Productivity impacts of pollution control when people and particulates are mobile

Anshuman Tiwari*

12 June, 2025

Abstract

How do productivity gains from pollution control change when accounting for migration? Productivity gains depend on whether regulation improves air quality and health in productive cities, and on how strongly resulting amenity and wage improvements attract workers to those cities. Because emissions sources differ in pollutant dispersion, quantifying benefits requires explicit dispersion modelling. I incorporate dispersion into a spatial equilibrium model and compare gains from scenarios with similar health benefits targeting localized urban emissions versus rural crop-burning in India, both major contributors to urban pollution. Migration-driven productivity gains are 18 times larger under urban regulation, driving 6-fold greater overall benefits.

* Energy Policy Institute at Chicago - India (email: atiwari2@uchicago.edu). I thank Gharad Bryan, Robin Burgess, Tamma Carleton, Kyle Meng, Gabriel Ahlfeldt, Kathy Bayliss, Tim Besley, Prashant Bhardwaj, Chris Costello, Olivier Deschenes, Marion Dumas, Teevrat Garg, Ludovica Gazze, Ben Groom, Eyal Frank, Nick Hagerty, Vernon Henderson, Allan Hsiao, Koichiro Ito, Kelsey Jack, Matt Lowe, Pierre Merel, Sefi Roth, Lutz Sager, Misato Sato, Daniel Sturm, Chris Timmins and various seminar audiences at LSE, UCSB, UCSD, UPenn PDRI, UCSB Occasional Workshop, PacDev, NEUDC, UC-EE Online, Ashoka university and ISI-Delhi for helpful discussions and suggestions. All errors are mine. Financial support from the Environmental Defense Fund, the Grantham Research Foundation and LSE STICERD is gratefully acknowledged.

1 Introduction

Standard benefit–cost analyses of air pollution regulation typically monetize only health improvements (US EPA 2015; Currie and Walker 2019). However, pollution also affects productivity directly by impairing worker health (Graff Zivin and Neidell 2012; Chang et al. 2016, 2019) and indirectly by influencing where workers choose to live (Chay and Greenstone 2005; Banzhaf and Walsh 2008; Freeman et al. 2017; Heblich et al. 2021; Chen et al. 2022). These productivity impacts could be especially large if pollution reduces wages and amenities in productive cities, discouraging worker migration to these urban centers (Lewis. 1954; Gollin 2014; Khanna et al. 2025). Whether pollution regulation enhances urban air quality and attracts workers depends crucially on the geographic dispersion of emissions. Regulating urban emissions clearly improves air quality, worker health, amenities and wages in those cities, thus attracting workers toward productive locations. Yet, controlling distant rural emissions can also significantly improve urban air quality downwind, similarly attracting workers to those cities. Thus, accurate estimation of productivity benefits requires explicit modeling of the interaction between pollution dispersion and worker migration. I study how the productivity benefits of pollution control depend on the interaction of dispersion and migration by incorporating pollution dispersion into a spatial equilibrium model of worker location choice.

Existing literature documents the health effects of pollution spillovers across jurisdictions, including both within and across countries (Deryugina et al. 2019; Heo et al. 2025). These studies typically focus only on health outcomes. However, a given health improvement in urban areas could yield larger productivity benefits compared to rural, due to the higher marginal product of labor in cities. Therefore, productivity gains from emissions regulation depend on the marginal product of labor in all the locations that are cleaned up. Moreover, these are partial equilibrium effects only. Pollution could also drive spatial sorting of workers, and this can have large productivity effects if workers reallocate toward high marginal product cities in general equilibrium (Khanna et al. 2025). Importantly, pollution regulation alters migration incentives *across all locations*, not solely in areas with directly improved air quality. Reallocation of labor across space alters prices, and induces congestion and agglomeration effects. For these reasons, standard causal inference tools cannot estimate the productivity effects of pollution regulation in general equilibrium.

This paper combines novel empirical evidence on pollution-driven migration in India with a spatial equilibrium model to understand productivity consequences of spatially targeted pollution control. I build a worker location choice model that can accommodate migration of workers and dispersion of pollution across fine-grained geographic units, while being tractable enough to conduct policy analysis. Specifically, I embed pollution dispersion across geographic units into a spatial equilibrium framework featuring agents with heterogeneous preferences for locations, and realistic migration costs that restrict mobility (Redding and Rossi-Hansberg 2017). Pollution can affect worker location choice for two reasons. First, the worker productivity effects of pollution may capitalize into wages if the aggregate marginal product of labor is reduced (Hanna and Oliva 2015;

Borgschulte et al. 2022; Hoffmann and Rud 2024). All else equal, a greater pollution differential between two locations would therefore increase the income differential, lowering the incentive to migrate to the more polluted location (Lewis. 1954; Harris and Todaro 1970). The income elasticity governs how worker migration responds to income differences across locations: a lower elasticity implies that higher incomes need to be paid on average to induce marginal workers to migrate. Second, pollution may lower quality of life and thus have a direct amenity value (Roback 1982). The amenity elasticity governs whether workers migrate away from pollution due to a preference for clean air. The model also accounts for other mechanisms by which pollution control could affect aggregate productivity: (1) labor reallocation toward cities can affect agglomeration economies (Au and Henderson 2006) and, (2) labor reallocation could also have congestion effects, for example through increased pollution or housing costs (Bayer et al. 2009).

