PEDL Research Papers

This research was partly or entirely supported by funding from the research initiative Private Enterprise Development in Low-Income Countries (PEDL), a Foreign, Commonwealth & Development Office (FCDO) funded programme run by the Centre for Economic Policy Research (CEPR).

Firm Presence, Environmental Quality, and Economic Activity: Evidence from Randomized Relocation of Small Firms

Michael Gechter* and Namrata Kala†

Abstract

Firm location decisions are one of the most important decisions managers make, optimizing factors such as proximity to customers, suppliers, and useful information. The inherent endogeneity of firm location decisions renders estimating the impact of firm presence difficult. In this paper, we use an environmental relocation policy that randomly moved over 20,000 small firms operating within city limits in New Delhi to industrial areas outside the city over several years. We find that a reduction in firm presence has no impact on measured air quality, but is costly for firms: relocated firms have a high rate of exit, which increases in the distance relocated. The lack of effects on ambient fine PM can be rationalized by the fact that industrial sources are not a large contributor to fine PM in Delhi. There are no long-term effects on population density or composition.

*Penn State University; Y-RISE. Email: mdg5396@psu.edu

†MIT Sloan School of Management; NBER; BREAD; CEPR; J-PAL. Email: kala@mit.edu.

We are grateful to the International Growth Center (IGC), MIT Sloan School of Management, the National Science Foundation (NSF), Penn State University, Private Enterprise Development in Low-Income Countries (PEDL) and Structural Transformation and Economic Growth (STEG) for financial support. We would like to thank Jonathan Eaton, Glenn Ellison, Bob Gibbons, Claudia Steinwender, Daniel Xu, and seminar audiences at MIT, Penn State University of California Berkeley for helpful comments, and Christina Antonakakis, David Cai, Zixu Chen, Valentina Goetz, Hanna Marsho, Astha Vohra, and Charlotte Wachter for excellent research assistance. All errors are our own.

1 Introduction

Firm location decisions are one of the most important decisions managers make, optimizing factors such as proximity to customers, suppliers, and useful information. At the same time, these decisions may have spillovers on local neighborhoods, by providing access to jobs, impacting environmental quality, and contributing to local economic activity (Glaeser, 2010). For this reason, numerous policies attempt to change location choices of firms, for instance, place-based policies that incentivize firms to locate in particular areas. The inherent endogeneity of firm location decisions renders estimating the impact of firm presence on the local economy difficult. Policies that shock firm location decisions can help overcome this, but in many cases, may be bundled with other features – for instance, place-based policies that incentivize firms to locate in certain areas often provide local infrastructure or tax benefits.

In this paper, we study a policy that is a common environmental policy tool in developing countries, namely firm relocation. These polices act as a form of ex-post zoning, moving existing firms operating in high-density areas to more remote locations. We study an environmental relocation policy that moved over 20,000 small firms operating within city limits in New Delhi to industrial areas outside the city over several years. A unique feature of this policy is that due to a shortage of industrial plots when relocation began, plot allotment was done via a series of lotteries between 2000 and 2005, with firms actually moving between 2006 and 2010. This generates random variation in firm presence over the time period, between neighborhoods with a greater number of firms receiving a plot earlier in the process, and those with a greater number of firms receiving a plot later in the process (condtional on the total number of firms relocated from a neighborhood).

Combining several data sources with administrative data on the relocation policy, we estimate the impact of firm presence on four main sets of outcomes: first, on environmental quality (neighborhood-level fine particulate matter concentrations), second, the size and composition of the resident population, third, on overall economic activity using Economic Census data, and fourth, on the relocated firms themselves. Controlling for the total number of firms that were relocated from a neighborhood, we compare neighborhoods that on average relocated a higher number of these firms earlier vs. later, with the timing randomly generated by the allotment of plots to firms via lottery.

We find that a reduction in firm presence, as measured by earlier neighborhood-level movement of firms, has no impact on measured air quality, but is costly for firms, as measured by firm survival. There are also no discernible long-term impacts on population density or the number of firms per square kilometer. These results are consistent with the fact that industrial emissions contribute less than 7% to overall fine PM emissions in Delhi (Sharma et al., 2016), and so this policy did not seem to translate into discernible impacts on ambient fine PM. In contrast, we are able to show how survival probabilities for relocated firms decrease in the distance they are relocated, indicating the importance of endogenous location choices and the costs of being moved from that location.

The paper is related to several literatures. First, we contribute to the literature on firm presence, environmental amenities, and real estate values. Location restrictions that seek to limit pollution exposure also have a long history, starting with the first zoning laws introduced in the early 20th century in New York in part to improve environmental quality Wilson et al. (2008). Harrison et al. (2015) study how Indian Supreme Court-ordered Action Plans for 17 cities affected firm decisions in corresponding districts to exit or invest in pollution abatement. A primary means to reduce pollution mentioned in these action plans was relocation of polluting industries to certain designated areas–for instance, 14 of 17 Action Plans in major cities mention industrial relocation (Harrison et al., 2015). We estimate the causal impacts of industrial relocation, which is a policy tool used by several countries, including India and China (Zhao and Yin, 2011).

More broadly, we contribute to the literature on the impact of environmental regulation in developing countries (Do et al., 2018; Duflo et al., 2013; Greenstone and Hanna, 2014; He et al., 2020). We use a unique experiment that forces firms to move out of populated neighborhoods, using the timing of lotteries to generate exogenous variation in firm presence. We focus on estimating how a policy that targets small firms' removal from populated areas impacts air quality, neighborhood population density, real estate values, and firm outcomes. We find limited impacts on these outcomes, except for relatively high costs for firms. The lack of effects on pollution can be rationalized by the fact that industrial emissions are not a large contributor to fine PM in Delhi (Sharma et al., 2016).

2 Context and Relocation Policy

In 1999, the Supreme Court mandated the relocation of manufacturing firms in Delhi that were operating in residential areas (there were exemptions for certain types of household industries). These firms comprised a large range of small manufacturing firms, including automobile parts, food processing, and rubber and plastics products. The Government started developing three industrial areas on the edges of the city to house these firms. However, since the number of industrial plots was limited (more were developed over time), they were allocated via a series of lotteries (about 5-6 large and many small) between 2001 and 2015.

Over 20,000 firms were relocated over the program time period, with the majority of relocation happening in the early to mid 2000s, and only about 500-600 firms relocated 2010 or later. Each firm was allotted an industrial plot ranging from 28 m^2 (which was a spot in a building housing several small firms) to standalone plots of 250 m^2 - the average plot size was between 100 and 150 m^2. Firms were given concessional loans to allow them to build their factories in the allotted plots, and were given leases for these plots. They were not allowed to sell or rent them, and were technically supposed to continue producing the same products they had done while located within Delhi. Of the three industrial areas, the largest one (Bawana) housed the majority of relocated firms (over 13,000).

Firms that were not allotted an industrial plot in the earlier lotteries could continue operating while they waited for a plot, and once allotted a plot had to move their operations within 3 years. In 2005 however, the Supreme Court directed the government to increase the pace of relocation, and so leases began to be given for the industrial areas in 2006. We will show that firms that "won" the lottery earlier also were allotted a lease earlier, generating exogenous variation in the timing of their departure from their original location. All the industrial areas were developed by the government body charged with implementing the relocation policy, and within the industrial area, the plot allocation was random conditional on plot size as well. Figures A1, A2, and A3 show Google Maps images of the largest industrial area where over 13,000 of these firms were relocated for the years 2000, 2001, and 2010, respectively.

3 Data

We combine several data sources to create a neighborhood-level dataset. Our definition of a neighborhood is a 2001 Census ward from the Indian Population Census, which is described in more detail below.

