Run Away?
Air Pollution and Emigration Interests in China

Yu Qin*

Hongjia Zhu†

November 1, 2015

Abstract

This paper investigates the impact of air pollution on people's interest in emigration. Using an online search index on "emigration" which is positively correlated with its search volume, we develop a city-by-day measurement of people's emigration sentiment. We find that searches on "emigration" will grow by approximately 2.3-4.7% the next day if today's air quality index (AQI) is increased by 100 points. In addition, such an effect is more pronounced when the AQI level is above 200, a sign of "heavily polluted" and "severely polluted" days. We also find that such effect differs by destination countries and by metropolitan areas.

Keywords: Emigration; Air Pollution; China

JEL classification: Q53; Q56; R23

*Corresponding author. E-mail: rstqyu@nus.edu.sg, Department of Real Estate, National University of Singapore.

†E-mail: irszh@nus.edu.sg, Institute of Real Estate Studies, National University of Singapore.

1 Introduction

International migration plays an important role in the globalization process, imposing substantial influences on the labor markets, productivity and economic growth. A growing body of literature has documented the profound consequences of emigration on the origination country, such as the "brain drain effect".[1] Therefore, in addition to measuring the socioeconomic consequences of international migration, it is no less important to understand the driving forces of emigration from various perspectives. Many previous works have documented important factors motivating emigration, such as a lower level of development/income/wages in the source country (Docquier and Rapoport, 2012; Hatton and Williamson, 2002; Hunt, 2006), higher income/wages in the destination countries (Karemera et al. 2000), increasing income per capita gap between the origin and destination countries (Ortega and Peri, 2009) and more relaxed immigration law in the destination countries (Ortega and Peri, 2009; Mayda, 2010).

Despite the high economic growth rate in China in the past few decades, the total number of emigrants in China continues to rise (Figure 1). As one of the largest suppliers of international migration, the total number of Chinese emigrants (above age 15) was 3.86 million in OECD countries in 2010-2011; thus, China ranked as the second largest supplier of emigrants after Mexico.[2] It is reported that approximately 1 million Chinese have obtained permanent-resident status in Canada or America in the past decade, placing Chinese migrants first in Canada and second in America behind Mexicans.[3] A large share of Chinese emigrants are high-skilled or high-income. Of these, 1.66 million of the 3.86 million Chinese emigrants in OECD countries are highly educated.[4] In addition, a survey by Barclays Bank in 2014 found that 47% of rich Chinese plan to emigrate in the next five years, compared with 23% of Singaporeans and 16% of Hong Kongers. These facts suggest that China is encountering increasing challenges in human capital and wealth flight due to international emigration.[5]

---

[1]Regarding whether the net "brain drain effect" is detrimental or beneficial to the source country remains controversial. For example, Vidal (1998) suggests that emigration to a higher return to skills country may encourage people in the source country to invest in human capital; Beine et al. (2008) find that most countries combining low levels of human capital and low migration rates of skilled workers tend to be positively affected by the brain drain, whereas the brain drain appears to have negative growth effects in countries where the migration rate of the highly educated is above 20% and/or where the proportion of people with higher education is above 5%. Agrawal et al. (2011) indicate that the emigration of highly skilled individuals weakens local knowledge networks (brain drain) but may help remaining innovators access valuable knowledge accumulated abroad (brain bank).

[2]See http://www.oecd.org/els/mig/World-Migration-in-Figures.pdf.

[3]See http://www.economist.com/news/china/21601305-more-middle-classes-are-leaving-search-cleaner-slower-life-yearning-breathe.

[4]See http://www.oecd.org/els/mig/World-Migration-in-Figures.pdf.

[5]Another example is that Chinese buyers spent more than 221 billion U.S. Dollars on property in

Emigration is a long-term decision and is attributed to many factors, such as income, politics, education and quality of life. As an important component of the quality of life, the environmental quality may affect the emigration decision. Although China has been experiencing high economic growth, it also has suffered from severe environmental degradation in recent years. For example, air pollution is now recognized as an increasing concern that affects China's public health, industrial development and economic growth (Brandt and Rawski, 2008). However, within the recent literature that studies international emigration, there is minimal research that has linked emigration incentives with the environmental degradation in the origin country. The major objective of this paper is to investigate the contemporary association between air pollution and the interest in international emigration in China.

In this paper, we test the hypothesis that a very short-term (daily) shock on air pollution levels may drive up the interest in emigration in a city. No papers have ever established a causal link between pollution and emigration possibly due to the scarcity of data on international migration. In addition, because emigration takes time to process, it is difficult to trace it to the time when the emigration decision was made. However, this is an important question to ask because pollution is likely to motivate emigration due to its negative impacts on health and subjective well-being.[6] In fact, anecdotal evidence in Hong Kong suggests that 26% of surveyed Hong Kong adults have considered emigration due to poor air quality. More importantly, those people who considered leaving were among the most competitive individuals in Hong Kong, including those with undergraduate and post-graduate degrees and high-income earners.[7]

To address the data scarcity and inaccessibility of international migration, we develop a proxy for people's intent to migrate by collecting a city-by-day search index on emigration ("*yi min*" in Chinese) via Baidu, the largest Chinese search engine. The value of the Baidu Index is positively correlated with people's search volume on the key word "emigration" in a city in a day; thus, it captures the contemporaneous aggregate interests on emigration in a city. We merge the search index with the air quality index (AQI) of 153 major Chinese cities in 2014. Our regression results show that the one-day lagged air pollution level signifi-

the U.S. alone, between April 2013 and March 2014. See http://www.rfa.org/english/news/china/flood-02122015104709.html

[6]A growing body of literature studies the health impacts of air pollution. For example, Chay and Greenstone (2003) and Currie and Walker (2011) estimate the significant effects of air pollution on the infant mortality rate, premature births, and low birth weight using the U.S. data. Schlenker and Walker (2011) focus on a shorter time span of the impacts and show the contemporaneous health impacts of air pollution for various population cohorts. Using China's data, Chen et al. (2013) find that the higher concentration levels of total suspended particulate (TSP) due to the winter heating policy in north China is responsible for approximately 5.5 years of lower life expectancy. Zhang et al. (2015) show that air pollution significantly reduces short-term hedonic happiness.

[7]See http://www.civic-exchange.org/en/publications/164987357.

cantly increases people's searches on emigration. On average, a 100 point increase in the air pollution level one day before significantly leads to an approximately 2.5% increase in the current search index. In addition, the estimated impacts illustrate nonlinear patterns. The impacts become more pronounced when the air pollution level achieves "heavily polluted" (AQI above 200 and below 300) and "severely polluted" (AQI above 300). We find that the magnitude of the air pollution's impact on searches for "emigration" is approximately one third of its impact on searches for "masks" and one tenth of its impact on searches for "smog" and "PM 2.5", because the top income earners alone could afford the option to emigrate.

In addition, we find heterogeneous impacts of air pollution on searches for emigration with different destination countries. Among the top four destinations of Chinese emigrates, i.e., Australia, Canada, New Zealand and the United States, the impact is most pronounced for emigration to the United States, which is the most popular destination of Chinese emigrants. A 100 points increase in air pollution significantly increases people's searches on "emigration to the U.S." by 2.6%. Moreover, we find heterogeneous impacts across the four largest metropolitan areas in China, including Beijing, Shanghai, Shenzhen and Guangzhou. In particular, the impact is the largest and most significant in Beijing, which is one of the most polluted large cities in China. Last, to address the possible concern regarding the potential manipulation problem of the official data of China, we verify the causal link between air pollution and search behavior using the hourly PM 2.5 data released by the US Embassy and Consulates in five cities. Similar results are found by using data from a different source.

This paper contributes to the literature in two ways. First, to the best of our knowledge, this is the first paper that investigates the causal link between pollution and emigration, which adds to the literature studying the determinants of emigration (Docquier and Rapoport, 2012; Hatton and Williamson, 2002; Karemera et al., 2000; Hunt, 2006; Ortega and Peri, 2009; Mayda, 2010). Second, this paper also adds to a growing body of literature that investigates the economic impacts of air pollution from a new perspective of lost human capital. In addition to the substantial literature that studies the long-term and short-term health impacts of air pollution, certain recent literature finds that air quality has significant impacts on workers' productivity (Graff Zivin and Neidell, 2012; Li et al., 2015) and academic outcomes (Currie et al., 2009; Stafford, 2015); this suggests that the impacts of air pollution may be limited to its direct impacts not only on the health, but also on a various socioeconomic outcomes. This paper attempts to associate air pollution to a new perspective, the interest in emigration. Although emigration is a long-term decision, our findings indicate that severe air pollution in the short run may increase emigration interest in the population intending to emigrate, particularly for the population who has not yet decided to migrate.[8]

2 Data Description

2.1 Search Index

A novelty of this study is that we collect the information of the daily variation of the sentiment to emigrate in all of the prefecture cities in China via Baidu. The Baidu search index, similar to Google Trends (GT), provides a measurement of the search volume of a key word in a given time period. There is a growing body of economic literature that uses GT to measure the searching interest of Internet users. For example, Tefft (2011) uses GT on the terms "depression" and "anxiety" to proxy the intent to seek treatment for psychological distress. Goel et al. (2010) and Choi and Varian (2012) find that GT has considerable prediction power on economic outcomes such as macroeconomics indicators, product sales, and consumer behavior.