I apply this framework to study productivity gains from targeted emissions control scenarios in India, which provides the ideal context to study this question for two reasons. First, it is the world's most populous and polluted country. Almost all of India's population in 2011 experienced pollution levels substantially higher than the World Health Organization (WHO) limit for particulate matter below 2.5 microns in size.[1] It is important to study the effects of air pollution in India due to the immense human cost involved (Greenstone 2022). Secondly, India is also witnessing rapid economic growth, with a concomitant migration of economic activity and labor toward urban areas. Productivity gains are important for developing countries since they increases their ability to finance investments for continued growth. Therefore, understanding whether pollution in urban centers in India is causing productivity losses due to direct health effects and indirect spatial misallocation of labor is important from a policy perspective. Whether pollution in urban centers is caused by local sources such as vehicular emissions or upwind, rural sources such as crop residue burning is a hotly contested policy debate (Guttikunda et al. 2023). This motivates the specific emissions regulation scenarios that I describe on the next page.

Before I delve into the productivity consequences of pollution regulation in India, I provide the first plausibly causal evidence that pollution affects worker location choice in India. I leverage novel pairwise migration data across Indian districts for two 5-year intervals preceding the 2011 population census.[2] In the decade preceding 2011, PM2.5 increased by an average of 10% across India, with substantial variation across districts. I test whether rising pollution in this period affects where workers choose to locate by relating changes in district out-migration and in-migration to changes in PM2.5 levels. Since changes in pollution, economic activity, and migration are jointly determined, I employ an instrumental variable strategy to isolate exogenous changes in pollution from upwind crop residue burning, which are plausibly unrelated to local economic activity in the

[1] This form of pollution is commonly knows as PM2.5 and is known to cause serious health effects. Both short-term and prolonged exposure to PM2.5 can lead to heart attacks, asthma, decreased lung function or cancer, stroke and a variety of other conditions, and cause premature mortality in people with heart or lung diseases (Greenstone 2022). The WHO annual average limit is 5 μg/m3. According to the WHO's PM2.5 database, Delhi's annual average PM2.5 level was 153 μg/m3 in 2013, in comparison with New York city's average of 14 and London's at 16.

[2] These are disaggregated by age, education, gender and the reported reason for migration.

downwind districts.[3] Using this IV strategy, I show that a 1% increase in PM2.5 levels is associated with a reduction in migrant worker inflows of 2.12% and an increase in outflows of 0.86%. This finding reinforces the necessity of using a general equilibrium model of worker location choice to understand productivity gains from regulating emissions to account for the resulting spatial reallocation of workers.

As discussed earlier, this migration response to pollution may be explained by either the income channel or the amenity channel. I estimate these income and amenity elasticities together in a gravity framework implied by the quantitative model, leveraging data on worker migration across the district pairs along with data on wage and pollution levels. I employ an instrumental variables strategy to deal with endogeneity concerns about unobserved, residual factors that affect migration and are correlated with wage or pollution.[4] Specifically, I construct an origin-destination specific instrument for destination wages that weights every out-of-state district's wage by the 2001 migration shares of that district to the destination. For pollution, I exploit the plausibly exogenous variation arising from exposure to upwind agricultural fires. Using this IV strategy, I find an income elasticity of 4.36 and an amenity elasticity of -0.25, both of which are statistically different from zero. Prior work on developing countries hypothesizes that this second channel is less important given the low willingness to pay to avoid air pollution damages as a consequence of low income levels (Greenstone et al. 2021; Greenstone and Jack 2015). On the contrary, evidence from India suggests that workers are willing to forego up to 35% of income to stay away from harsh conditions in the city such as low quality housing or substandard transportation (Imbert and Papp 2020b). Such living conditions may also increase the exposure of these workers to high pollution levels so that it becomes a small but salient disamenity. In the following sections, I also outline estimation strategies that rely on the structure of the model and sources of plausibly exogenous variation to identify other important parameters.