3.1 Population Census

We use the 2001 and 2011 Population Census, which includes information on neighborhood-level total population, population by gender, as well as population composition such as the number of working and non-working residents (a Census neighborhood is called a ward, and we use these interchangeably). Our level of analysis for neighborhood is the 2001 census ward, of which there are 145 in New Delhi. We assign population variables from the 2011 census to the 2001 census ward based on the percentage of area in each 2011 ward that was in the 2001 ward. We use the 2001 data to test for balance i.e. test whether neighborhoods with more firms that were lotteried out earlier are different in terms of baseline population levels. We use the 2011 data to test for any impacts of the relocation policy on population density and composition.

3.2 Air Pollution

3.2.1 Van Donkelaar Data

We use data on annual fine particulate matter (PM 2.5) created by van Donkelaar et al. (2016). The data are available monthly at the 1km by 1km resolution, from 1997-2019. We use 1999 as our baseline year, and 2010 as our endline year, and create neighborhood-level measures of PM 2.5 concentrations by taking the mean of all points within a neighborhood, as well as as minimum and maximum values within the neighborhood. These data are constructed by combining Aerosol Optical Depth (AOD) data from several satellite sources, followed by a calibration to pollution monitor data using a Geographically Weighted Regression (GWR) (for more details, please refer to van Donkelaar et al. (2016)).

3.2.2 Aerosol Optical Depth (AOD) Data

We use Aerosol Optical Depth (AOD) data as a proxy for fine PM, and calculate neighborhood-level statistics for two measures of AOD-at 0.47 micrometers and at 0.55 micrometers from satellite images between 2000 and 2019. We use the MCD19A2 Version 6 data product, which is a Moderate Resolution Imaging Spectroradiometer (MODIS) Terra and Aqua combined Multi-angle Implementation of Atmospheric Correction (MAIAC) Land Aerosol Optical Depth (AOD) gridded Level 2 product. These are available daily at the 1km by 1 km level globally. As with the fine PM 2.5, we calculate neighborhood-level statistics including average, minimum and maximum values of these AOD measures at the neighborhood-year level. We present results using the first measure in the paper, but results are similar using the second measure as well.[1]

3.3 Real Estate Values

To obtain a proxy for real estate values, we rely on a range of values that the Delhi Government assigns to micro-regions in Delhi, which are used to determine property taxes.[2] Each micro-region (named "colony") is assigned a grade from A to H, with A being the

[1]These estimates are omitted for brevity, but available upon request.

[2]Real estate prices are unfortunately not available for Delhi during this time period.

highest price regions and H the lowest. These grades, called "circle rate categories", are periodically re-graded by the government, and we obtain these data for 2011 and 2014.[3] We use data from 2011 as an outcome in the analysis, but results are similar if we use data from 2014 instead.[4] The 2011 circle rate data had information on 2,638 micro-regions. We geo-coded these using Google maps, and assigned each region to the 2001 Census neighborhood that it lay within. From these, we are able to calculate the average, minimum, and maximum values of these circle rates for each neighborhood.

3.4 Administrative Data on Relocation Policy

The administrative data on the relocation policy, is available from the government body that was responsible for the relocation, Delhi State Industrial and Infrastructure Development Corporation Ltd. (DSIIDC). The data include firm name, original address from where they were relocated, details such as applicant name, date of the lottery when they were allotted a plot, firm products, as well as final location in the industrial area. It also includes details on the timing of the dates when the firm received the formal allotment letter, when the firm's lease began, and when they started paying rent.

Geocoding the original addresses using Google's API, we assign a firm to a neighborhood. Combining this information with the timing of when a firm was given a plot, we create a neighborhood-level dataset of the number of firms that were allotted a plot via lottery each year between 2000 and 2005 (95% of firms were assigned a plot via lottery by then), as well as the number of firms whose lease began in each year between 2006 and 2010. Figure 1 presents the cumulative probability of having won a plot lottery by a given year as well as the probability of having initiated a lease in an industrial area by year, showing a positive relationship between the two measures.

3.5 Surveyor Vists to Baseline Addresses of Relocated Firms

To ensure that any measurement error on assigning firms to neighborhoods is independent of lottery timing, we collected data for about 14,483 firms which were geoded in-person by surveyors (about 63% of the total sample).[5] Of these, Google and the surveyors place an address in the same ward about 73% of the time (9,883 of 13,511). The probability that the neighborhood is not the same according to Google and the surveyors is not correlated with the year of the lottery (see Table A5), nor is the probability that geocodes are

[3]We were unable to get data earlier than 2011.

[4]These are omitted for brevity, but available upon request.

[5]We sent surveyors to about 15,811 firms but could not find about 1,328 addresses.

missing from either of these sources. Furthermore, our measure of early relocation, which is the number of firms relocated by 2003 according to Google maps and the surveyors are highly correlated, with a correlation of 0.98.

4 Empirical Strategy

To estimate how having a greater number of firms relocated early impacted neighborhood-level outcomes (our proxy for firm presence), we use the following difference-in-differences specification:

$$\begin{aligned} Y_{it} = &\alpha + \beta_t \text{Log(Firms per Sq Km Relocated by 2003)}_i \times \mathbb{1}[\text{Year} > 2005]_t + \psi_i + \delta_t \\ &+ \gamma_t \text{Log(Total Number of Firms Relocated per Sq Km)}_i \times \mathbb{1}[\text{Year} > 2005]_t + \epsilon_{it} \end{aligned} \tag{1}$$

where Y_{it} is the outcome for neighborhood i at time t. The main coefficients of interest are β_t, which measure the marginal impact of a greater number of firms relocated earlier in the process i.e. by 2003, in each year. Note that this timing is plausibly exogenous due to the lottery governing the allotment of plots, and the fact that a delay in winning the lottery resulted in a delay in the beginning of a firm's lease on average (as shown in Table 1). We control flexibly for the total firms relocated per sq km, by including interactions of this variable with year fixed effects, and also include neighborhood and year fixed effects (ψ_i and δ_t, respectively). We consider 2005 to be the year before firms' leases began to be approved in the industrial area, and so hypothesize that treatment effects should be zero prior to that year (since neighborhoods with early vs later relocation should have similar pre-trends in their outcomes), and any differences would be visible starting the year after i.e. 2006.[6] Therefore, the interactions terms are omitted for the year 2005.

We also show robustness to defining the neighborhood to be a 2km by 2km grid rather than the Census neighborhood for outcomes that are defined at a more spatially disaggregated level (air pollution and long-run firm presence from the 2012 Directory of Establishments). We cluster standard errors at the level of the neighborhood. The analogous differences-in-differences specification is as follows:

$$\begin{aligned} Y_{it} = &\alpha + \beta \text{Log(Firms per Sq Km Relocated by 2003)}_i \times \mathbb{1}[\text{Year} > 2005]_t + \psi_i + \delta_t \\ &+ \gamma \text{Log(Total Number of Firms Relocated per Sq Km)}_i \times \mathbb{1}[\text{Year} > 2005]_t + \epsilon_{it} \end{aligned} \tag{2}$$

[6] A small number of firms, 106, were given leases in 2005, with the implementation beginning on a full scale in 2006, when 2,950 firms were given leases. Our results are similar if we use 2004 to be the omitted year.

where β is the coefficient of interest.