China has experienced very rapid growth in its internet penetration rate. The number of internet users increased from 0.33 billion in 2009 to 0.63 billion in 2014, which is approximately twice the U.S. population.[9] However, because Google services are blocked in China by the Great Fire Wall, the conventional Google Trends data used in the previous literature cannot reflect the fluctuation of the real search volume of Chinese Internet users. Instead, Baidu, which was founded in 2000, is the largest search engine in China and has more than 50% of the market share.[10]. The Baidu Index is a data product analogous to GT, which measures the search frequencies of the selected terms. According to the official explanation, the Baidu Index provides a weighted sum of the search volume on a key word in a given period. We have no information regarding the specific formula it adopts to transform search volume into an index because it is a trade secret. However, the search index is very likely to be linearly correlated with the search volume of a key word based on the small sample of search volume and search index data that we obtained online (refer to Appendix Figure A1).

Because international emigration is typically a long-term decision process, searching the related information regarding it on the internet can be interpreted as a revelation of emigration interest. Therefore, we collect daily Baidu Index data on the Chinese keyword "*yi min*" (emigration) in the 153 matched cities with the air pollution data in 2014. Fortunately, there is no ambiguity of the term "*yi min*" in Chinese. Moreover, this term primarily

[8]We assume that emigration searches by people who have planned to move are less likely to be affected by temporary shocks, such as air pollution.

[9]The numbers refer to https://www.cnnic.net.cn/hlwfzyj/jcsj/index.htm

[10]See https://en.wikipedia.org/wiki/Baidu

indicates the meaning of international emigration in a single-word context. Therefore, we believe that the Baidu Index for this specific term provides a credible measurement of the overall interest of a city's international emigration. Figure 2 presents the spatial disparity of the average Baidu Index of "emigration" in our selected cities in 2014. Compared with their counterparts, Beijing, Guangzhou and Shanghai have the highest values, which is partially due to their large population size and higher development levels.

To investigate the potential heterogeneity of emigration interests by destination countries, we collect key word searches for the largest four destination countries of Chinese emigrants, including Australia, Canada, New Zealand and the United States. Specifically, we collect Baidu Index data for the Chinese keywords search on "emigration to Australia", "emigration to Canada", "emigration to New Zealand" and "emigration to the United States". In addition, we also test the potential impacts of air pollution on certain other searching interests directly associated with air pollution. To do this, we collect Baidu Index data on "PM 2.5", "smog" (*wu mai*) and "mask" (*kou zhao*) covering the same sample period as the outcome variables in the estimations. A few recent literature studies such as Mu and Zhang (2014) and Zheng et al. (2015) use the online purchase data and find that there is significant impact of air pollution on the purchasing of masks. Because these online searches on pollutants, pollution and preventative measures are likely to be triggered by the pollution episodes, our purpose is to compare the potential impacts, if any, of air pollution on emigration and on these other related searches.

Table 1 provides the summary statistics for Baidu Index on all of the above keywords. The average index on "emigration" is approximately 72, with the maximum value as 1006, and the minimum as zero. There is no direct means to interpret the economic meaning of the numbers because we cannot precisely match them to the corresponding search volume. However, it is possible to compare the relative magnitude across different indices. For example, the country-specific emigration indices have a lower mean compared with the overall emigration index. In addition, the search index for emigration to New Zealand has the highest mean value among all of these four countries, whereas the index is relatively lower for keyword searches on emigration to the United States. It is noteworthy that the differences on search frequency may be due to many reasons, such as the general interest in a specific country, the respective immigration policies and other socioeconomic factors. In addition, it is found that on average, there are more searches on "PM 2.5" and "smog" in Chinese cities compared with searches on emigration. However, the search on "mask" has a lower index compared with emigration.

2.2 Air Pollution Data

We collect daily air pollution data from the Ministry of Environmental Protection (MEP) for the empirical analysis of this paper. Our data source is the same as certain previous studies related to air pollution in China, such as Viard and Fu (2015) and Mu and Zhang (2014). Since 2001, the MEP has published daily air pollution data on its website. Before 2012, the air pollution index (API), rather than air quality index (AQI), was reported on the website. A major difference between API and AQI is that the later considers the concentration level of PM2.5, which is one of the major pollutants in many cities that have caused the public's increasing concern in recent years.

MEP began to report the new measurement of AQI on its website in 2014. However, there have been 153 cities publicizing their daily AQI since January of 2014, whereas the remainder of the cities was gradually included in latter months. To construct a balanced panel for our analysis, we collect the daily AQI information solely in these 153 cities and then matched the data with the emigration indices. The core information of our dataset is the daily AQI measurement of each city in 2014, which is the average of the daily AQI readings from different monitor stations in each city. Additionally, the levels of air pollution are also provided according to the reported AQI. In China, air quality levels are classified into six categories: excellent, good, lightly polluted, moderately polluted, heavily polluted, and severely polluted, with corresponding AQI cutoff points at 50, 100, 150, 200, 300, and 500.

In our data, the daily AQI readings range from 12 to 500, with significant spatial heterogeneity across different regions. Figure 3 presents the average AQI of the 153 cities in 2014. Air pollution is more severe in northern China. Beijing, for example, had only 28 days of good air quality ($AQI < 50$) in 2014, but had 14 heavily polluted days and 28 severely polluted days. In contrast, Shenzhen, a southern high-technological city and the home of over 18 million people, had 155 good air quality days and there was no day with AQI value higher than 150 in 2014. If we divide the cities into 20 sub-groups based on their average AQI levels in 2014 and then plot against the average Baidu Index on emigration searches of each sub-group, we can observe a positive correlation between these two variables as suggested in Figure 4.

A potential concern of the official AQI data is that it may be manipulated by the government. As an effort for pollution abatement, the air quality of a city is now associated with the promotion of the local government officers. Therefore, this promotion opportunity may provide incentives for government officers to manipulate the reported air pollution data. Indeed, using the daily data of API, Chen et al. (2012) finds that there is evidence of downward manipulation at the cutoff point at 100, which was the threshold of defining a "blue-sky day". Although there is no recent evidence on whether the newly adopted AQI also suffers from data manipulation, we address this concern by providing evidence using another data source from the U.S. Embassy and Consulates in five cities in China, which are not manipulated by the local government.

The five U.S. Embassy and Consulates in China measure and publicize the hourly reading of PM 2.5 in Beijing, Shanghai, Guangzhou, Chengdu and Shenyang since 2008. However, it should be noted that, due to different measurements, the readings from MEP and the U.S. Embassy and Consulates are not directly comparable. In addition to the credibility of the data quality, another advantage of using the U.S. data is that it allows us to study the response of the search index with respect to the air pollution in different time periods. A simple statistics of the hourly air quality shows that the air is typically more polluted during the night time than the day time. Because the night-time air pollution is less observable than the day-time pollution, it may potentially lead to heterogeneous effects on searching behavior. Therefore, using the hourly data, we are able to test whether the daily change of the search index is more sensitive to the pollution during working hours.

2.3 Weather Data

Weather may also affect peoples intention to migrate. For example, people are more likely to be depressed on rainy days than sunny days; thus, they are more likely to be dissatisfied with their current living conditions. The short-term psychological effects of weather conditions have been examined in the finance literature using security market data (Saunders, 1993; Hirshleifer et al., 2003). It is also worth noting that weather conditions have considerable effects on the local air quality, as discussed in Viard and Fu (2015) and Mu and Zhang (2014). If weather conditions are omitted from our regressions, the correlations between weather and local air quality as well as the interests of emigration will lead to biased estimations.

We address this problem by including a series of daily weather measurements in our regressions. Specifically, we control for daily temperature, dew point, wind speed, and precipitation during the entire year of 2014. The weather data are from the National Climatic Data Center under the US National Oceanic and Atmospheric Administration (NOAA), which provides rich daily weather information at the monitor station level. To match the daily weather data with air pollution and Baidu Index data, we search for the nearest weather station for each geographic city center.

3 Empirical Strategies

The major objective of this paper is to examine the potential short-term effects of air pollution on the interest in emigration in China. To begin, we specify the following estimation function:

$$Baidu_{i,t} = Constant + FE + \sum_{n=0}^{6} \beta_n AQI_{i,t-n} + \lambda Weather_{i,t} + \epsilon_{i,t} \quad (1)$$

As described in the data section, we use the search frequency of the keyword "emigration" in Chinese as a proxy for people's emigration sentiments. The daily AQI is the core independent variable to measure the air quality. However, because we are not sure whether air pollution has both contemporaneous and lagged effects on searching behavior, we include both the current day as well as six lagged days of air quality in Equation 1. Therefore, the vector of parameters β_n represents the potential effects of air pollution on searching behavior in different periods. In addition, to capture the potential effects of weather conditions on both the emigration interests and local air quality levels, we control a list of weather variables including wind speed, rainfall, maximum temperature, minimum temperature, mean temperature, and dew points. We also include different sets of fixed effects in different specifications. In particular, in the most conservative regression form, we control for city-week fixed effects, public holiday fixed effects, and day-of-week fixed effects.