I study two emissions control scenarios that are motivated by the wider policy debate on pollution in India (Guttikunda et al. 2023). The first scenario imposes a 10% reduction of emissions from agricultural fires originating in north India (hereafter referred to as the *Rural* scenario).[5] To understand which locations would benefit in this scenario, I utilize historical data on wind patterns, crop burning events and PM2.5 levels over 16 years to estimate a pollution dispersion model for smoke. This model predicts annual smoke-driven pollution with a ~80% accuracy against a state-of-the-art Chemical Transport Model (CTM), with residue burning responsible for up to 15% of annual pollution in north Indian cities (McDuffie et al. 2021; Singh et al. 2021). A CTM models both the Physics and Chemistry of pollutant dispersion from an emissions source, and can take days to complete a single run for a single source. This makes it infeasible to use for a counterfactual scenario that

[3] Crop burning increased by an average of 370% across Indian districts within this period. These increases are attributable to many factors including exogenous policy changes such as the National Rural Employment Guarantee Act (Behrer 2023) and groundwater conservation policies in the states of Punjab and Haryana (Tiwari 2025).

[4] For example, the quality of housing stock or pre-existing origin-destination migrant networks.

[5] The two states of Punjab and Haryana in northwestern India accounted for ~56% of all burning events in the country in 2011. Jack et al. (2025) conduct an RCT in Punjab to show that a payment of between INR 2700-4050 per acre to farmers incentivizes them to reduce burning by 10%.

require multiple runs for each source for one scenario. The pollution dispersion model can finish multiple runs for multiple scenarios *in under one second*. The second scenario imposes restrictions on localized emissions such as those from vehicles within the 10 largest cities of India (the *Urban* scenario). To show that such emissions affect pollution largely within the city itself, I combine hourly data from geocoded monitoring stations on pollution and wind with distance from nearest major roads for the largest Indian cities, and find that pollution decays to background levels within 5-10 kilometers of the road.[6] Other localized sources include residential solid fuel burning for cooking and heating (McDuffie et al. 2021). I hold the total reduction in population exposure constant between the *Rural* and *Urban* scenarios to ensure that well-established health benefit calculations score both scenarios as nearly identical.[7]

Next, I utilize the quantitative model to understand productivity gains from these scenarios. Accounting only for differences in where regulation improves air quality, the *Urban* scenario increases GDP by three times more than the *Rural* scenario.[8] Allowing for a worker migration response, the *Urban* scenario increases GDP by six times more than the *Rural* scenario. In absolute terms, aggregate gains without migration in the *Rural* scenario are 0.09% of baseline Indian GDP, and increase to 0.11% of GDP with migration.[9] In contrast, under the *Urban* scenario, aggregate gains without migration are 0.25% of GDP, increasing to 0.68% of national GDP when allowing for migration.[10] In absolute terms, this is a difference of 0.68% - 0.25% = 0.43% between gains from the two scenarios; in comparison, the Government of India budgeted ~0.38% of GDP for the National Rural Employment Guarantee Scheme for poverty alleviation in 2012.

Gains from migration from the *Urban* scenario are 18 times larger compared to the *Rural* scenario.[11] What explains this large difference under the two scenarios even though they would score similarly on health benefits? The key insight is that these gains are larger when air quality improvements reallocate workers to productive cities rather than rural areas. Residue burning control improves air quality along the entire pollution dispersion path that includes some cities, but it improves air quality relatively more in rural areas. As a result, migration toward urban centers is muted, and productivity gains from migration are smaller. On the other hand, regulating localized emissions

[6] Background levels in these cities are usually around 40-60% of the level at the road. Estimated rates in the scientific literature suggest much faster decay, within 1 kilometer (Liu et al. 2019).

[7] The precise calculations depend on the shape of the dose-response function. The EPA uses a linear function that would produce identical benefits from these two scenarios, whereas a concave dose-response would score the *Rural* scenario higher due to larger marginal health benefits for rural residents who are exposed to lower baseline PM2.5 levels (US EPA 2015)

[8] Since total population is unchanged, percentage changes in GDP equal changes in GDP per capita, the our main measure of aggregate productivity.

[9] To give a sense for the magnitudes involved, the Government of India (GoI) in 2012 budgeted ~0.02% of GDP for grants to extremely poor households under the Homestead scheme of the Indira Awaas Yojana, ~0.06% of GDP to provide lower interest loans to farmers, and ~0.1% for the Rashtriya Krishi Vikas Yojana to enhance public investment in agriculture (Government of India, Ministry of Finance 2013a, 2013b).

[10] For comparison, the 2012 budget of the GoI provided ~0.2% of GDP for the Mid-Day meal scheme that provided school-age children with free meals to improve nutrition and ~0.9% to the annual fertilizer subsidy for all farmers (Government of India, Ministry of Finance 2013b).

[11] $\frac{(0.681-0.254)}{(0.112-0.089)}$ ~18.

in the largest cities improves air quality solely within those cities. This sharply increases incentives to migrate to those cities, reinforcing agglomeration economies in these already productive cities and leading to higher productivity gains from improved spatial allocation of labor.