For cross-sectional outcomes such as longer-run firm presence and real estate values, we estimate the following specification at the neighborhood-level:

$$\begin{aligned} Y_i = \gamma + \pi \text{Log(Firms per Sq Km Relocated by 2003)}_i + \\ \theta \text{Log(Total Firms per Sq Km Relocated)}_i + \phi \text{Log (Area-Sq Km)}_i + v_i, \end{aligned} \tag{3}$$

where π is the coefficient of interest. We control for both the density of relocated firms in total for neighborhood i, as well as the area.[7]

Furthermore, since we use the timing generated by the lottery to proxy for firm presence, we also show that a greater number of firms lotteried by 2003 strongly predicts the number of firms with a lease in the industrial area starting by 2010 or 2015– this indicates that the timing of the lottery governed the timing of the leases being issued and therefore the timing of firms' movement to the industrial area. This indicates that the timing of the lottery of the plot assignment governed the timing of the actual movement to the industrial area. We estimate this "first-stage" equation both at the individual-firm level, as well as at the neighborhood-level, the latter being the relevant unit of analysis for neighborhood-level effects. The individual specification is given by:

$$\text{Year of Lease Execution}_j = \alpha + \mu \text{Year of Lottery Win}_j + \epsilon_j \tag{4}$$

where j denotes the firm. The coefficient of interest is μ, which measures the marginal impact of a one-year delay in "winning" the lottery for firm j. The neighborhood-level specification is given by:

$$\begin{aligned} \text{Log(Firms per Sq Km with Lease by 2010)}_i = \nu + \zeta \text{Log(Firms per Sq Km Relocated by 2003)}_i + \\ \theta \text{Log(Total Firms per Sq Km Relocated)}_i + \phi \text{Log (Area-Sq Km)}_i + v_i \end{aligned} \tag{5}$$

The coefficient of interest is ζ, which estimates the impact of a percent increase in firms "winning" the lottery to move by 2003 on firms with leases in the industrial area by 2010 (we show this effect persists for later in time as well, in 2015).

[7]Note that for robustness checks where we run these regressions at the 2 km by 2km grid-level rather than neighborhood, controlling for area is not necessary, and firm density measures would be the same as total firms measures i.e. firms relocated by 2003 per sq km would give the same estimates as firms relocated by 2003, since all grid squares have the same area.

5 Results

5.1 The Impact of Lottery Timing on Movement to the Industrial Area

Since the empirical strategy relies on the timing of the relocation lotteries causing exogenous variation in earlier vs. later firm removal from a neighborhood, we show that the timing of the lottery impacts actual firm presence in the industrial area in two ways. The first is at the firm-level, and estimates Equation 4. Results are presented in Table 1. We consider three measures of firm departure from its original address and its presence in the industrial area - the year in which it took possession of the plot in the industrial area, the year in which its lease was executed and the year in which its first year of rent became due. Across these three outcomes, we see a positive and statistically significant effect. A one year delay in the firm "winning" the lottery to be relocated causes a 1.1 year delay in the year in which it acquires physical possession of its plot in the industrial area, an effect that is precisely estimated. The other two outcomes show impacts of similar magnitude and precision, indicating that indeed the timing of the lottery impacted impacted a firm's presence in the industrial area (and departure from the original address).

Next, we show that these firm-level effects aggregate up to the neighborhood-level, in Table 2, which present results from estimating Equation 5. Results show that in neighborhoods with more firms winning a lottery by 2003, more firms had leases that began by 2010, and also 2015 (Columns 1 and 2, respectively). We also present results using the proportion of firms in a neighborhood relocated earlier (by 2003), standardized to have 0 mean and variance 1, and find similar results in magnitude and statistical significance. Results at the 2 km grid-level are shown in Table A6, and are similar to the neighborhood-level results. Thus, these results indicate that the lottery timing generated variation in the timing of firm removal, allowing us to estimate the causal impacts of firm presence.

5.2 Air Pollution

We begin by showing impacts for environmental quality, measured by levels of fine PM. Figure 2 shows the event study for the AOD measure (Figures 2a shows average annual values of AOD while, 2b and 2c shows effects for minimum and maximum AOD, respectively.) Neighborhoods with a greater number of firms relocated by 2003 do not have differential pollution before 2005, but also do not show any changes relative to neighborhoods with a smaller number of firms relocated by 2003 post-2005. Difference-in-differences estimates are presented in Table 3, and are consistent with the event study estimates, showing no impact of the departure of these firms.

Results are similar we use the alternative measure of fine PM, from van Donkelaar et al. (2016). Figure 3 shows event study estimates (with Figures 3a showing impacts for annual mean levels, and Figures 3b and 3c for minimum PM and maximum PM, respectively), showing no effect of earlier vs. later relocation.

5.3 Impacts on Neighborhood-level Firm Presence

To test whether neighborhoods with earlier relocation experienced persistent changes to firm presence, we use the specification from Equation 3, using data from the 2012 (sixth) Economic Census (available at the neighborhood-level), as well as the corresponding Directory of Establishments, which has the address for all firms with eight or more employees (we use the geo-coded addresses to assign firms to a neighborhood).[8] Note that as we see in Figure 1 that the large majority of firms had signed a lease in the industrial area by 2012, the effects we measure here principally represent spillover effects from the relocation onto firms other than the relocated ones. This is in contrast to the air pollution effects in the previous subsection which we measure contemporaneously with relocated firms leaving their original neighborhoods.

Results from the Census are presented in Table 7, using the log of the number of firms per sq km as the outcome variable. Column 1 presents estimates for overall firm density, and Columns 2, 3, and 4 restrict the sample to firms only in 3-digit National Industrial Classification (NIC) sectors with at least 50, 100, and 500 relocated firms. We find that firm density is very similar for neighborhoods with a greater number of firms that were relocated by 2003, indicating that earlier relocation did not impact firm presence in any persistent manner. Results using the Directory of Establishments data are shown in Table 8, and are consistent with those in Table 7. We also show that these null results are similar across the firm size distribution- Table A1 presents results using the same specification and sample definitions as Table 8, but only for firms with 30 or fewer employees. Tables A2 and A3 do the same, but for firms with between 30 to 99 employees, and with 100 or more employees, respectively, and show similar results.

[8] The two specifications address different forms of measurement error. The error in the full Economic Census data arises from mapping the 2011 neighborhood-level data to 2001 neighborhoods according to the share of land each 2011 neighborhood takes up in each 2001 neighborhood, implictly assuming uniform firm density. The error in the directory of establishments arises from geocoding mistakes.

5.4 Impact on Neigborhood-level Population and Real Estate Prices

In this section, we present results on the impact of earlier vs. later relocation on population density and composition, as well as the impact on real estate values. Table 5 presents a balance table for population density, as well as the fraction of workers in the population using the 2001 Population Census, since these should not be correlated with the intensity of firm lottery winning by 2003– as expected, we find that this is not the case. Table 6 presents results on these outcomes using the 2011 Census, to test whether earlier relocation led to long-term persistence in neighborhood population. We find that this is not the case, with neither of these impacts being precisely measured. The coefficient on the proportion of workers is negative, indicating that earlier relocation led to more non-workers in these neighborhoods, but is also imprecisely measured. Overall, these results do not indicate that earlier relocation is associated with precise changes to neighborhood-level population density or composition.

In Table A4, we estimate the impact of early relocation on 2011 real estate values, proxied using circle rate categories.[9] Recall that these are the Delhi Government's assessment of real estate values, which assign one of eight values to each small region, and is used for taxation purposes. Since 2001 Census neighborhhods have multiple circle rate regions within them, we estimate impacts on minimum, maximum, and average values at the neighborhood-level. We find no impacts on these outcomes, indicating that early relocation also did not impact real estate values in the longer term.

5.5 Impact on Relocated Firms

In this section, we consider how the relocation impacted firms that were moved outside the city. First, we use exit from the industrial area to show that the probability firms survived in the industrial area in the long run is strongly decreasing in the distance they were moved, highlighting that location choice is an endogenously chosen parameter by firms. Second, we consider two different types of counterfactual death rates for the firm to understand how the policy may have led to a lower firm survival probability.

We use data on firm survival in 2017 from a census conducted by the government in the largest industrial area where the majority of firms were moved, as our measure of firm survival. The data includes information on whether each of the 15,849 firms that were allotted a plot this area (the remainder of the relocated firms were allocated plots in two much smaller industrial areas) were still operating out of the industrial area. The data indicate that the vast majority of firms did not survive the move, with only about

[9] Results are similar if we use 2005 values instead.