It is noteworthy that the values of the Baidu Index on "emigration" are intrinsically non-negative count numbers. Figure 5 shows a typically skewed distribution of the Baidu Index, which has 32% of the values at 0. Therefore, the normal distribution assumption of error terms tends not to hold in the ordinary least square (OLS) regression if there is no transformation on the dependent variable, which may subsequently result in invalid statistical inferences. Instead, the Poisson regression model is an appropriate alternative estimation method in this context, which is also applied in Mu and Zhang (2014) where the online sales index of mask is used as the outcome variable.[11] The interpretation of β_n in the Poisson model is analogous to the log-linear model, which implies that every 1 unit increase of $AQI_{i,t-n}$ will lead to β_n% change of the search frequencies in Equation 1.

The estimation of Equation 1 assumes that the there is a linear effect of air pollution on people's searching behavior. That is, every 100 increase in the AQI values at different levels has the same impact on search frequency. However, recent literature finds that the effects of air pollution display significant nonlinear features. In particular, the average impacts of air pollution at higher levels are usually disproportionately larger than the air pollution at lower levels. To test whether there is a nonlinear pattern of the impacts of air pollution in this study, we collapse the continuous measurement of $AQI_{i,t-n}$ into six dummy indicators, AQI^m, to denote the pollution levels as defined by the government.[12] Therefore, the estimated

[11] In our analysis, we find that the results from OLS regressions and Poisson regressions all point to similar conclusions.

[12] To avoid the perfect multicollinearity problem in the regression, we omit the first category AQI^1 in Equation 2.

coefficient of β_m captures the partial effect of a day in the mth pollution level on the local search frequency of keyword "emigration". If there is a nonlinear pattern of these effects, it is expected that the estimates of these coefficients show particular trends either in statistical significance or in magnitudes.

$$Baidu_{i,t} = Constant + FE + \sum_{m=2}^{6} \beta_m AQI_{i,t}^{m} + \lambda Weather_{i,t} + \epsilon_{i,t} \tag{2}$$

4 Main Findings

4.1 Dynamic Effects of Air Pollution on Baidu Index

Because we do not have a priori evidence regarding the dynamic effects of air pollution on the change of search interest on "emigration", finding the effective timing of air pollution on search behavior is fundamental to our analysis. To answer this key question, we use the emigration index as the dependent variable and include the contemporary AQI, and we lagged one day to six days of AQI values on the right side.

Table 2 presents the dynamic regression results. As discussed in the previous section, we adopt the Poisson regression method and control for a rich set of fixed effects in the estimations. To facilitate the interpretation of the coefficient, we rescale the value of AQI by dividing the original AQI values by 100. In addition to controlling for day-of-week fixed effects, public holidays and weather conditions, we use different alternative levels of city and time fixed effects across column 1 to 3 to test the sensitivity of regression results. Specifically, in the first column, we control for month and prefecture city fixed effects to capture the unobservable factors in both the temporal and spatial dimensions. In the second column, we control for city-by-month fixed effect, which eliminates biases due to potentially heterogeneous time patterns in different cities. Last, the third column provides evidence using the most conservative regression form, i.e., a city-by-week fixed effect, which further allows different growth patterns in different weeks within the cities.

The results in Table 2 suggest a positive and significant impact of the lagged-one value of the AQI on online searches for "emigration". The positive coefficients range from 0.015 to 0.025, which indicates that a 100 point increase in the AQI leads to an approximately 1.5 to 2.5% growth in the Baidu Index one day after. Although the AQI has a significant and positive contemporary effect on the Baidu Index in the estimation of column 1, the effect becomes statistically insignificant in our preferred estimation in column 3, which controls for the variations within a city-week unit. In summary, according to the findings above, we conclude that the variation of the Baidu Index are primarily affected by the lagged one day

AQI values; this nominates the lagged one day AQI as the key independent variable in our subsequent analysis.

Table 3 presents the results of regressing the Baidu Index on the lagged one day value of AQI. Again, different combinations of fixed effects are controlled for across three columns. The results suggest a positive impact of lagged air pollution on online searches for "emigration". Specifically, a 100 point increase in the AQI level leads to an approximate 4.7% growth in terms of the Baidu Index on "emigration" the next day, as indicated in column 1. The magnitude of effect decreases to approximately 2.5% after imposing stronger fixed effects in column 2 and 3, but it remains statistically significant.

Overall, our findings above indicate a short-term effect of air pollution on the international emigration interest in China. These findings are in accordance with recent evidence from Zhang et al. (2015), who find that the air pollution has immediate negative effects on shorter-term hedonic happiness. However, because we continue to lack evidence regarding the association between online search interest and the long-term decision process, the evidence we have provided can be interpreted as a pulse response of emigration interest because of air pollution episodes. However, it is difficult to translate these findings into effects on the ultimate emigration behavior at this stage.

4.2 Nonlinear Effects

Certain previous studies find that the impacts of air pollution tend to be nonlinear (Li et al., 2015; Currie et al., 2009; Graff Zivin and Neidell, 2012; Schlenker and Walker, 2011). In this section, we will test whether there is any nonlinear pattern of the estimated effects of air pollution on the Baidu Index. To do so, we categorize the AQI index into six groups based on the guidance by the Ministry of Environmental Protection of China: "excellent" (AQI $\subset$[0,50]); "good" (AQI $\subset$[51,100]); "lightly polluted" (AQI $\subset$[101,150]); "moderately polluted" (AQI $\subset$[151,200]); "heavily polluted" (AQI $\subset$[201,300]); and "severely polluted"(AQI $\subset$[301,500]). We transfer the daily AQI value into the corresponding dummy variable. To avoid the perfect multi-collinearity problem in the regression, the dummy for "good" is absorbed as the baseline group.

Table 4 reports the regression results for the nonlinear effect of air pollution on the Baidu Index on "emigration". Interestingly, the estimated coefficient increases in magnitude as AQI achieves higher levels. The impact of air pollution on emigration searches is most significant after AQI exceeds 200 (corresponding to "heavily polluted" and "severely polluted"). A "heavily polluted" day with AQI between 201 to 300 leads to approximately 5.5 to 10.8% growth in the emigration index on the subsequent day. Moreover, a "severely polluted" with an AQI exceeding 300 leads to approximately 7.9 to 12.7% growth in the Baidu search index on emigration, which is approximately two times as high as the effects of "heavily polluted" days and six times as high as the effects of "moderately polluted" days.

5 Discussions

5.1 Using Air Pollution Data from the U.S. Embassy

There is concern raised in the previous literature regarding the accuracy of official AQI data in China. For example, Chen et al. (2012) use data from the air pollution index (API) from 2000 to 2009 and find that, as a result of local government officers' promotion incentives, the air pollution data may be manipulated to achieve the number of "blue sky days" in a year. Such manipulation typically affects the API values that are slightly above 100, which is the threshold for defining a "blue sky day". However, because our results suggest that the AQI level does not lead to increased emigration searches until it hits 200, the manipulation for "blue sky" is not likely to affect our main findings.

To completely eliminate the possibility of biased results due to data manipulation, we also collect the PM2.5 data reported by the U.S. Embassy in Beijing and its Consulates in four cities, including Shanghai, Guangzhou, Shenyang and Chengdu. A salient feature of the U.S. data is that it provides the hourly readings of the PM 2.5 concentration level, thus allowing us to investigate the effects of pollution within different time periods. Therefore, we create four outcome variables using the hourly data: the mean and maximum values of the PM 2.5 level during working hours (9 am to 6 pm) and the whole day, respectively.

Similar to the previous analysis, we conduct regressions by using the continuous measurement of air pollution as well as exploring the nonlinear patterns of it. However, it is noteworthy that although the PM 2.5 is the major pollutant in certain cities, the PM 2.5 data cannot be directly compared with the AQI data in levels because AQI is a composite index composed of six different pollutants. Therefore, it is not sensible to create dummy variables for PM2.5 levels based on the classification method of AQI levels. Instead, we create dummy variables for each 50 points of PM 2.5 increments capped at 300, which results in 7 dummy variables that represent the pollution levels.