I also find that if the residue burning scenario were to be become policy, it would more than pay for itself: the benefit-cost ratio for GDP gains with labor reallocation from a 10% control of residue burning is 31:1. I use the upper-end estimate of Indian Rupees 4050 to reduce burning per-acre from Jack et al. (2025) to do this calculation.[12] While similar abatement cost estimates do not exist for localized emissions reductions, these costs may be reasonably low given that pollution control programs in developing countries may be far from efficient (Greenstone and Jack 2015). Given that the aggregate benefits for the localized emissions scenario are also much higher at 0.68% of GDP, control policies are likely to have a high return on investment.[13]

This paper makes several contributions. Firstly, I contribute to the air pollution literature that documents well-identified, partial equilibrium effects of pollution on productivity (Graff Zivin and Neidell 2012; Chang et al. 2016, 2019; Fu et al. 2021; Borgschulte et al. 2022; Hoffmann and Rud 2024) but has fewer papers on the impact of adjustments in general equilibrium. I demonstrate that pollution reduces aggregate productivity by causing spatial misallocation of labor, and that the extent of this misallocation depends on the geography of the emissions source (its location and dispersion tendency). Incorporating this geographic element into the analysis is a vital difference with the spatial equilibrium framework in Khanna et al. (2025), who quantify aggregate productivity losses from workers' preference for cleaner air in China. For example, not accounting for migration would miss gains of ~0.43% of GDP from controlling localized emissions within cities, but only ~0.023% of GDP from reducing residue burning, a large 18x difference in productivity gains between health-similar scenarios. Understanding where pollution comes from is critical to quantifying productivity effects of pollution control when accounting for migration in general equilibrium. Localized sources within cities create larger productivity wedges by directing workers away from productive cities, and the size of those migration-driven losses is far larger than what we would anticipate based on partial equilibrium, health-based analysis.

Secondly, I provide evidence for a migration response to pollution in India. Currently evidence documents such responses in richer economies (Chay and Greenstone 2005; Banzhaf and Walsh 2008; Heblich et al. 2021), or in China where the government has widely publicized information on air quality (Freeman et al. 2017; Chen et al. 2022; Khanna et al. 2025). India had a GDP per-person less than half of China's in 2011. Seen in this light, the finding of a sorting response to

[12] Jack et al. (2025) estimate a cost of INR 2700-4050 per-acre to reduce burning. Using the upper estimate, the total cost of reducing burning on 10% of cultivated rice area as a percentage of 2011 GDP is 0.0036%. Aggregate gains are 0.11% of GDP, giving a ratio of 0.11/0.0036=~31. Detailed calculations are provided in section 8. This calculation assumes that emissions are linear in cultivated area; I provide supporting evidence in the appendix. A caveat for these GDP benefits come from a model that is estimated on data from 2011.

[13] For example, promising ways to reduce vehicular emissions include scrapping the oldest and most polluting vehicles, improving testing under the mandatory Pollution Under Control (PUC) program or implementing congestion pricing. Adoption of cleaner burning cook stoves has been shown to cost-effectively reduce pollution from residential solid fuel burning (Berkouwer and Dean 2022).

pollution in India may seem unexpected; if clean air is a normal good demand for it should grow with income. But this migration response is driven largely by the effect of pollution on income, which is relatively more important than the amenity channel in explaining the migration effect. The small amenity channel echoes findings from Imbert and Papp (2020b) who show that workers are willing to forego up to 35% increases in income to avoid urban disamenities in Indian cities. For most workers, it may be difficult to avoid pollution-related damages through the use of low-cost technologies like masks, while more effective (yet imperfect) technologies such as air purifiers are too expensive for all but the richest households. My results show that some workers respond to extremely high ambient pollution concentrations by avoiding these cities.

Finally, I contribute to the macro-development literature that views migration toward urban areas as being synonymous with increases in aggregate productivity (Lewis. 1954), and documents various frictions that slow this process (Gollin 2014). I identify urban air pollution as a previously overlooked barrier (Khanna et al. 2025). Importantly, I document that localized urban pollution sources cause much larger spatial misallocation of labor by deterring migration and trapping workers in less-productive areas. By quantifying the aggregate output loss from this environment-induced misallocation, this paper shows that cleaning urban pollution is not only a health intervention but also a productivity-enhancing structural transformation policy.

The rest of the paper is structured as follows. Section 2 builds intuition through a two-location example, section 3 describes the Indian context and the pollution control scenarios, section 4 describes the data used in the paper, section 5 provides empirical results on a migration response to pollution in India. Section 6 presents the model while section 7 describes the estimation of the parameters governing equilibrium. Section 8 describes the results from model counterfactuals and section 9 concludes.