26% of firms still present and functioning in the industrial area by 2017. While this is a large number, firms in developing countries have high exit rates on average, and so all of these exits cannot be attributed to the policy. To estimate a counterfactual death rate for these firms, we take two approaches - the first is estimate how the death rate varies as a function of the distance which the firm was moved. In Figure 4, we present evidence that the further a firm was moved, the more likely it is to exit, and Table 9 presents regression results for this relationship. Column 2 of Table 9 includes 2001 Census neighborhood fixed effects, comparing firms from nearby baseline locations. Each kilometer relocated increases the probability of firm exit by 0.009 percentage points, which implies that the average firm, which was relocated 20.37 kilometers, the increase in the probability of firm exit is 18.3 percent higher than a firm that was not relocated at all. This is a substantial increase in the risk of firm exit, although much lower than the average rate of 74%. A second approach compares these exit rates to those estimated for Indian firms more broadly - for instance, Sengupta and Singh (2019) find that the probability of firm survival in India over 20 years for registered firms is about 50%, about a risk of 2.5% each year. Therefore, seventeen years after the policy, about 42.5% of these firms should have exited, yielding a treatment effect on firm death of 31.5%.

To create back of the envelope estimates of profits and wages lost due to the policy, we use data from the Center for Monitoring the Indian Economy (CMIE) on firm wages and profits. Restricting the sample to 2-digit NIC sectors and to years 2000 or earlier (when the policy implementation started), the median firm earns about ₹70,000 in annual profits and pays ₹130,000 in annual wages. Applying the increased probability of firm death of between 18.3 and 31.5% to the 15, 849 firms moved to Bawana, we get between ₹0.24 billion to ₹0.42 billion in lost profits per year, and between ₹0.45 billion to ₹0.78 billion in lost wages per year, respectively (the counterfactual exit rates of 18.3% and 31.5% imply exit of 3,470 and 5,973 firms caused by the policy). If we assume that the survival-distance gradient was similar for firms relocated to the other industrial areas and apply them to all 26,540 firms, these estimates scale up to ₹0.34 billion to ₹0.59 billion for lost profits, and between ₹0.63 and ₹1 billion for lost wages, respectively. Therefore, costs for firms as a result of this policy were considerable in terms of lost profits, even if we assume that workers were able to transition to new jobs, which may not have been the case.

6 Conclusion

Firm location decisions have important spillovers to the neighborhoods they locate in. These spillovers can be positive, generating employment and knowledge flows, or negative, such as increasing pollution exposure. Several policies regulate firm presence to direct these externalities and increase welfare, but the trade-offs in doing so are not well known. We find that small firm presence does not impact neighborhood-level ambient environmental quality in New Delhi, but their removal does impact relocated firms, decreasing their survival probabilities. Furthermore, removal of firms may have important equity implications, by increasing commuting costs or moving costs for workers, as well as impacting the affordability of a neighborhood.

Note that our results do not indicate that relocation itself is not a policy that can have benefits– indeed, policies that focus on relocating large point sources of pollution (such as large polluting plants) that are a significant contributor to environmental degradation would presumably have large welfare gains. However, since the firms in this policy were largely very small, and industrial emissions are a very small contributor to fine PM (Sharma et al., 2016), contributing less than 7% to this burden, relocating these firms does not seem to have led to lowered PM 2.5 emissions. Furthermore, it is also possible that it impacted other pollutants that data availability constraints preclude estimating the impacts for (such as noise or water pollution).

Our results are able to say something definitive about the costs to firms of being relocated, as well as how these increase with distance. These indicate that if relocation is warranted, a focus on relocating firms closer to industrial areas would lower the burden on relocated firms. Which types of industries should be relocated to provide the maximum environmental benefits and minimize costs to firms and workers, and whether lump sum transfers instead of allocating them land in a fixed place is better for firms' survival, remain interesting questions for future work.

References

Do, Q.-T., Joshi, S., and Stolper, S. (2018). Can environmental policy reduce infant mortality? evidence from the ganga pollution cases. *Journal of Development Economics*, 133:306–325.

Duflo, E., Greenstone, M., Pande, R., and Ryan, N. (2013). Truth-telling by third-party auditors and the response of polluting firms: Experimental evidence from india. *The Quarterly Journal of Economics*, 128(4):1499–1545.

Glaeser, E. L. (2010). Introduction. In *Agglomeration Economics*, pages 1–14. University of Chicago Press.

Greenstone, M. and Hanna, R. (2014). Environmental regulations, air and water pollution, and infant mortality in india. *American Economic Review*, 104(10):3038–72.

Harrison, A., Hyman, B., Martin, L., and Nataraj, S. (2015). When do firms go green? comparing price incentives with command and control regulations in india.

He, G., Wang, S., and Zhang, B. (2020). Watering down environmental regulation in china. *The Quarterly Journal of Economics*, 135(4):2135–2185.

Sengupta, R. and Singh, M. (2019). Firm survival in India: Status of firms formed over the last 30 years. *Ideas for India*.

Sharma, S., Mandal, T., Jain, S., Saraswati, Sharma, A., and Saxena, M. (2016). Source apportionment of pm 2.5 in delhi, india using pmf model. *Bulletin of environmental contamination and toxicology*, 97:286–293.

van Donkelaar, A., Martin, R. V., Brauer, M., Hsu, N. C., Kahn, R. A., Levy, R. C., Lyapustin, A., Sayer, A. M., and Winker, D. M. (2016). Global Estimates of Fine Particulate Matter using a Combined Geophysical-Statistical Method with Information from Satellites, Models, and Monitors. *Environmental Science and Technology*, 50:3762–3772.

Wilson, S., Hutson, M., and Mujahid, M. (2008). How planning and zoning contribute to inequitable development, neighborhood health, and environmental injustice. *Environmental Justice*, 1(4):211–216.

Zhao, X. and Yin, H. (2011). Industrial relocation and energy consumption: Evidence from china. *Energy Policy*, 39(5):2944–2956.