Table 5 reports the findings using the continuous PM 2.5 data in four various measurements. To save space, we only report the regression results using the finest week-by-city fixed effect for each outcome variable. The coefficients on lagged PM 2.5 are positive but not significant, which could possibly be explained by the reduced power due to a smaller sample size. However, the nonlinear effects found in the AQI data also appear when the PM 2.5 data are used. As shown in Table 6, the impact of lagged PM 2.5 on the emigration index is most pronounced when PM 2.5 achieves 300. The lagged one day PM 2.5 value higher than 300, on average, leads to approximately 5.7 to 11% growth in the Baidu Index of "emigration". Moreover, we also show the heterogeneous effects of different pollution level definitions. Overall, the dummy variables of the level-7 pollution category are all statistically significant across column 1 to 4. Particularly, the effects of the mean and maximum pollution levels during working hours are generally larger than the effects of the mean and maximum pollution levels of the entire day, suggesting that people may be more affected by the pollution during the time that they will interact in a day.

5.2 Heterogeneous Effects by Country of Destination

Our main results suggest that a higher level of air pollution leads to more searches on emigration. In this section, we take a further step to investigate the potential heterogeneous effects of air pollution by different destination countries of emigration. Specifically, we use the Chinese keyword searches on "emigration to Australia", "emigration to Canada", "emigration to New Zealand" and "emigration to the United States" as outcome variables. The reason for selecting the Baidu Index for this is that these four countries are the top choices for Chinese emigrants.[13]

Table 7 reports the results for the four popular destination countries. We solely report the results of regression specifications with week-by-city fixed effects. Interestingly, the impact of air pollution on searches for emigration to the United States is the most significant among all of these four countries. A 100 point increase in the lagged PM 2.5 leads to a 2.6% growth in the Baidu Index on "emigration to the U.S.". Table 8 presents the nonlinear effect of air pollution on emigration searches by destination countries. Again, the effect is more pronounced for searches on "emigration to the U.S." when PM 2.5 hits 200 and 300. One explanation for these results is that the United States is the most popular destination country for Chinese emigrants potentially due to higher income, cleaner air, lower housing prices and better education. As suggested by the statistics, in 2012, 81,784 Chinese emigrants obtained permanent residence in the United States, compared with 33,018 Chinese emigrants in Canada, 25,509 in Australia, and 7,223 in New Zealand.[14]

5.3 Effect in Four Largest Metropolitan Areas

In this section, we test the effect of air pollution on emigration searches in the four largest metropolitan areas in China, namely Beijing, Shanghai, Guangzhou and Shenzhen. These cities host large populations of highly educated, high income and wealthy individuals. For each city, we include a full set of control variables (including weather, day-of-week fixed effect and holiday dummy) and the weekly fixed effect.

[13] http://www.rfa.org/english/news/china/flood-02122015104709.html
[14] http://en.ccg.org.cn/Research/view.aspx?Id=512

Table 9 presents the regression results on these four respective cities. Interestingly, the impact is most significant in Beijing but primarily muted in other three cities. A 100 point increase in the lagged AQI level leads to an approximately 2.9% increase in emigration searches in Beijing. Such an effect is particularly large when the AQI level achieves "heavily polluted". As shown in Table 10, the search index on "emigration" in Beijing grows by approximately 12.7% when AQI achieves 300. A possible explanation of such heterogeneity is that Beijing is on average more polluted than the other three cities. The average AQI in Beijing is 127 in 2014, whereas the number is approximately 80 in Shanghai and Guangzhou in the same period and 55 in Shenzhen.

5.4 Magnitude

To gain a better understanding of the magnitude of air pollution's impact on emigration sentiment, we compare the impact of air pollution on emigration with its impact on other pollution-related keyword searches. Specifically, we select the Baidu Index of keyword searches on "PM 2.5", "smog" and "mask" in Chinese as the outcome variables. We apply the same regression methods and presents the findings in Table 11. In summary, a higher pollution level leads to significantly more searches for all three keywords the next day. A 100 point increase in the AQI level leads to 19.5-27.1% more searches on "smog", 23.3-32.1 % more searches on "PM 2.5" and 10.3-12.7% more searches on "mask".

In comparison, the impact on searches for "emigration" is approximately 2.3-4.7%, which is approximately one third of the impact of searching on "mask" and approximately one tenth of the impact of searching on "smog" and "PM 2.5". It appears that the air pollution effect is much smaller on emigration than keyword searches on PM 2.5 and smog. However, this is an economically reasonable magnitude because a very small proportion of the overall population may consider emigration.[15] In contrast, the majority of the population tends to obtain additional information on pollution as well as related self-protective measures as a rational response to air pollution shocks. Our finding is related to the study by Mu and Zhang (2014), who find that air pollution increases people's purchases of masks as self-protection. However, the relatively smaller coefficient of air pollution on the Baidu Index on "emigration" than other terms does not imply that the real effect is negligible. In contrast, it is worth noting that people who consider emigration as an option are likely to be much wealthier and more educated compared with the general population. Thus, the economic consequences of emigration could be important and need further study in future research.

[15]The accumulated number of emigrants is approximately 9.34 million by the end of 2013, of a total population of approximately 1.4 billion.

5.5 Controlling for Anti-Corruption Campaign

Another concern on our estimation regards about the existence of other confounders such as the anti-corruption campaigns. China conducted waves of anti-corruption campaigns in 2014, targeting corrupt leaders from high ranks to low ranks. Therefore, such campaigns may motivate the corrupted leaders and their relatives to emigrate as the last option. To tease out the effect of an anti-corruption campaign on emigration searches, we collect the Baidu search Index search data on "anti-corruption" (*fan fu* in Chinese) and use it as a control variable in the main regressions. Table 12 reports the impact of the lagged AQI on the emigration index using both the continuous measure and the dummies for AQI classifications. The results remain very similar to Table 3 and 4, indicating that our results are robust to the exclusion of this significant movement.

6 Conclusion

In this paper, we study the impact of air pollution on people's interest in international emigration in China. Using the Baidu search index on emigration related Chinese keywords as the measurement of overall expressed interest, we establish a causal relationship between air pollution and the interest in emigration at the city-by-day level. Specifically, we find that the searches on "emigration" will grow by approximately 2.3-4.7% the next day if the AQI today is raised by 100 points, which suggests that air pollution increases people's interest in emigration. In addition, this effect is more pronounced when the AQI level is above 200, which indicates "heavily polluted" and "severely polluted" days. We also find that air pollution's impact on the emigration interest differs by destination countries. Among all of the top four destination countries, the impact is most significant for the United States, which is the number one destination country of Chinese emigrants. Moreover, the effect is large and significant in Beijing but not in the other three largest metropolitan areas, i.e., Shanghai, Guangzhou and Shenzhen, which are possibly driven by the low pollution levels in these cities.

Emigration is a long-term decision. However, our findings indicate that severe air pollution in the short run may switch on people's interest in emigration, particularly for the marginal population who are not yet determined to emigrate. Because there is no literature documenting the relationship between the interest in emigration and the real behavior of emigration, we cannot estimate the impact of pollution on the number of new emigrants, which is a major caveat of this paper. We leave it for future research and await new data collection on emigration behaviors. However, given that a positive correlation between online searching behavior and a final decision usually exists, the policy makers may consider the potential human capital and wealth flight due to emigration in the cost benefit analysis of pollution abatement.

Acknowledgement

We are grateful to Sumit Agarwal, Yongheng Deng, Shihe Fu, Xiaobo Zhang, and seminar participants at NUS for their valuable comments and suggestions. Yu Qin acknowledges financial support from Academic Research Fund - Tier 1 (WBS: R-297-000-129-133). All remaining errors are ours.

References

Agrawal, A., D. Kapur, J. McHale, and A. Oettl (2011). Brain drain or brain bank? the impact of skilled emigration on poor-country innovation. *Journal of Urban Economics 69*(1), 43–55.

Beine, M., F. Docquier, and H. Rapoport (2008). Brain drain and human capital formation in developing countries: Winners and losers*. *The Economic Journal 118*(528), 631–652.

Brandt, L. and T. G. Rawski (2008). *China's great economic transformation*. Cambridge University Press.

Chay, K. Y. and M. 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.

Chen, Y., A. Ebenstein, M. Greenstone, and H. Li (2013). Evidence on the impact of sustained exposure to air pollution on life expectancy from chinas huai river policy. *Proceedings of the National Academy of Sciences 110*(32), 12936–12941.

Chen, Y., G. Z. Jin, N. Kumar, and G. Shi (2012). Gaming in air pollution data? lessons from china. *The BE Journal of Economic Analysis & Policy 12*(3).

Choi, H. and H. Varian (2012). Predicting the present with google trends. *Economic Record 88*(s1), 2–9.

Currie, J., E. A. Hanushek, E. M. Kahn, M. Neidell, and S. G. Rivkin (2009). Does pollution increase school absences? *The Review of Economics and Statistics 91*(4), 682–694.

Currie, J. and R. Walker (2011). Traffic congestion and infant health: Evidence from e-zpass. *American Economic Journal: Applied Economics 3*(1), 65–90.

Docquier, F. and H. Rapoport (2012). Globalization, brain drain, and development. *Journal of Economic Literature*, 681–730.