2 Two-location example

In order to fix ideas about how pollution regulation might affect productivity, let us consider a two-location example. In particular, we are interested in understanding the interaction between pollution dispersion and worker migration. There are two locations Urban and Rural, indexed by $d \in \{U, R\}$. These locations start off identical in all respects such as labor, technology, amenities etc., except for two major differences between them: (1) the Urban location has access to greater capital and (2) prevailing winds carry emissions from location R to location U. Identical workers populate both locations.

Each worker derives utility V_d from clean air (the inverse of pollution Z_d) and the real wage $W_d(Z_d)$ which is negatively affected by pollution. Utility is given by $V_d = W_d(Z_d)Z_d^{-\lambda} = \tilde{W}_d Z_d^{-\beta} Z_d^{-\lambda} = \tilde{W}_d Z_d^{-\kappa}$ Here, κ summarizes two separate mechanisms by which pollution can affect utility: by lowering worker health and productivity it may lower wages (β), and it may also directly cause

disutility (λ).[14] Since location U has greater capital, its marginal product of labor is higher ($W_U > W_R$) to begin with.[15] Emissions scale with output, and pollution Z_d is linear in emissions. Now, because U is more productive, output in U is higher than in R. Therefore, $Z_U > Z_R$, both because U produces and emits more, and because emissions in R also disperse to downwind U.

Relative worker utility is $\frac{V_U}{V_R} = \frac{\tilde{W_U}}{\tilde{W_R}}(\frac{Z_U}{Z_R})^{-\kappa}$. Assuming that κ is smaller than 1, higher relative wages outweigh higher relative pollution in U such that utility in U is greater. Therefore, this creates an incentive to migrate from R to U. If migration is frictionless, workers move until utilities are equalized. Given downward sloping demand curve for labor, migration causes relative wages $\frac{\tilde{W_U}}{\tilde{W_R}}$ to shrink whereas relative pollution $\frac{Z_U}{Z_R}$ increases, until they balance out to equate utility across R and U in equilibrium. Note that this reallocation of labor increases overall aggregate productivity, which is equivalent to the labor-weighted average wage in this setting.

Now, let us consider what changes are induced by targeted pollution regulation in R from the initial equilibrium. Because emissions in R disperse pollution to U, regulation in R reduces pollution in both R and U. Both locations become more productive and livable, creating uncertainty regarding which location ultimately offers higher relative utility. The direction of change in relative pollution $\frac{Z_U}{Z_R}$ and relative wages $\frac{W_U}{W_R}$ depends critically on the extent of pollution dispersion. In general equilibrium, some workers will relocate to the location offering greater utility. If workers relocate to the Urban location U, average productivity will increase substantially along with an increase in wages in both locations that is driven by lower pollution. Conversely, if workers move toward the Rural location R, the improvement in average productivity may be modest despite the increase in wages.

Now, consider regulation that targets only the Urban location U in the initial equilibrium, reducing pollution exclusively within U. Such regulation clearly raises relative wages $\frac{W_U}{W_R}$ while simultaneously reducing relative pollution $\frac{Z_U}{Z_R}$. As a consequence, the Urban location U becomes substantially more productive and liveable. Productivity and amenity gains are entirely concentrated among workers in U. In general equilibrium, these improvements incentivize migration from location R to location U until utilities again equalize due to relative wage and pollution adjustments. In the new equilibrium where workers from R relocate to the Urban location U, average productivity across the two locations unambiguously increases.

These two scenarios should make clear that aggregate productivity gains depend on: (1) the marginal product of labor in the locations that pollution regulation cleans up, and (2) whether the change in relative utilities leads workers to migrate to locations with higher marginal product. While this example provides intuition, it ignores important heterogeneity across workers and locations, and several endogenous mechanisms such as congestion and agglomeration. In section 6, I will develop a full quantitative model of worker location choice that can accommodate heterogeneity and includes important mechanisms that govern migration response to pollution

[14] $\tilde{W_d}$ would be the wage if pollution did not impact worker productivity.
[15] Under the relatively benign assumption that capital and labor are complements.

control.

3 Context and pollution control scenarios

In the rest of this paper, I will build on the intuition from the two-location example and develop a spatial equilibrium model of location choice in the context of India. I will then use this model to study productivity gains from targeted pollution control scenarios In this section, I examine why India is a good context to study this question, and then discuss the two archetypal pollution control scenarios through which I understand the interaction of pollution dispersion and worker migration.

3.1 Pollution in India

India experiences some of the highest levels of air pollution globally, with particulate matter concentrations below 2.5 microns in size (PM2.5) significantly exceeding WHO recommended thresholds. Figure 1 shows the spatial heterogeneity in pollution across Indian districts using reanalysis data from Hammer et al. (2020).[16] In 2013, Delhi's annual average PM2.5 concentration was over 150 $\mu g/m^3$, nearly tenfold higher than in major cities such as New York or London. Even though figure 1 shows a large area of pollution over North India, it should be noted that *none* of India's districts meet the WHO annual PM2.5 guideline of 5μg/m3, and 61% of the population lived in districts above India's own substantially higher guideline of 40 μg/m3.[17] Exposure to such severe pollution leads to substantial health damage, including increased incidence of heart attacks, strokes, lung disease, and premature mortality. Existing evidence primarily documents reductions in individual worker productivity (Graff Zivin and Neidell 2012), labor supply and earnings (Hanna and Oliva 2015; Borgschulte et al. 2022; Hoffmann and Rud 2024).