Figure 1: Timing of Plot Lottery and Beginning of the Lease

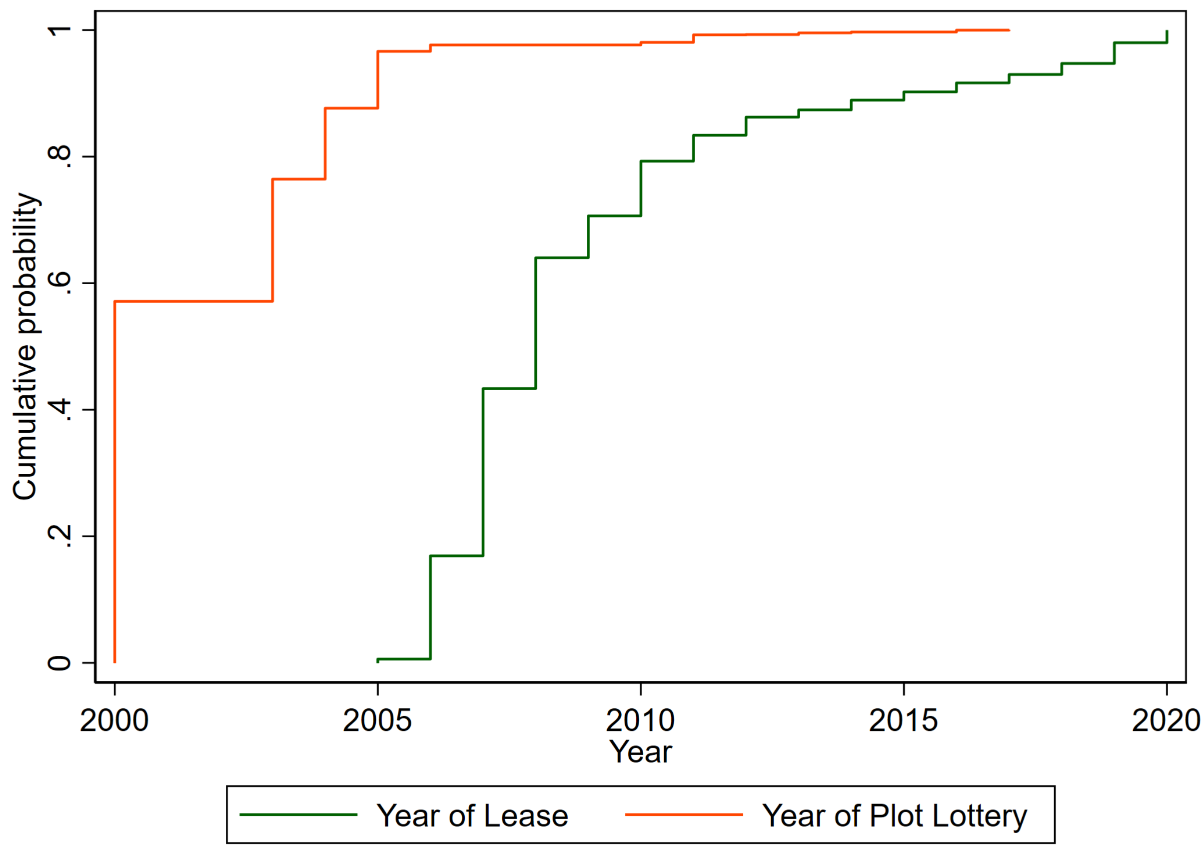
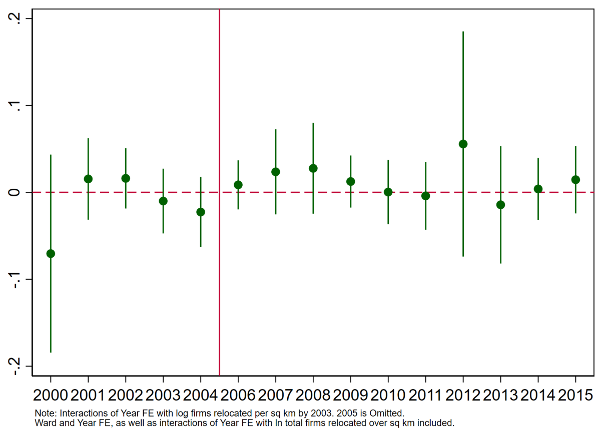
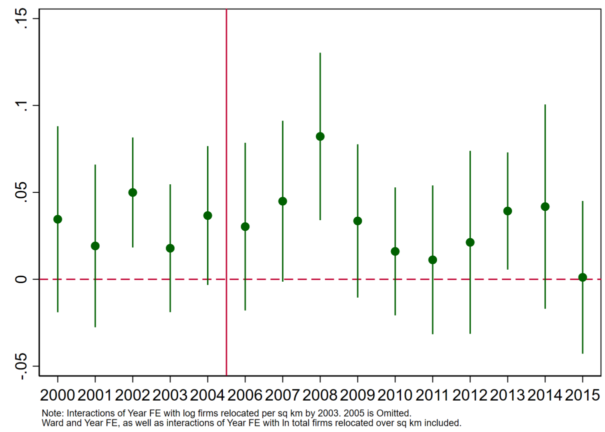


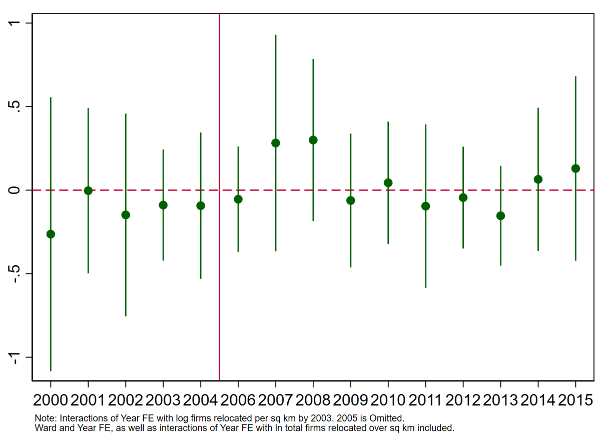
Figure 2: Aerosol Optical Depth (Annual)



(a) Mean

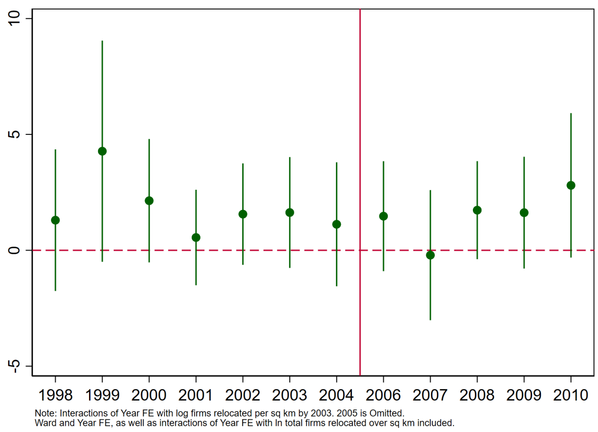


(b) Minimum

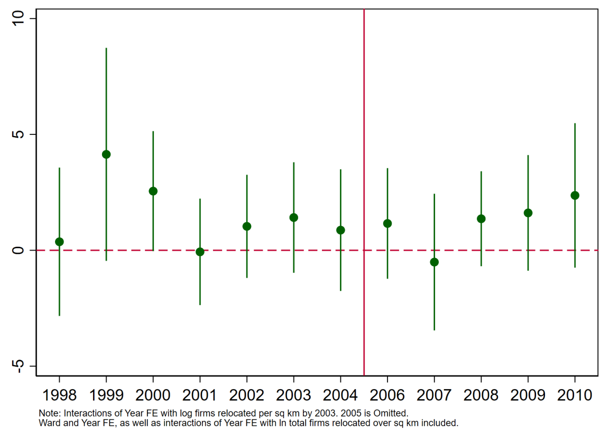


(c) Maximum

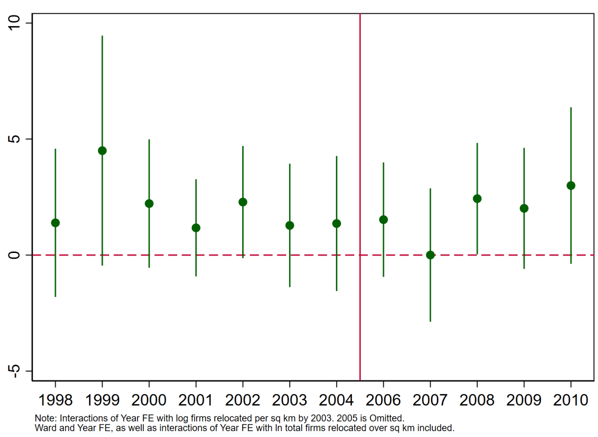
Figure 3: Fine Particulate Matter (Annual)