Goel, S., J. M. Hofman, S. Lahaie, D. M. Pennock, and D. J. Watts (2010). Predicting consumer behavior with web search. *Proceedings of the National academy of sciences 107*(41), 17486–17490.

Graff Zivin, J. and M. Neidell (2012). The impact of pollution on worker productivity. *American Economic Review 102*(7), 3652–73.

Hatton, T. J. and J. G. Williamson (2002). What fundamentals drive world migration? Technical report, National Bureau of Economic Research.

Hirshleifer, D., T. Shumway, et al. (2003). Good day sunshine: Stock returns and the weather. *Journal of finance 58*(3).

Hunt, J. (2006). Staunching emigration from east germany: Age and the determinants of migration. *Journal of the European Economic Association 4*(5), 1014–1037.

Karemera, D., V. I. Oguledo, and B. Davis (2000). A gravity model analysis of international migration to north america. *Applied Economics 32*(13), 1745–1755.

Li, T., H. Liu, and A. Salvo (2015, March). Severe Air Pollution and Labor Productivity. IZA Discussion Papers 8916, Institute for the Study of Labor (IZA).

Mayda, A. M. (2010). International migration: A panel data analysis of the determinants of bilateral flows. *Journal of Population Economics 23*(4), 1249–1274.

Mu, Q. and J. Zhang (2014). Air pollution and defensive expenditures: Evidence from particulate-filtering facemasks. *Available at SSRN 2518032*.

Ortega, F. and G. Peri (2009). The causes and effects of international migrations: Evidence from oecd countries 1980-2005. Technical report, National Bureau of Economic Research.

Saunders, E. M. (1993). Stock prices and wall street weather. *The American Economic Review*, 1337–1345.

Schlenker, W. and W. R. Walker (2011). Airports, air pollution, and contemporaneous health. Technical report, National Bureau of Economic Research.

Stafford, T. M. (2015). Indoor air quality and academic performance. *Journal of Environmental Economics and Management 70*, 34–50.

Tefft, N. (2011). Insights on unemployment, unemployment insurance, and mental health. *Journal of Health Economics 30*(2), 258–264.

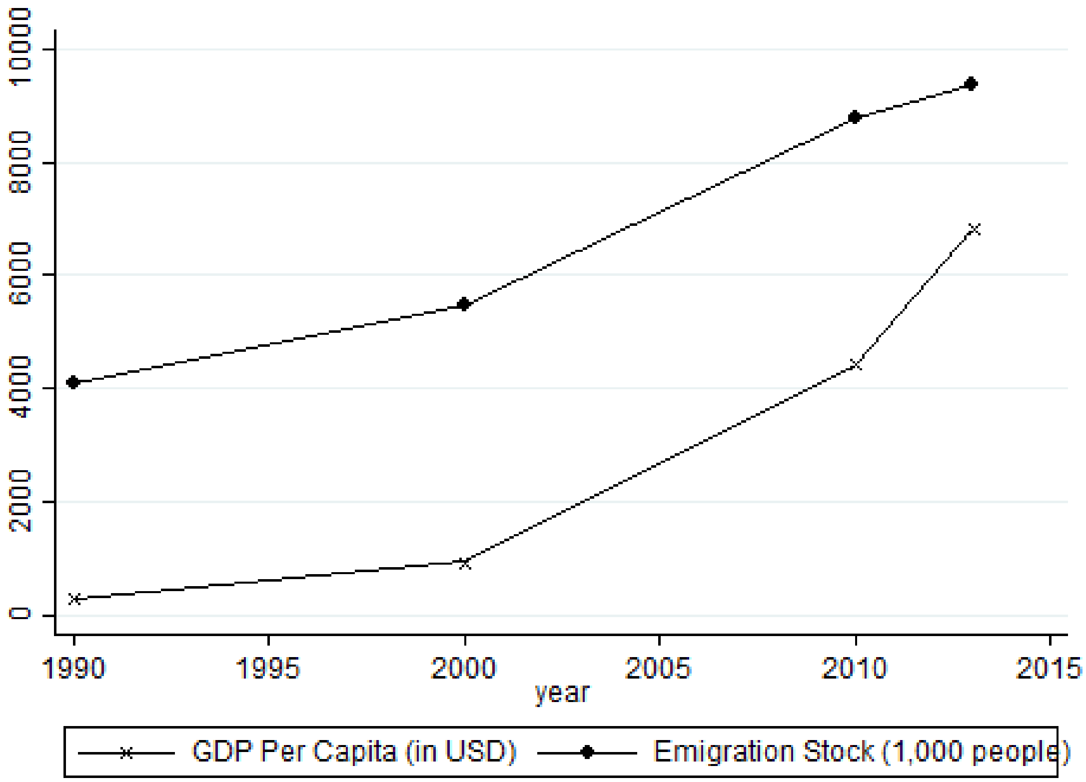
Viard, V. B. and S. Fu (2015). The effect of beijing's driving restrictions on pollution and economic activity. *Journal of Public Economics 125*, 98–115.

Vidal, J.-P. (1998). The effect of emigration on human capital formation. *Journal of Population Economics 11*(4), 589–600.

Zhang, X., X. Zhang, X. Chen, et al. (2015). Happiness in the air: How does a dirty sky affect subjective well-being? Technical report, Institute for the Study of Labor (IZA).

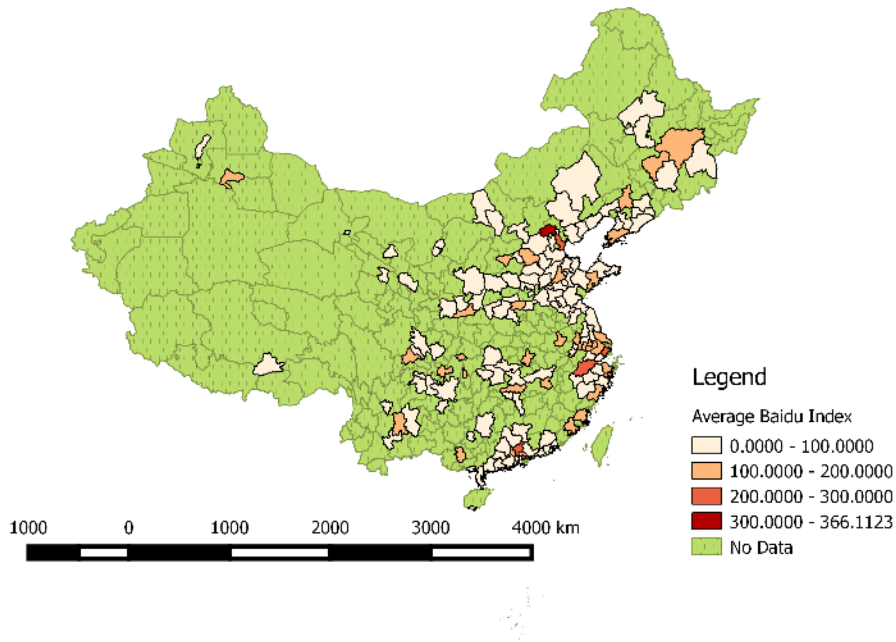
Zheng, S., C. Sun, and M. E. Kahn (2015). Self-protection investment exacerbates air pollution exposure inequality in urban china. Working Paper 21301, National Bureau of Economic Research.

Figure 1: China's GDP per capita and Overseas Emigration



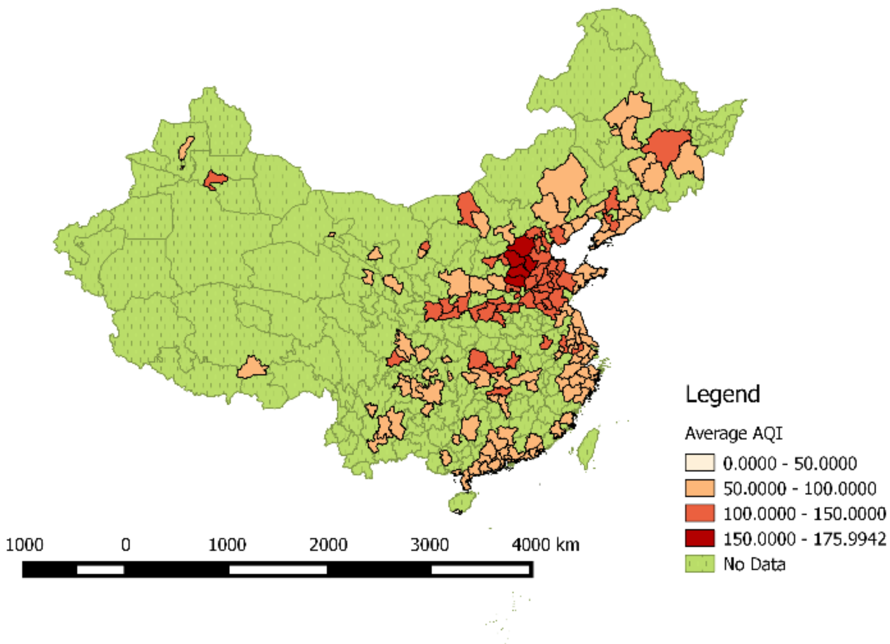
Data source: Emigration data: United Nation Department of Economic and Social Affairs *http : //www.un.org/en/development/desa/population/migration/data/estimates2/estimatesorigin.shtml*; GDP per capita data: the World Bank.