While extensive research highlights health consequences, relatively little is known about the aggregate productivity impacts of such high levels of pollution. Even less understood is how pollution shapes the spatial allocation of labor through its impact on market-level wages and amenities, potentially leading workers to relocate to less productive regions. Such spatial misallocation of labor can substantially limit aggregate productivity growth if workers are directed away from productive cities by pollution, slowing down an important channel of productivity growth (Lewis. 1954).

3.2 Rural sources contribute significantly to urban pollution in India

An entire field of air pollution science, known as source apportionment, is dedicated to identifying and quantifying the origins of pollution. Source apportionment studies utilize chemical transport

[16] The next section describes these data in more detail.

[17] Health impacts start arising at almost any non-zero level of PM2.5 exposure. For this reason, the WHO and the US EPA have been tightening their guidelines over time (US EPA 2015).

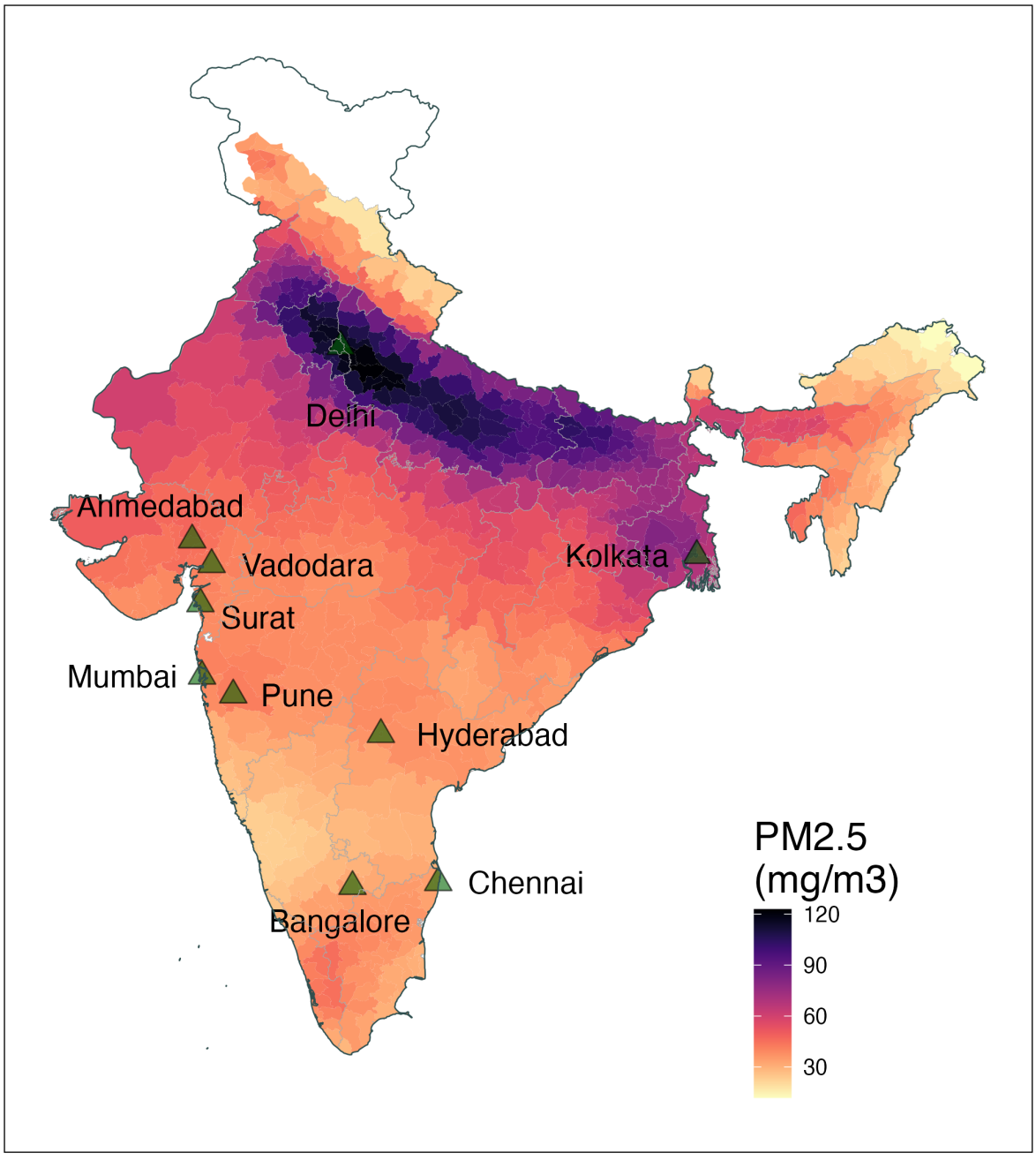


Figure 1: Average PM2.5 concentration in Indian districts (2010). Constructed from satellite-based reanalysis data. See text for more details.