(a) Mean



(b) Minimum



(c) Maximum

Figure 4: Distance Relocated and Firm Exit

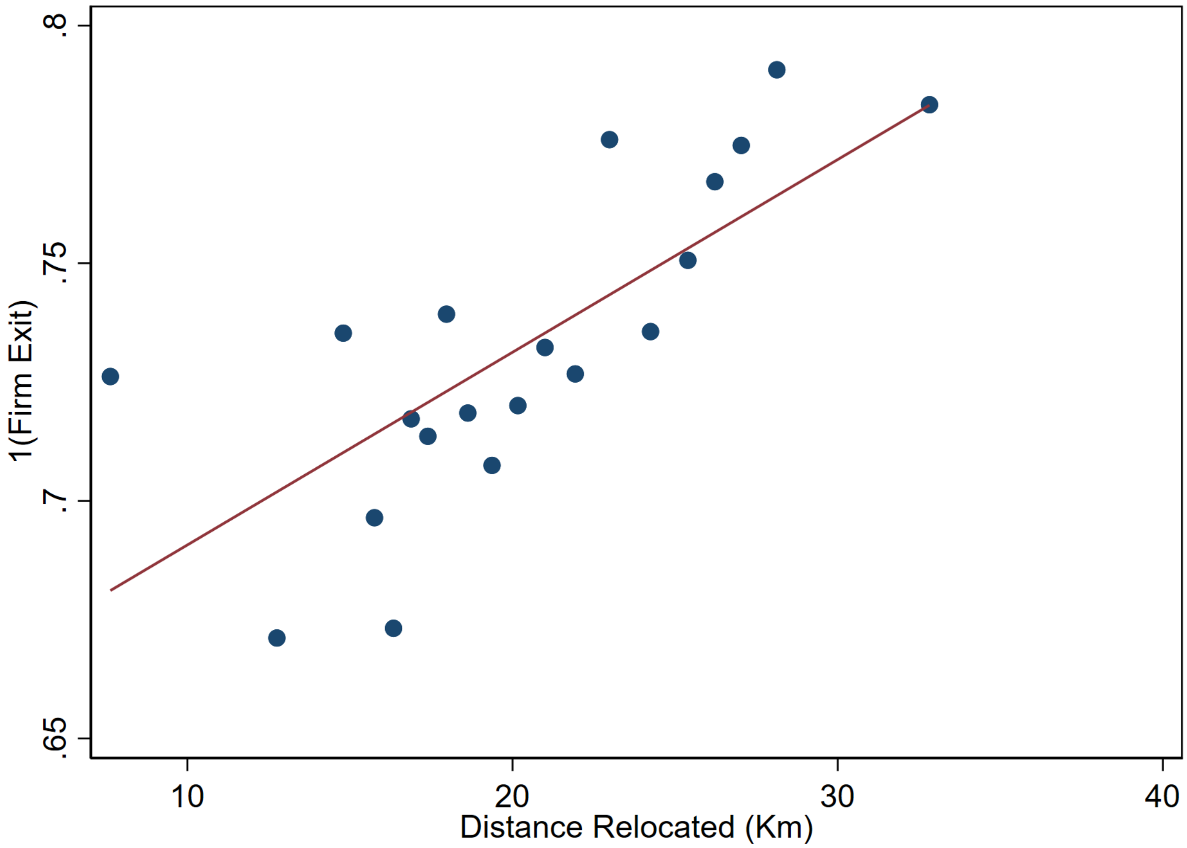


Table 1: Firm-Level Lottery Timing and Timing of Move to Industrial Area

	(1) Year of Physical Possession	(2) Year of Lease Execution	(3) First Due Year Of Lease Ground Rent
Year(Draw Date)	1.128****	0.725****	1.128****
	(0.0112)	(0.0105)	(0.0112)
Observations	17654	16250	17654

Robust standard errors in parentheses.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$, **** $p < 0.001$

Table 2: Impact of Neighborhood-Level Lottery Timing on Timing of Firm Leasing in the Industrial Areas

	(1) Log(Firms per Sq Km Leased by 2010)	(2) Log(Firms per Sq Km Leased by 2015)	(3) Proportion Firms with Lease by 2010 (Std)	(4) Proportion Firms with Lease by 20015 (Std)
Log (Number of Firms Relocated per Sq Km)	0.251**	0.448****	-0.0572	-0.0669
	(0.0970)	(0.0744)	(0.109)	(0.136)
Log(Firms per Sq Km Relocated by 2003)	0.750****	0.552****		
	(0.0955)	(0.0742)		
Log(Area-Sq Km)	0.000995	0.00204	-0.0779	-0.0811
	(0.0234)	(0.0261)	(0.165)	(0.198)
Proportion Firms Relocated by 2003 (Std)			0.617****	0.542****
			(0.0806)	(0.0731)
Constant	-0.118	-0.0632	2.245**	4.861****
	(0.0947)	(0.0974)	(0.962)	(1.107)
Mean of the Dependent Variable (Levels)	41.04	46.65	8.513	10.30
N	134	134	134	134
R2	0.997	0.997	0.385	0.302

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$, **** $p < 0.001$

Restricted to lottery years 2005 and earlier, since over 90% of firms received a plot by 2005.

Table 3: Impacts on Aerosol Optical Depth

	(1) Mean	(2) Minimum	(3) Maximum
Log(Firms Relocated Per Sq Km by 2003) × $\mathbb{1}$(Year>2005)	0.0248	0.00580	0.141
	(0.0192)	(0.0104)	(0.148)
Log(Total Firms Relocated Per Sq Km) × $\mathbb{1}$(Year>2005)	-0.0222	-0.00742	-0.0941
	(0.0191)	(0.0103)	(0.148)
Constant	0.619****	0.0753****	2.796****
	(0.00457)	(0.00252)	(0.0370)
Dependent Variable Mean	0.620	0.0700	2.850
Observations	2144	2144	2144

Standard errors clustered at the 2001 Census neighborhood level in parentheses. Neighborhood and year fixed effects included.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$, **** $p < 0.001$

Table 4: Impacts on Fine PM

	(1) Mean	(2) Minimum	(3) Maximum
Log(Firms Relocated Per Sq Km by 2003) × $\mathbb{1}$(Year>2005)	0.317	0.231	0.408
	(0.685)	(0.726)	(0.708)
Log(Total Firms Relocated Per Sq Km) × $\mathbb{1}$(Year>2005)	-0.216	-0.169	-0.275
	(0.686)	(0.728)	(0.711)
Constant	106.8****	106.3****	107.4****
	(0.126)	(0.135)	(0.133)
Mean of the Dependent Variable	106.9	106.3	107.5
Observations	1474	1474	1474

Standard errors clustered at the 2001 Census neighborhood level in parentheses. Neighborhood and year fixed effects included.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$, **** $p < 0.001$

Table 5: Population Census: Balance

	(1) Log(2001 Population Dens)	(2) 2001 Proportion Workers
Log (Number of Firms Relocated per Sq Km)	1.769	0.0000492
	(1.640)	(0.0000620)
Log(Firms per Sq Km Relocated by 2003)	-1.659	-0.0000523
	(1.573)	(0.0000584)
Log(Area-Sq Km)	-0.855****	-0.00000857
	(0.127)	(0.00000908)
Constant	10.10****	0.000127***
	(0.968)	(0.0000471)
Mean of the Dependent Variable	35046.6	0
N	134	132
R2	0.567	0.0356

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$, **** $p < 0.001$

Table 6: Population Census: Impacts

	(1) Log(2011 Population Density)	(2) 2011 Proportion Workers
Log(2001 Population Density)	0.114	
	(0.105)	
Log (Number of Firms Relocated per Sq Km)	-0.381	0.0330
	(0.450)	(0.0290)
Log(Firms per Sq Km Relocated by 2003)	0.525	-0.0349
	(0.440)	(0.0285)
Log(Area-Sq Km)	-0.307**	-0.00362
	(0.120)	(0.00410)
2001 Proportion Workers		179.6**
		(84.85)
Constant	9.077****	0.318****
	(1.207)	(0.0233)
Mean of the Dependent Variable	28989.1	0.343
N	134	132
R2	0.730	0.110

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$, **** $p < 0.001$

Table 7: Sixth Census: Impacts

	(1)	(2)	(3)	(4)
	Log (Firms Per Sq Km)			
Log (Number of Firms Relocated per Sq Km)	-0.340	-0.119	-0.150	0.858
	(0.641)	(0.737)	(0.758)	(0.841)
Log(Firms per Sq Km Relocated by 2003)	0.598	0.512	0.547	-0.234
	(0.619)	(0.714)	(0.734)	(0.836)
Log(Area-Sq Km)	-0.403****	-0.309****	-0.309***	-0.0879
	(0.0748)	(0.0909)	(0.0937)	(0.0959)
Constant	7.080****	4.686****	4.587****	1.628****
	(0.415)	(0.490)	(0.504)	(0.477)
Sector Restriction on Number of Relocated Firms	None	50	100	500
Mean of the Dependent Variable (Levels)	1817.3	349.0	317.5	74.75
N	134	134	134	134
R2	0.721	0.676	0.669	0.700