Figure 2: Spatial Distribution of Annual Mean Baidu Index of "Emigration" in 2014



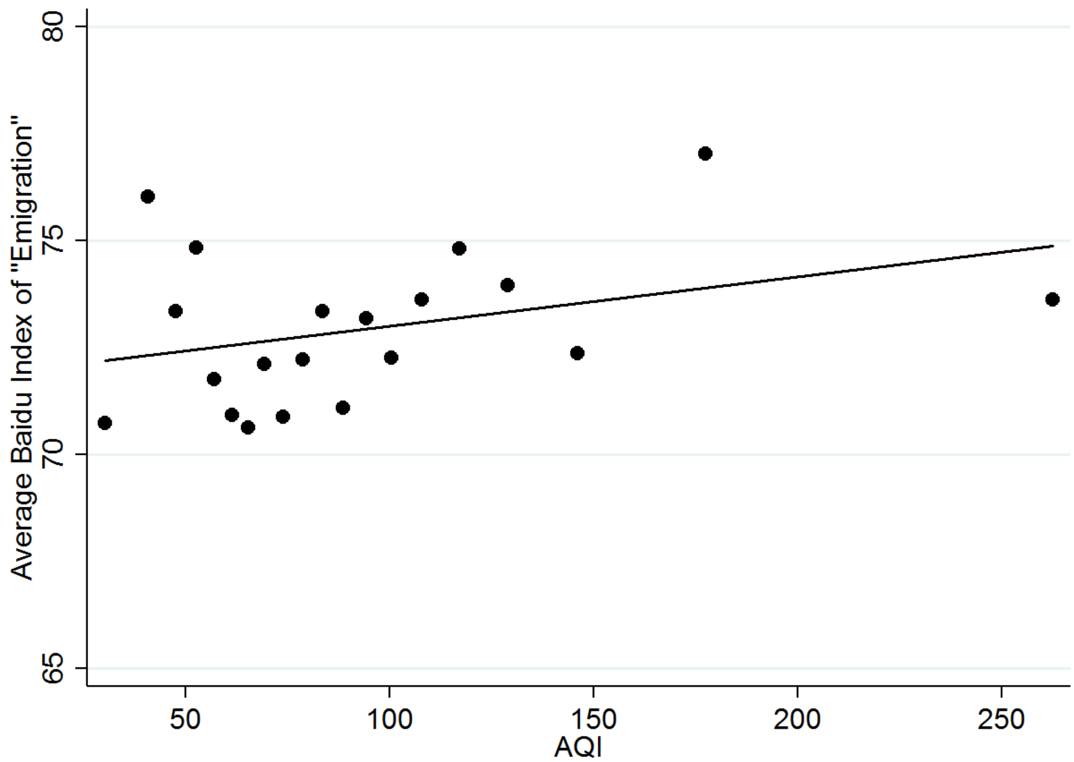
Data source: Emigration Index is collected from Baidu; map is provided by China Data Online.

Figure 3: Spatial Distribution of Annual Mean AQI in 2014



Data source: AQI data are from the Ministry of Environmental Protection in China; map is provided by China Data Online.

Figure 4: City-Level Mean AQI and Emigration Index in 2014



Data source: AQI data are from the Ministry of Environmental Protection in China; Emigration Index is collected from Baidu.
Note: We divide the sample into 20 sub-groups based on their AQI levels and then present the mean values of search in each group.

Figure 5: Histogram of Baidu Index of “Emigration” in 2014

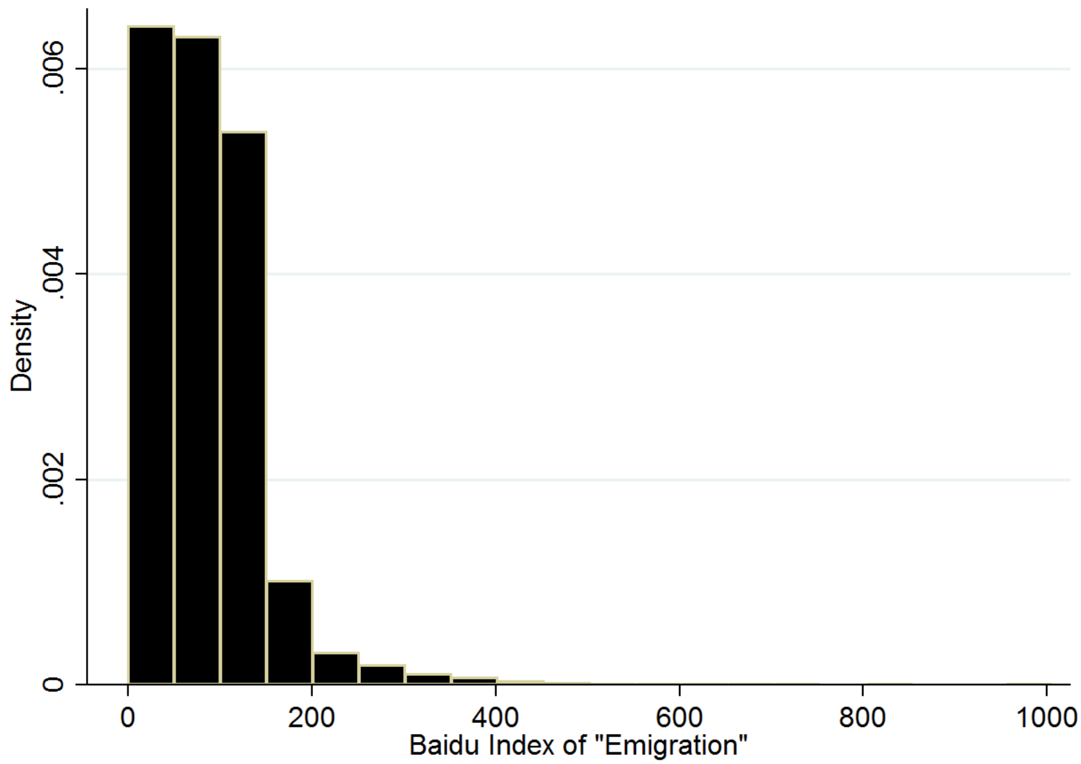


Table 1: Summary Statistics

	description	count	mean	sd	min	max
AQI	Air Quality Index	52515	94.712	54.883	12	500
emigration	Baidu Index of "emigration"	52515	72.917	72.204	0	1006
us_migrate	Baidu Index of "emigrate to the U.S."	52515	26.858	43.611	0	811
can_migrate	Baidu Index of "emigrate to Canada"	52515	37.241	50.945	0	899
au_migrate	Baidu Index of "emigrate to Australia"	52515	34.128	47.280	0	353
nz_migrate	Baidu Index of "emigrate to New Zealand"	52515	42.705	70.954	0	3588
pm2.5	Baidu Index of "PM 2.5"	52515	142.148	407.695	0	44444
smog	Baidu Index of "smog"	52515	97.474	147.343	0	12162
mask	Baidu Index of "mask"	52515	38.406	51.883	0	1812

Note: 1. The full dataset covers 153 prefecture cities which have daily AQI readings in 2014. 2. AQI data is from the Ministry of Environmental Protection in China; 3. All the Baidu indices are collected from Baidu.

Table 2: Dynamic Effects of Air Pollution on Baidu Index of "Emigration"

Dependent Variable: Baidu Index of "Emigration"			
	(1)	(2)	(3)
AQI/100	0.019**	0.002	-0.000
	(0.009)	(0.007)	(0.008)
l1AQI/100	0.025***	0.015**	0.015**
	(0.007)	(0.007)	(0.007)
l2AQI/100	0.011	0.004	0.004
	(0.008)	(0.007)	(0.008)
l3AQI/100	0.001	-0.005	-0.008
	(0.007)	(0.007)	(0.007)
l4AQI/100	0.015**	0.009	0.009
	(0.007)	(0.007)	(0.007)
l5AQI/100	-0.000	-0.006	-0.002
	(0.008)	(0.007)	(0.007)
l6AQI/100	0.013	0.005	0.005
	(0.010)	(0.007)	(0.008)
Month FE	Yes	No	No
City FE	Yes	No	No
City-month FE	No	Yes	No
City-week FE	No	No	Yes
Public Holiday FE	Yes	Yes	Yes
Day-of-Week FE	Yes	Yes	Yes
Weather	Yes	Yes	Yes
N	41124	40725	38929

Notes: 1. AQI/100 is measured by the daily AQI readings divided by 100; 2. Weather variables, including wind speed, precipitation, minimum temperature, maximum temperature, average temperature and dew point, have been controlled in all the regressions. 3. Public holiday fixed effect and day-of-week fixed effect have also been controlled in all the specifications. 4. The first specification controls for month and city fixed effect; the second specification allows different patterns of emigration searches in different months of the same city by controlling for city-month fixed effect; the last specification further controls for city-week fixed effect to eliminate the omitted variables at the city-week level. 5. *** $p \leq 0.01; **p \leq 0.05; *p \leq 0.1$.

