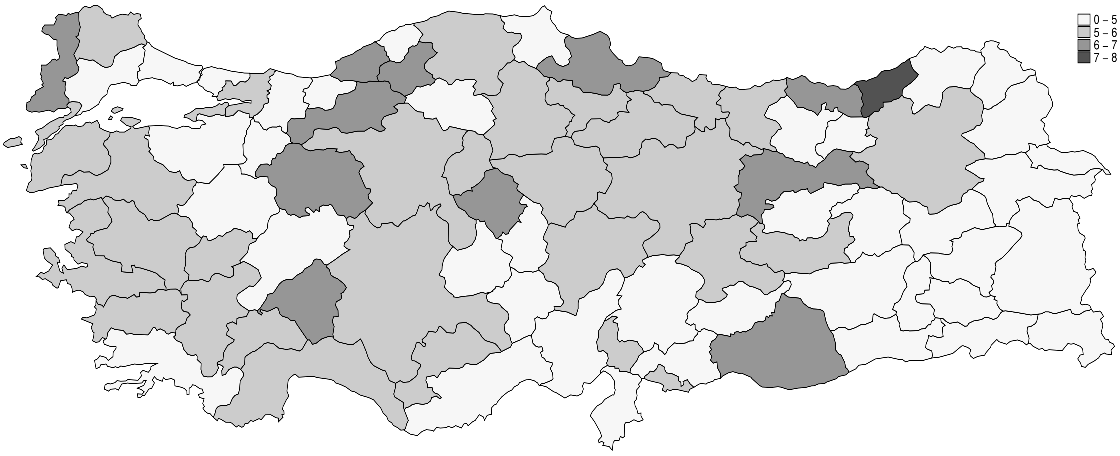


Figure A.2 Visits to Hospitals, 2018

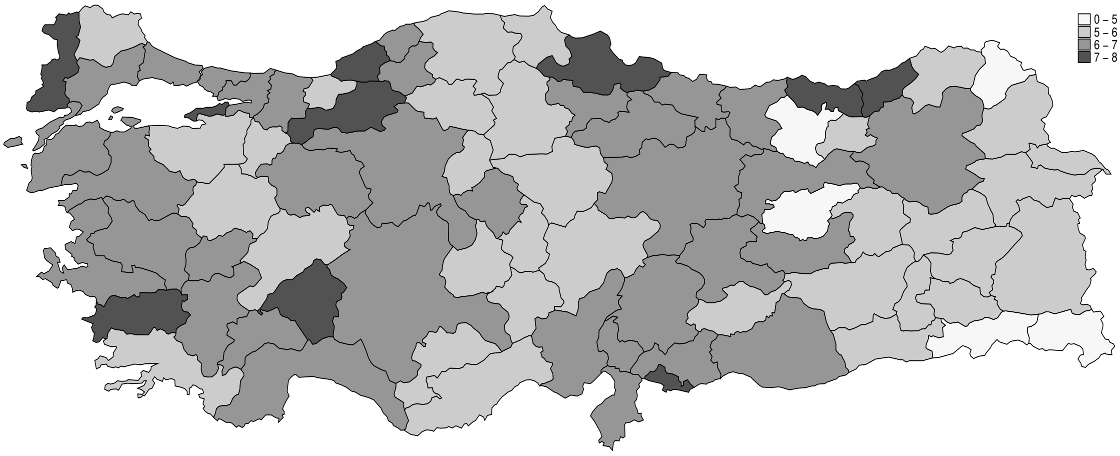


Figure A.3 Visits to Healthcare Facilities, 2012

Figure A.4 Visits to Healthcare Facilities, 2018

APPENDIX B

Can Natural Gas Expansion Be Treated as Random?

This section focuses on the identification strategy for the vector of controls that contains time-varying observable province characteristics. The following exercises aim to understand whether or not there are different trends in the adoption of natural gas between the provinces with respect to the vector of controls. If such characteristic determine the way that natural gas network is constructed, then it leads to a non-random expansion of natural gas. This also may indicate that some of the controls are confounding variables while estimating the impact of natural gas on the healthcare utilization. If there are no significant differences regarding the vector of controls between the provinces that stimulate natural gas use, then the change in the visits to healthcare facilities can be attributed to the of variation in the natural gas utilization. In order to see if this is the case, one should account for province and time fixed effects, time trends and region-by-year fixed effects while running the regressions and the coefficient estimates for province characteristics should not have significant impact on the natural gas use in province i at time t.