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$, **** $p < 0.001$

Table 8: Directory of Establishments from Sixth Census: Impacts

	(1)	(2)	(3)	(4)
	Log (Firms Per Sq Km)			
Log (Number of Firms Relocated per Sq Km)	0.814	0.487	0.495	0.486
	(1.007)	(1.358)	(1.356)	(1.279)
Log(Firms per Sq Km Relocated by 2003)	-0.577	-0.104	-0.121	-0.170
	(0.990)	(1.330)	(1.328)	(1.262)
Log(Area-Sq Km)	-0.247*	0.0546	0.0473	0.156
	(0.131)	(0.182)	(0.182)	(0.148)
Constant	2.637****	0.975	0.956	0.177
	(0.631)	(0.882)	(0.882)	(0.760)
Sector Restriction on Number of Relocated Firms	None	50	100	500
Mean of the Dependent Variable (Levels)	54.32	33.21	32.36	14.81
N	133	133	133	133
R2	0.315	0.185	0.178	0.104

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$, **** $p < 0.001$

Table 9: Impacts of Distance Relocated on Firm Exit

	(1)	(2)
	$\mathbb{1}$(Firm Exit)	$\mathbb{1}$(Firm Exit)
Distance Relocated (km)	0.00406***	0.00939***
	(0.000589)	(0.00339)
Constant	0.650***	0.540***
	(0.0125)	(0.0694)
Neighborhood Fixed Effecsts	No	Yes
Mean of the Dependent Variable	0.733	0.732
N	16343	16173
R2	0.00291	0.0151

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

7 Appendix

Figure A1: Bawana Industrial Area-2000

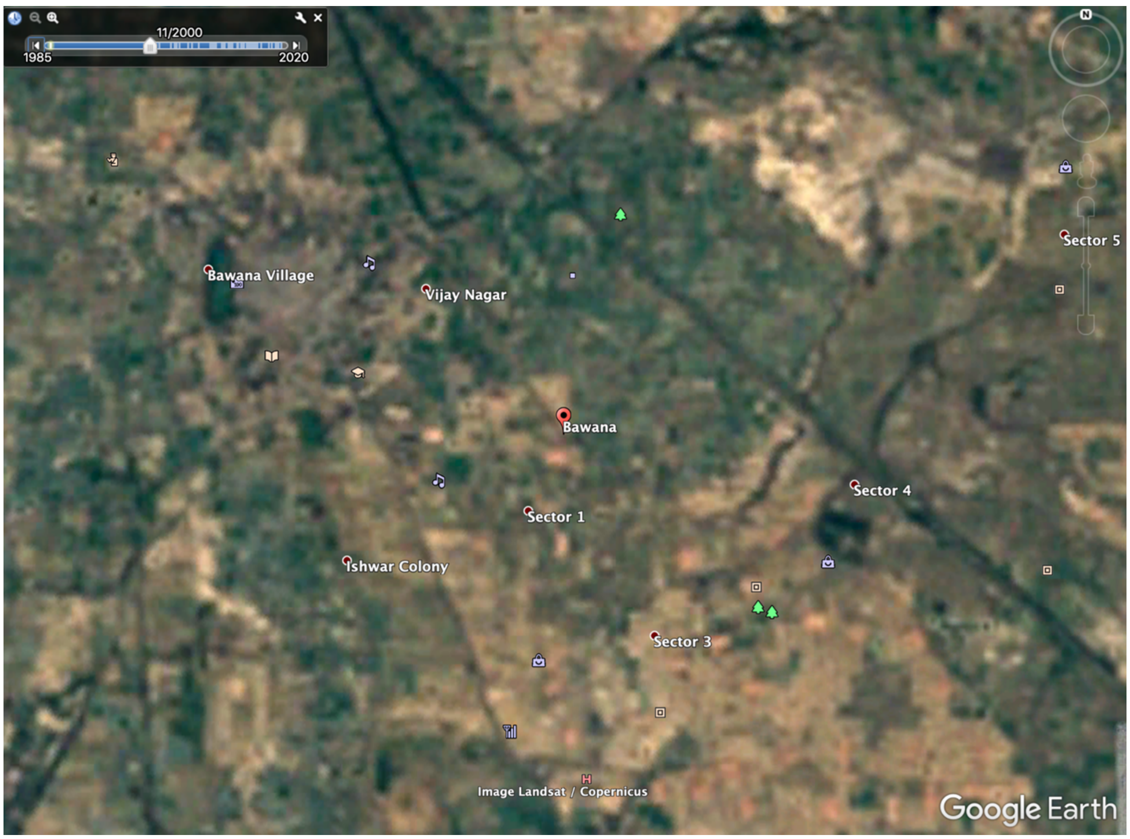


Figure A2: Bawana Industrial Area-2001

Figure A3: Bawana Industrial Area-2010

Table A1: Directory of Establishments from Sixth Census: Impacts on Firms with Fewer than 30 Employees

	(1)	(2)	(3)	(4)
	Log (Firms Per Sq Km)			
Log (Number of Firms Relocated per Sq Km)	1.055	0.390	0.374	0.549
	(0.899)	(1.225)	(1.222)	(1.082)
Log(Firms per Sq Km Relocated by 2003)	-0.799	-0.00806	0.000346	-0.248
	(0.886)	(1.200)	(1.197)	(1.067)
Log(Area-Sq Km)	-0.231*	0.0690	0.0653	0.164
	(0.120)	(0.167)	(0.167)	(0.122)
Constant	2.161****	0.720	0.707	-0.0651
	(0.561)	(0.793)	(0.792)	(0.630)
Sector Restriction on Number of Relocated Firms	None	50	100	500
Mean of the Dependent Variable (Levels)	38.08	22.59	22.02	8.229
N	133	133	133	133
R2	0.356	0.199	0.191	0.115

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$, **** $p < 0.001$

Table A2: Directory of Establishments from Sixth Census: Impacts on Firms With Between 30 to 99 Employees

	(1)	(2)	(3)	(4)
	Log (Firms Per Sq Km)			
Log (Number of Firms Relocated per Sq Km)	0.354	-0.0488	-0.0571	-0.286
	(1.069)	(1.179)	(1.173)	(1.048)
Log(Firms per Sq Km Relocated by 2003)	-0.227	0.211	0.214	0.411
	(1.048)	(1.164)	(1.158)	(1.034)
Log(Area-Sq Km)	-0.151	0.0313	0.0223	0.0598
	(0.145)	(0.145)	(0.145)	(0.132)
Constant	1.539**	0.632	0.637	0.391
	(0.714)	(0.737)	(0.736)	(0.666)
Sector Restriction on Number of Relocated Firms	None	50	100	500
Mean of the Dependent Variable (Levels)	13.97	9.540	9.343	6.022
N	133	133	133	133
R2	0.120	0.0460	0.0457	0.0263

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$, **** $p < 0.001$

Table A3: Directory of Establishments from Sixth Census: Impacts on Firms with 100 Employees or More

	(1)	(2)	(3)	(4)
	Log (Firms Per Sq Km)			
Log (Number of Firms Relocated per Sq Km)	-0.645	-0.333	-0.276	-0.180
	(0.674)	(0.494)	(0.476)	(0.377)
Log(Firms per Sq Km Relocated by 2003)	0.713	0.442	0.384	0.262
	(0.674)	(0.508)	(0.489)	(0.385)
Log(Area-Sq Km)	-0.0234	0.0839	0.0797	0.0696
	(0.0896)	(0.0724)	(0.0718)	(0.0649)
Constant	0.846**	0.0712	0.0334	-0.0427
	(0.414)	(0.291)	(0.284)	(0.247)
Sector Restriction on Number of Relocated Firms	None	50	100	500
Mean of the Dependent Variable (Levels)	2.276	1.074	0.994	0.561
N	133	133	133	133
R2	0.0370	0.0433	0.0459	0.0397