Table 3: Impact of Air Pollution on Baidu Index of “Emigration”

Dependent Variable: Baidu Index of ”Emigration”			
	(1)	(2)	(3)
l1AQI/100	0.047***	0.023***	0.025***
	(0.009)	(0.005)	(0.006)
Month FE	Yes	No	No
City FE	Yes	No	No
City-month FE	No	Yes	No
City-week FE	No	No	Yes
Public Holiday FE	Yes	Yes	Yes
Day-of-Week FE	Yes	Yes	Yes
Weather	Yes	Yes	Yes
N	52239	51941	49671

Notes: 1. l1AQI/100 is measured by the lagged one day AQI readings divided by 100; 2. Weather variables, including wind speed, precipitation, minimum temperature, maximum temperature, average temperature and dew point, have been controlled in all the regressions. 3. Public holiday fixed effect and day-of-week fixed effect have also been controlled in all the specifications. 4. The first specification controls for month and city fixed effect; the second specification allows different patterns of emigration searches in different months of the same city by controlling for city-month fixed effect; the last specification further controls for city-week fixed effect to eliminate the omitted variables at the city-week level. 5. *** $p \leq 0.01; **p \leq 0.05; *p \leq 0.1$.

Table 4: Impact of Air Pollution on Baidu Index of “Emigration”: Nonlinear Effects

	Dependent Variable: Baidu Index of ”Emigration”		
	(1)	(2)	(3)
l1AQI_level2	-0.002	0.008	0.008
	(0.009)	(0.008)	(0.008)
l1AQI_level3	0.010	0.013	0.007
	(0.011)	(0.009)	(0.009)
l1AQI_level4	0.036**	0.019	0.015
	(0.015)	(0.012)	(0.013)
l1AQI_level5	0.108***	0.055***	0.057***
	(0.023)	(0.015)	(0.015)
l1AQI_level6	0.127***	0.079***	0.096***
	(0.029)	(0.025)	(0.027)
Month FE	Yes	No	No
City FE	Yes	No	No
City-month FE	No	Yes	No
City-week FE	No	No	Yes
Public Holiday FE	Yes	Yes	Yes
Day-of-Week FE	Yes	Yes	Yes
Weather	Yes	Yes	Yes
N	52239	51941	49671

Notes: 1. l1AQI_level2 to l1AQI_level6 are dummy variables indicating the lagged one day AQI level based on the pollution levels defined by the government. l1AQI_level1 is the default category. 2. Weather variables, including wind speed, precipitation, minimum temperature, maximum temperature, average temperature and dew point, have been controlled in all the regressions. 3. Public holiday fixed effect and day-of-week fixed effect have also been controlled in all the specifications. 4. The first specification controls for month and city fixed effect; the second specification allows different patterns of emigration searches in different months of the same city by controlling for city-month fixed effect; the last specification further controls for city-week fixed effect to eliminate the omitted variables at the city-week level. 5. *** $p \leq 0.01; **p \leq 0.05; *p \leq 0.1$.

Table 5: Impact of Air Pollution on Baidu Index of “Emigration” using U.S. Embassy Data

	Dependent Variable: Baidu Index of ”Emigration”			
	(1)	(2)	(3)	(4)
PM2.5 Definitions	max	working hour max	mean	working hour mean
l1PM2.5/100	0.003	0.010	0.012	0.014
	(0.007)	(0.008)	(0.012)	(0.011)
City-week FE	Yes	Yes	Yes	Yes
Public Holiday FE	Yes	Yes	Yes	Yes
Day-of-Week FE	Yes	Yes	Yes	Yes
Weather	Yes	Yes	Yes	Yes
N	1798	1790	1798	1790

Notes: 1. l1PM2.5/100 is measured by the lagged one day PM2.5 readings divided by 100; 2. Weather variables, including wind speed, precipitation, minimum temperature, maximum temperature, average temperature and dew point, have been controlled in all the regressions. 3. City-week fixed effect, public holiday fixed effect and day-of-week fixed effect have also been controlled in all the specifications. 4. *** $p \leq 0.01; **p \leq 0.05; *p \leq 0.1$.

Table 6: Impact of Air Pollution on Baidu Index of “Emigration” using U.S. Embassy Data: Nonlinear Effect

	Dependent Variable: Baidu Index of ”Emigration”			
	(1)	(2)	(3)	(4)
PM2.5 Definitions	max	working hour max	mean	working hour mean
l1PM2.5_level2	0.024	0.012	0.002	0.014
	(0.019)	(0.012)	(0.011)	(0.013)
l1PM2.5_level3	0.009	0.003	-0.029*	-0.025*
	(0.018)	(0.013)	(0.016)	(0.013)
l1PM2.5_level4	-0.015	0.002	0.007	-0.016
	(0.023)	(0.016)	(0.022)	(0.018)
l1PM2.5_level5	-0.023	-0.015	-0.007	0.057
	(0.023)	(0.025)	(0.029)	(0.042)
l1PM2.5_level6	0.002	-0.016	0.034	0.008
	(0.027)	(0.029)	(0.055)	(0.040)
l1PM2.5_level7	0.057*	0.080**	0.101**	0.110**
	(0.033)	(0.038)	(0.040)	(0.044)
City-week FE	Yes	Yes	Yes	Yes
Public Holiday FE	Yes	Yes	Yes	Yes
Day-of-Week FE	Yes	Yes	Yes	Yes
Weather	Yes	Yes	Yes	Yes
N	1798	1790	1798	1790

Notes: 1. l1PM2.5_level2 to l1PM2.5_level7 are dummy variables indicating the lagged one day PM 2.5 levels for each 50 points increment of PM 2.5. l1PM2.5_level1 is the default category. l1PM2.5_level7 is the highest level, with values of 300 or above. 2. Weather variables, including wind speed, precipitation, minimum temperature, maximum temperature, average temperature and dew point, have been controlled in all the regressions. 3. City-week fixed effect, public holiday fixed effect and day-of-week fixed effect have also been controlled in all the specifications. 4. *** $p \leq 0.01; **p \leq 0.05; *p \leq 0.1$.

Table 7: Impact of Air Pollution on Baidu Index of "Emigration" by Destination Countries

Dependent Variable:	Baidu Index of "Emigration" by Destination Countries			
	US	Canada	Australia	New Zealand
	(1)	(2)	(3)	(4)
l1AQI/100	0.026**	0.003	0.005	-0.004
	(0.012)	(0.008)	(0.009)	(0.011)
N	35558	40069	40025	41981
City-Week FE	Yes	Yes	Yes	Yes
Public Holiday FE	Yes	Yes	Yes	Yes
Day-of-Week FE	Yes	Yes	Yes	Yes
Weather	Yes	Yes	Yes	Yes

Notes: 1. l1AQI/100 is measured by the lagged one day AQI readings divided by 100; 2. Weather variables, including wind speed, precipitation, minimum temperature, maximum temperature, average temperature and dew point, have been controlled in all the regressions. 3. Public holiday fixed effect and day-of-week fixed effect have also been controlled in all the specifications. 4. The first specification controls for month and city fixed effect; the second specification allows different patterns of emigration searches in different months of the same city by controlling for city-month fixed effect; the last specification further controls for city-week fixed effect to eliminate the omitted variables at the city-week level. 5. *** $p \leq 0.01; **p \leq 0.05; *p \leq 0.1$.

Table 8: Impact of Air Pollution on Baidu Index of "Emigration" by Destination Countries: Nonlinear Effect

	Dependent Variable: Baidu Index of "Emigration"			
	US	Canada	Australia	New Zealand
	(1)	(2)	(3)	(4)
l1AQI_level2	-0.039**	0.004	0.019	-0.004
	(0.018)	(0.014)	(0.016)	(0.019)
l1AQI_level3	-0.028	0.017	0.031*	0.032
	(0.021)	(0.016)	(0.018)	(0.023)
l1AQI_level4	-0.002	0.012	0.019	-0.003
	(0.026)	(0.021)	(0.022)	(0.025)
l1AQI_level5	0.052*	0.002	0.001	-0.025
	(0.031)	(0.023)	(0.026)	(0.028)
l1AQI_level6	0.093*	-0.003	0.012	0.002
	(0.055)	(0.039)	(0.041)	(0.047)
N	35558	40069	40025	41981
City-Week FE	Yes	Yes	Yes	Yes
Public Holiday FE	Yes	Yes	Yes	Yes
Day-of-Week FE	Yes	Yes	Yes	Yes
Weather	Yes	Yes	Yes	Yes

Notes: 1. l1AQI_level2 to l1AQI_level6 are dummy variables indicating the lagged one day AQI level based on the pollution levels defined by the government. 2. Weather variables, including wind speed, precipitation, minimum temperature, maximum temperature, average temperature and dew point. 3. City-week fixed effect, public holiday fixed effect and day-of-week fixed effect have also been controlled in all the specifications. 4. *** $p \leq 0.01; **p \leq 0.05; *p \leq 0.1$.