The results of the first identification exercise are presented in Table B.1. It shows the estimates of regression of dummy variable for natural gas existence on the jointly specified time-varying observable province characteristics. The equation that it's results are shown in Table B.1 is:

$$DummyNG_{it} = \alpha + \beta_1 \mathcal{X}_{it} + \theta_t + \gamma_i + \beta_2 \tau_{it} + \beta_3 \tau_{it}^2 + \omega_{ti} + \varepsilon_{it}$$

where $DummyNG_{it}$ represents the provinces with natural gas infrastructure if it equals one, $\mathcal{X}_{it}$ is the vector of independent variables, θ_t is the set of year fixed effects, γ_i is the set of province fixed effects, τ_{it} and τ_{it}^2 are linear and quadratic time trends respectively, and finally ω_{ti} stands for the set of region-by-year fixed effects. The regression is run on the years from 2012 and 2018.

In column 1, we see that the estimates of regression with year fixed effects mostly coincide with the main assumption of this section. Except for the healthcare indicators, the estimates are significantly different between the groups; meaning that the characteristics significantly associate with the presence of a natural gas infrastructure. In column 2, we see that when province fixed effects are added, the number of significantly different characteristics fall. With the addition of province-specific linear and quadratic time trends (columns 3 and 4), the number of significantly different variables decline to two, which are education characteristics of the provinces. Eventually with the addition of region-by-year fixed effects (column 5) we see that the statistical significance for the logarithm of income per capita, high school and college graduate rates and the logarithm of physicians per 1 Million remain. The significance of these coefficients reflects that the odds of having natural gas network in the province is higher for a wealthier province with more educated residents.

Table B.1 The Estimates of Dummy for Natural Gas on Jointly Specified Time-varying Observable Province Characteristics

VARIABLES	(1)	(2)	(3)	(4)	(5)
Log Income per capita	-0.09*	0.23	0.60	0.60	0.60**
	(0.05)	(0.44)	(0.37)	(0.37)	(0.29)
Log Hospitals per 1 Million	-0.02	0.01	-0.19	-0.19	-0.19
	(0.04)	(0.14)	(0.14)	(0.14)	(0.12)
Log Hospital Beds per 1 Million	0.20***	-0.11	0.11	0.11	0.11
	(0.04)	(0.15)	(0.13)	(0.13)	(0.13)
Log Physicians per 1 Million	0.09	0.50***	0.20	0.20	0.20*
	(0.10)	(0.18)	(0.14)	(0.14)	(0.11)
Log FPCs per 1 Million	0.03	-0.27	-0.01	-0.01	-0.01
	(0.10)	(0.22)	(0.05)	(0.05)	(0.12)
Log Automobiles per 1000	0.30***	-0.88***	0.05	0.05	0.05
	(0.03)	(0.33)	(0.48)	(0.48)	(0.43)
Log Student per Teacher	0.43***	0.36*	-0.03	-0.03	-0.03
	(0.13)	(0.18)	(0.17)	(0.17)	(0.13)
High School Graduate Rate	0.01**	0.03	0.07***	0.07***	0.07***
	(0.01)	(0.04)	(0.02)	(0.02)	(0.02)
College Graduate Rate	-0.02***	0.05	0.11*	0.11*	0.11***
	(0.01)	(0.07)	(0.06)	(0.06)	(0.04)
Unemployment Rate by Region	0.01***	-0.00	-0.00	-0.00	-0.00
	(0.00)	(0.01)	(0.01)	(0.01)	(0.00)
Constant	-3.71	4.08	-8.64**	-23.84	-48.52*
	(2.43)	(7.08)	(4.15)	(18.45)	(28.90)
Observations	567	567	567	567	567
R-squared	0.42	0.18	0.86	0.86	0.86
F test	2.597	2.597	2.597	2.597	2.597
Prob > F	0.00	0.00	0.00	0.00	0.00
Year fixed effects	Yes	Yes	Yes	Yes	Yes
Province fixed effects	No	Yes	Yes	Yes	Yes
Province-specific linear time trends	No	No	Yes	Yes	Yes
Province-specific quadratic time trends	No	No	No	Yes	Yes
Region-by-year fixed effects	No	No	No	No	Yes

Notes. Robust standard errors, are in parantheses.