models, receptor modeling, and detailed chemical analyses of particulate matter to trace pollutants back to their origins (Vallero 1973). Such studies are regularly conducted for major cities around the world, including Indian cities, to understand the precise contribution of various pollution sources. For example, studies of air quality in Delhi have consistently shown significant contributions from both localized urban sources—such as vehicular emissions, construction dust, and residential cooking—and distant, rural sources—particularly agricultural residue burning in Punjab and Haryana, located hundreds of kilometers upwind (Singh et al. 2021; Guttikunda et al. 2023). The seasonal surge in crop burning emissions from these rural regions typically leads to severe winter pollution episodes across North Indian cities, significantly amplifying local pollution levels (McDuffie et al. 2021). Effective pollution regulation must not only target localized urban sources but also account for these regional spillovers, recognizing the broader spatial interactions of pollutants. This approach has been adopted in other contexts as well; for instance, the Acid Rain Program in the United States was explicitly designed to control sulfur dioxide emissions from coal-fired power plants across multiple states, successfully addressing regional pollution spillovers and demonstrating significant improvements in air quality and public health (US EPA 2015).

3.3 Pollution control scenarios

In this paper, I study two archetypal pollution control scenarios that are motivated by a real-world policy dilemma facing policymakers on how to clean urban air. Urban policymakers are likely to emphasize upwind emissions from mostly rural areas so that they can avoid taking blame for local emissions that also contribute to high local pollution levels.[18] At the same time, they also have the option of claiming credit if they do clean up local sources. Understanding which of these options yields greater productivity benefits and tax revenues could help policymakers prioritize scarce resources.

These scenarios differ in the location of emissions and the dispersion of pollution. Our interest is in understanding how productivity gains are determined by the interaction of worker migration with these geographic features of the source. The two scenarios under consideration are as follows:

Rural scenario The *Rural* scenario imposes a 10% reduction in emissions from burning in the states of Punjab and Haryana, which accounted for more than half of all burning in 2010. The figure of 10% comes from Jack et al. (2025) who conduct an RCT in Punjab to show that a payment of between INR 2700-4050 per acre to farmers incentivizes them to reduce burning by that amount. They show some evidence that most farmers who responsed to payment chose to use balers to achieve this reduction. Their cost estimates also allow us to conduct benefit-cost calculations for this scenario. Importantly, while the *Rural* scenario targets emissions in rural areas of Punjab and Haryana, it should also reduce pollution in downwind cities in other states due to long-distance

[18] For example, this happens with alarming frequency in the city of Delhi, where seasonal crop residue burning sharply increases pollution levels during the winter and spring months.

dispersion from crop burning. I develop a pollution dispersion model in section 7.1 to understand the extent of pollution reduction in downwind areas.

Urban scenario The *Urban* scenario reduces pollution by a uniform percentage amount in the largest cities such that the total population x change in pollution (population-exposure reduction) is equal for both these scenarios. This scenario reduces pollution exclusively in the 10 largest cities due to the nature of sources it targets. For example, vehicular emissions at roads are known to decay rapidly-even within a few hundred meters-to background levels (Liu et al. 2019). In appendix section 12.2, I combine hourly data from geocoded monitoring stations on pollution and wind with distance from nearest major roads for the largest Indian cities. I find that PM2.5 measured by monitoring stations decays to background levels within 1 kilometer of the road, on average.[19] Therefore, lowering emissions from vehicles in a city is likely to reduce pollution within the city itself since the largest fraction of vehicle-kilometers are driven on roads within the city (McDuffie et al. 2021).

This policy dilemma can also be seen in two recent pollution control programs initiated by the Government of India. The Commission on Air Quality Management Act, 2021 (CAQM) is a statutory body recently set up through legislation with the aim to curb pollution in north India, an area that includes cities such as Delhi, Agra and Kolkata. One of the sources explicitly targeted by the CAQM is crop residue burning in rural areas of north India that disperse smoke over hundreds of kilometers to cities in north India. At the same time, the National Clean Air Program, 2019 (NCAP) is an attempt to clean up urban pollution in cities that are classified as "non-attainment" (i.e. air pollution levels exceed certain set thresholds). Phase 1 of NCAP targets "localized" emissions sources within cities such as vehicles that tend not to affect distant, downwind areas (Ganguly et al. 2020). As yet, there is little evidence to suggest that either of these programs have been successful in reducing pollution. This paper estimates productivity benefits from pollution control scenarios that are closely related to these real-world policies, if they were implemented effectively.