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$, **** $p < 0.001$

Table A4: Category of 2011 Real Estate Prices: Impacts

	(1)	(2)	(3)
	Mean	Minimum	Maximum
Log (Number of Firms Relocated per Sq Km)	0.0542	-0.687	0.815
	(1.045)	(0.879)	(1.728)
Log(Firms per Sq Km Relocated by 2003)	-0.201	0.618	-1.046
	(1.017)	(0.819)	(1.730)
Log(Area-Sq Km)	-0.284*	-0.492****	-0.0243
	(0.150)	(0.132)	(0.220)
Constant	4.187****	3.130****	5.246****
	(0.737)	(0.733)	(0.975)
Mean of the Dependent Variable (Levels)	3.421	1.985	4.925
N	133	133	133
R2	0.0538	0.294	0.0614

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$, **** $p < 0.001$

Table A5: Correlation Between Missing or Noisy Geocoding with Timing of Firm Lottery

	(1)	(2)	(3)
	$\mathbb{1}$[Surveyor Geocode and Google Gecode in Different Wards]	$\mathbb{1}$[Surveyor Geocode Missing]	$\mathbb{1}$[Google Geocode Missing]
Lottery Year	-0.000619	-0.00125	-0.000170
	(0.00193)	(0.00115)	(0.000930)
Constant	1.574	2.601	0.421
	(3.858)	(2.303)	(1.862)
Mean of the Dependent Variable	0.334	0.0980	0.0800
N	14061	15010	20186
R2	0.00000728	0.0000749	0.00000167

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Restricted to lottery years 2005 and earlier, since over 90% of firms received a plot by 2005.

Table A6: Impact of Neighborhood-Level Lottery Timing on Timing of Firm Leasing in the Industrial Areas: 2km by 2km Grid Level

	(1)	(2)	(3)	(4)
	Log(Firms Leased by 2010)	Log(Firms Leased by 2015)	Proportion Firms with Lease by 2010 (Std)	Proportion Firms with Lease by 20015 (Std)
Log (Number of Firms Relocated)	0.416****	0.713****	-0.0538**	-0.0411
	(0.0729)	(0.0594)	(0.0243)	(0.0297)
Log(Firms Relocated by 2003)	0.565****	0.271****		
	(0.0749)	(0.0609)		
Proportion Firms Relocated by 2003 (Std)			0.699****	0.551****
			(0.0851)	(0.108)
Constant	-0.109***	-0.0971***	1.266****	2.509****
	(0.0338)	(0.0324)	(0.319)	(0.411)
Mean of the Dependent Variable (Levels)	58.70	66.50	3.165	4.010
N	223	224	237	237
R2	0.987	0.989	0.495	0.307

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$, **** $p < 0.001$

Table A7: Directory of Establishments from Sixth Census: Impacts at the 2km by 2km Grid Level

	(1)	(2)	(3)	(4)
		Log (Firms)		
Log (Number of Firms Relocated)	0.809**	0.584	0.545	0.459
	(0.342)	(0.367)	(0.375)	(0.335)
Log(Firms Relocated by 2003)	-0.236	0.0374	0.0852	0.0576
	(0.331)	(0.352)	(0.360)	(0.320)
Constant	1.873****	0.721***	0.636**	0.207
	(0.237)	(0.267)	(0.270)	(0.246)
Sector Restriction on Number of Relocated Firms	None	50	100	500
Mean of the Dependent Variable (Levels)	118.9	67.61	64.92	33.27
N	213	213	213	213
R2	0.423	0.379	0.382	0.305

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$, **** $p < 0.001$

Restricted to lottery years 2005 and earlier, since over 90% of firms received a plot by 2005.

Table A8: Directory of Establishments from Sixth Census: Impacts on Firms with Fewer than 30 Employees at the 2km by 2km Grid Level

	(1)	(2)	(3)	(4)
		Log (Firms)		
Log (Number of Firms Relocated)	0.793**	0.503	0.427	0.305
	(0.324)	(0.346)	(0.350)	(0.303)
Log(Firms Relocated by 2003)	-0.220	0.0837	0.166	0.163
	(0.313)	(0.331)	(0.335)	(0.287)
Constant	1.543****	0.550**	0.494*	0.121
	(0.226)	(0.252)	(0.252)	(0.227)
Sector Restriction on Number of Relocated Firms	None	50	100	500
Mean of the Dependent Variable (Levels)	78.02	42.90	41.21	18.52
N	213	213	213	213
R2	0.454	0.380	0.384	0.308

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$, **** $p < 0.001$

Table A9: Directory of Establishments from Sixth Census: Impacts on Firms Between 30 to 99 Employees at the 2km by 2km Grid Level

	(1)	(2)	(3)	(4)
		Log (Firms)		
Log (Number of Firms Relocated)	0.551*	0.310	0.313	0.345
	(0.322)	(0.303)	(0.303)	(0.263)
Log(Firms Relocated by 2003)	-0.0876	0.106	0.0982	-0.0508
	(0.315)	(0.292)	(0.293)	(0.254)
Constant	0.870****	0.204	0.173	0.0348
	(0.220)	(0.224)	(0.223)	(0.201)
Sector Restriction on Number of Relocated Firms	None	50	100	500
Mean of the Dependent Variable (Levels)	33.84	21.70	21.00	13.21
N	213	213	213	213
R2	0.302	0.232	0.227	0.148

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$, **** $p < 0.001$

Table A10: Directory of Establishments from Sixth Census: Impacts on Firms with 100 Employees or More at the 2km by 2km Grid Level

	(1)	(2)	(3)	(4)
		Log (Firms)		
Log (Number of Firms Relocated)	0.550**	0.134	0.111	0.0955
	(0.268)	(0.182)	(0.170)	(0.132)
Log(Firms Relocated by 2003)	-0.291	0.0314	0.0506	0.0138
	(0.260)	(0.172)	(0.160)	(0.125)
Constant	0.241	0.0682	0.0443	0.0162
	(0.179)	(0.142)	(0.137)	(0.108)
Sector Restriction on Number of Relocated Firms	None	50	100	500
Mean of the Dependent Variable (Levels)	7.033	3.014	2.718	1.549
N	213	213	213	213
R2	0.167	0.0972	0.0997	0.0718

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$, **** $p < 0.001$

Table A11: Impacts on Aerosol Optical Depth: 2km by 2km Grid Level

	(1) Mean	(2) Minimum	(3) Maximum
Log(Firms Relocated Per Sq Km by 2003) × $\mathbb{1}$(Year>2005)	-0.00129	-0.00554	-0.0485
	(0.00675)	(0.00389)	(0.0534)
Log(Total Firms Relocated Per Sq Km) × $\mathbb{1}$(Year>2005)	0.00277	0.00436	0.0799
	(0.00690)	(0.00402)	(0.0544)
Constant	0.612****	0.0643****	2.829****
	(0.00266)	(0.00151)	(0.0190)
Dependent Variable Mean	0.620	0.0600	2.910
Observations	2704	2704	2704

Standard errors clustered at the 2km by 2km grid cell in parentheses. Grid cell and year fixed effects included.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$, **** $p < 0.001$

Table A12: Impacts on Fine PM: 2km by 2km Grid Level

	(1) Mean	(2) Minimum	(3) Maximum
Log(Firms Relocated Per Sq Km by 2003) × $\mathbb{1}$(Year>2005)	-0.419	0.0398	-1.419
	(0.558)	(0.544)	(2.521)
Log(Total Firms Relocated Per Sq Km) × $\mathbb{1}$(Year>2005)	0.334	-0.0356	1.586
	(0.555)	(0.572)	(2.487)
Constant	106.7****	40.16****	219.2****
	(0.159)	(0.203)	(0.751)
Dependent Variable Mean	106.6	40.16	219.8
Observations	3006	3006	3006

Standard errors clustered at the 2km by 2km grid cell in parentheses. Grid cell and year fixed effects included.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$, **** $p < 0.001$