*, ** or *** indicates significance at the 90%, 95% or 99% levels respectively.

The results of the second identification exercise are presented in Table B.2. The practice follows the same steps as in the first exercise. However, instead of regressing binary variable for natural gas existence, this time the variable for natural gas intensity is regressed on the jointly specified time-varying observable province characteristics. The regression function is:

$$Intensity_{it} = \alpha + \beta_1 \mathcal{X}_{it} + \theta_t + \gamma_i + \beta_2 \tau_{it} + \beta_3 \tau_{it}^2 + \omega_{ti} + \varepsilon_{it}$$

where $Intensity_{it}$ stands for natural gas intensity per capita, $\mathcal{X}_{it}$ is the vector of independent variables, θ_t is the set of year fixed effects, γ_i is the set of province fixed effects, τ_{it} and τ_{it}^2 are linear and quadratic time trends respectively, and finally ω_{ti} stands for the set of region-by-year fixed effects.

It is interesting to see that even in column 1 without controlling for time trends and fixed effects except year fixed effects, we only see 2 of the 10 variables for province characteristics as having statistically significant coefficients. Through columns 2 to 5, the introduction of province fixed effects, time trends and region-by-year fixed effects one after the other eliminates the significant differences between the characteristics of the two groups. Only the logarithm of automobiles per 1000 remain significantly different between the two groups, hence having a relation with natural gas intensity.

Comparing the results with the previous table, one could observe the distinct realizations of estimates for natural gas existence and natural gas intensity. Lending support to the commentaries made in Section 3.2.4, the results from Table B.1 and Table B.2 indicates a higher probability of having a natural gas pipeline in a province if that province is wealthier and more developed. However, once the pipeline is established, the consumers' choice to switch to natural gas does not depend on the time variant socio-economic province characteristics of development.

Table B.2 The Estimates of Natural Gas Intensity on Jointly Specified Time-varying Observable Province Characteristics

VARIABLES	(1)	(2)	(3)	(4)	(5)
Log Income per capita	0.09***	0.12	0.01	0.01	0.01
	(0.02)	(0.11)	(0.03)	(0.03)	(0.04)
Log Hospitals per 1 Million	0.01	-0.00	-0.02*	-0.02*	-0.02
	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)
Log Hospital Beds per 1 Million	-0.00	-0.01	0.02	0.02	0.02
	(0.01)	(0.02)	(0.01)	(0.01)	(0.02)
Log Physicians per 1 Million	-0.12***	0.00	0.00	0.00	0.00
	(0.03)	(0.02)	(0.01)	(0.01)	(0.02)
Log FPCs per 1 Million	-0.08	-0.08	0.03	0.03	0.03
	(0.05)	(0.10)	(0.06)	(0.06)	(0.05)
Log Automobiles per 1000	0.01	0.06*	0.15**	0.15**	0.15**
	(0.01)	(0.04)	(0.06)	(0.06)	(0.07)
Log Student per Teacher	0.00	0.04	-0.01	-0.01	-0.01
	(0.03)	(0.02)	(0.02)	(0.02)	(0.02)
High School Graduate Rate	0.00*	0.00	0.00	0.00	0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
College Graduate Rate	0.00	0.02	0.01**	0.01**	0.01
	(0.00)	(0.02)	(0.00)	(0.00)	(0.01)
Unemployment Rate by Region	-0.00	0.00	-0.00	-0.00	-0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Constant	1.59	0.07	-1.77	2.55	12.97**
	(1.23)	(1.40)	(1.57)	(3.20)	(5.23)
Observations	413	413	413	413	413
R-squared	0.39	0.31	0.98	0.98	0.98
F test	1.168	1.168	1.168	1.168	1.168
Prob > F	0.313	0.313	0.313	0.313	0.313
Year fixed effects	Yes	Yes	Yes	Yes	Yes
Province fixed effects	No	Yes	Yes	Yes	Yes
Province-specific linear time trends	No	No	Yes	Yes	Yes
Province-specific quadratic time trends	No	No	No	Yes	Yes
Region-by-year fixed effects	No	No	No	No	Yes

Notes. Robust standard errors, are in parantheses.

*, ** or *** indicates significance at the 90%, 95% or 99% levels respectively.