Table 9: Impact of Air Pollution on Baidu Index of "Emigration" by Cities

Dependent Variable: Baidu Index of "Emigration"				
	Beijing	Shanghai	Guangzhou	Shenzhen
	(1)	(2)	(3)	(4)
l1AQI/100	0.029**	-0.025	0.078	-0.015
	(0.011)	(0.017)	(0.049)	(0.049)
N	341	340	335	335
Week FE	Yes	Yes	Yes	Yes
Public Holiday FE	Yes	Yes	Yes	Yes
Day-of-Week FE	Yes	Yes	Yes	Yes
Weather	Yes	Yes	Yes	Yes

Notes: 1. l1AQI/100 is measured by the lagged one day AQI readings divided by 100. 2. Weather variables, including wind speed, precipitation, minimum temperature, maximum temperature, average temperature and dew point, have been controlled in all the regressions. 3. Week fixed effect, public holiday fixed effect and day-of-week fixed effect have also been controlled in all the specifications. 4. *** $p \leq 0.01; **p \leq 0.05; *p \leq 0.1$.

Table 10: Impact of Air Pollution on Baidu Index of "Emigration" by Cities: Nonlinear Effect

	Dependent Variable: Baidu Index of "Emigration"			
	Beijing	Shanghai	Guangzhou	Shenzhen
	(1)	(2)	(3)	(4)
l1AQI_level2	-0.018	0.017	0.036	-0.007
	(0.026)	(0.017)	(0.031)	(0.017)
l1AQI_level3	-0.017	0.016	0.070	-0.006
	(0.033)	(0.019)	(0.044)	(0.047)
l1AQI_level4	0.023	-0.028	0.186**	
	(0.033)	(0.028)	(0.085)	
l1AQI_level5	-0.004	-0.032		
	(0.036)	(0.070)		
l1AQI_level6	0.127***			
	(0.046)			
N	341	340	335	335
Week FE	Yes	Yes	Yes	Yes
Public Holiday FE	Yes	Yes	Yes	Yes
Day-of-Week FE	Yes	Yes	Yes	Yes
Weather	Yes	Yes	Yes	Yes

Notes: 1. l1AQI_level2 to l1AQI_level6 are dummy variables indicating the lagged one day AQI level based on the pollution levels defined by the government. 2. Weather variables, including wind speed, precipitation, minimum temperature, maximum temperature, average temperature and dew point, have been controlled in all the regressions. 3. City-by-week fixed effect, public holiday fixed effect and day-of-week fixed effect have also been controlled in all the specifications. 4. *** $p \leq 0.01; **p \leq 0.05; *p \leq 0.1$.

Table 11: Impact of Air Pollution on Other Keyword Searches

	Dependent Variable: Baidu Index		
	(1)	(2)	(3)
	Keyword: Smog		
l1AQI/100	0.271***	0.253***	0.195***
	(0.049)	(0.028)	(0.026)
N	52239	51945	50329
	Keyword: PM 2.5		
l1AQI/100	0.321***	0.292***	0.233***
	(0.068)	(0.043)	(0.036)
N	52239	52161	51615
	Keyword: Mask		
l1AQI/100	0.127***	0.124***	0.103***
	(0.029)	(0.019)	(0.017)
N	52239	50671	42248
Month FE	Yes	No	No
City FE	Yes	No	No
City-month FE	No	Yes	No
City-week FE	No	No	Yes
Public Holiday FE	Yes	Yes	Yes
Day-of-Week FE	Yes	Yes	Yes
Weather	Yes	Yes	Yes

Notes: 1. AQI/100 is measured by the daily AQI readings divided by 100. 2. Weather variables, including wind speed, precipitation, minimum temperature, maximum temperature, average temperature and dew point, have been controlled in all the regressions. 3. Public holiday fixed effect and day-of-week fixed effect have also been controlled in all the specifications. 4. The first specification controls for month and city fixed effect; the second specification allows different patterns of emigration searches in different months of the same city by controlling for city-month fixed effect; the last specification further controls for city-week fixed effect to eliminate the omitted variables at the city-week level. 5. *** $p \leq 0.01; **p \leq 0.05; *p \leq 0.1$.

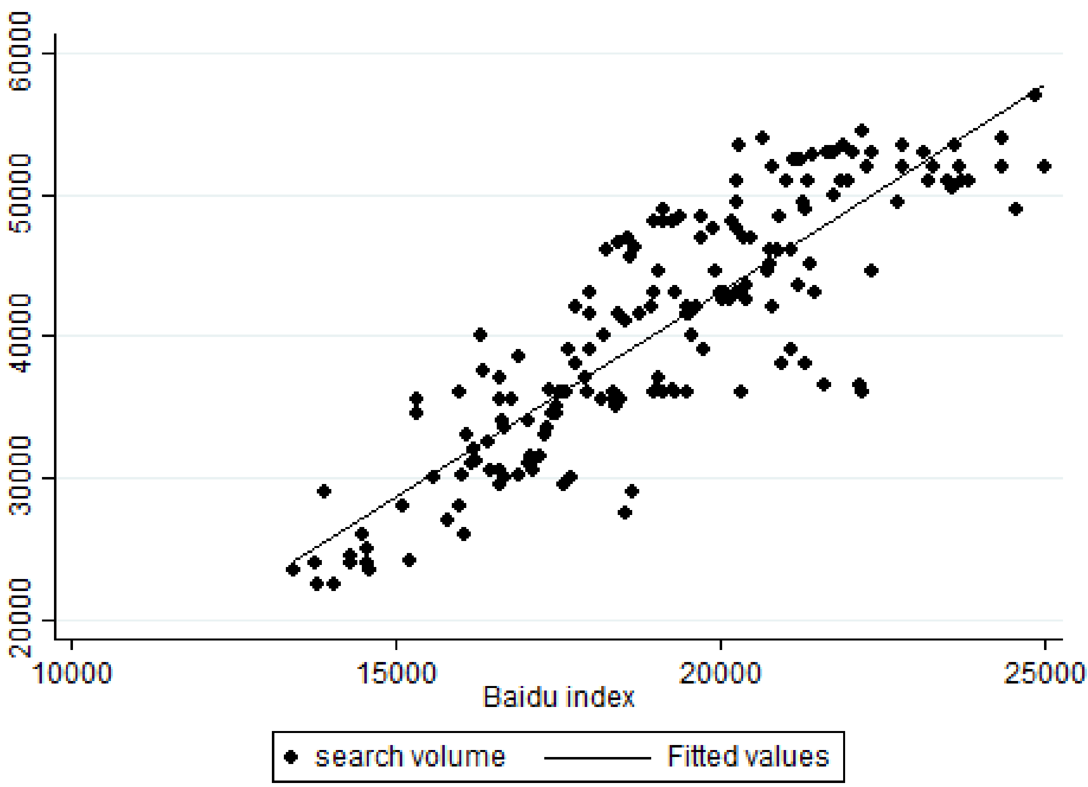
Table 12: Robustness Check: Controlling for Searches on Anti-Corruption Campaign

	Dependent Variable: Baidu Index of "Emigration"						
	(1)	(2)	(3)		(4)	(5)	(6)
l1AQI/100	0.046***	0.023***	0.025***	l1AQI/100_level2	-0.003	0.008	0.007
	(0.009)	(0.005)	(0.006)		(0.009)	(0.008)	(0.008)
				l1AQI/100_level3	0.009	0.013	0.007
					(0.011)	(0.009)	(0.009)
				l1AQI/100_level4	0.034**	0.018	0.015
					(0.015)	(0.012)	(0.013)
				l1AQI/100_level5	0.105***	0.054***	0.057***
					(0.023)	(0.014)	(0.015)
				l1AQI/100_level6	0.124***	0.077***	0.096***
					(0.029)	(0.025)	(0.027)
N	52209	51911	49635		52209	51911	49635
Month FE	Yes	No	No		Yes	No	No
City FE	Yes	No	No		Yes	No	No
City-month FE	No	Yes	No		No	Yes	No
City-week FE	No	No	Yes		No	No	Yes

Notes: 1. Panel A replicates the three columns in Table 2 with the inclusion of search index on anti-corruption campaigns. 2. Panel B replicates the three columns in Table 3 with the inclusion of search index on anti-corruption campaigns. 3. *** $p \leq 0.01$; $** p \leq 0.05$; $*p \leq 0.1$.

A Figures

Figure A.1: Linear Relationship between Search Volume and Baidu Index



Data source: zhihu.com