4 Data and measurement

4.1 Air quality

An important consideration for air quality data is complete geographical coverage. Ground-level monitoring station coverage in India is extremely sparse (Greenstone and Hanna 2014). Observations from these stations also may be more susceptible to manipulation (Greenstone et al. 2022; Ghanem and Zhang 2014). On the other hand, satellite imagery-based products provide complete coverage and cannot be manipulated by local actors. The source of remote sensing data on air quality in this paper is Hammer et al. (2020), a gridded reanalysis product of global surface PM2.5 concentrations at a resolution of 0.01°. This product combines satellite imagery data on Aerosol

[19] Background levels in these cities are usually around 40-60% of the level at the road.

Optical Depth with state-of-the-art chemical transport models, and calibrates the output to global ground-based observations. This product has been used in the literature to measure PM2.5 levels in settings where ground level observations are sparse (Khanna et al. 2025). These data are aggregated up to the district level using spatial averaging for analysis.

4.2 Migration

The source of data on migration in this paper is the Population Census of India, 2011.[20]. I leverage novel pairwise migration data across Indian districts sourced directly from the Census Office. A migrant is defined as a person who moved to the district of enumeration at least 6 months ago, and has been living in that district since that move. This definition is likely to capture permanent migration since seasonal migration usually occurs for a few months at a time. Since we are interested in permanent changes in the spatial allocation of labor due to pollution regulation, this type of migration is the appropriate measure to use.

These pairwise data are disaggregated by the period in which the migration occured (last 5 years, 6-10 years and more than 10 years), the educational level (higher degree or not), reason for migration, age groups and gender of migrants. For the reduced form analysis of the effect of pollution on migration, I summarize pairwise migration for "Work/Employment" into migrant inflows and outflows for each district within the 5 years preceding the census (2006-2010), and the previous 5 year period (2001-2005). I relate changes in these inflows and outflows with changes in average pollution over these two periods for each district, as described in section 5.[21] The spatial equilibrium approach has specific data requirements in the form of migration shares from all possible origins to all possible destinations, in order to estimate the amenity and income elasticities of migration. These requirements are satisfied by the pairwise migration data from the census since it provides complete data across all districts.

4.3 Wages

Wage data are utilized mainly in the estimation of migration elasticities and the labor productivity elasticity of pollution. These data come from the National Sample Survey Organization's Employment and Unemployment round 68 for the year 2010-11. The sample provides microdata on individual earnings and work hours within the last 7 days from date of enumeration along with information on 5-digit industry codes; whether the work contract was permanent (salaried) or made on the spot (casual labor); and whether work was done within household, for an employer or on public works. I restrict the sample to individuals aged 18-59, who work for an employer regardless

[20] See https://censusindia.gov.in/census.website/data/census-tables

[21] The primary self-reported reason for female migration in India is for marriage, since the social norm is for newly-married women to move to their husband's town or village. In contrast, the dominant reason for male migration is work-related.

of contract type as the most representative group of people. I also construct average district wage using provided weights for estimation.

4.4 Crop burning

Crop burning, also referred to as agricultural fires, is the practice of setting fire to leftover residue after crop harvest. There are no representative ground-level observations of this phenomenon, but the National Aeronautics and Space Administration (NASA) agency of the United States produces the Fire Information for Resource Management System (FIRMS) product that is widely used to identify such fires. This product provides information on daily, pixel-level fire detection across the world. FIRMS provides a few related products: a Near-Real Time (NRT) fires using the MODIS instrument aboard Terra and Aqua satellites, a quality-controlled standard product from the same instrument but with a 2-3 month lag and another NRT product using the VIIRS instrument from the Suomi-NPP and NOAA-20 satellites. The main difference between the first two and the third is the resolution of the data. MODIS products are at 1 km resolution and are available from 2000 (with higher reliability from 2002 onward when the Aqua satellite was launched) whereas VIIRS products are at 375 m but only available from 2012.

Since the main analysis in this paper relies on data from before 2012, I am unable to use the higher-resolution VIIRS-based product. The primary analysis utilizes the MODIS quality-controlled standard product which differs from the NRT data in that corrections are made to the imprecise location of the Aqua satellite in the NRT data. Imagery data from Aqua and Terra satellites is available at least four times daily for each pixel on Earth and is processed by NASA using a proprietary algorithm to isolate a ground-level fire signal from other signals such as solar flares.[22]

I combine this data with land use data from the European Space Agency Climate Change Initiative's land cover map (version 2.07) [23]. This allows the subset of fires that is found on agricultural land to be separated from natural forest fires since this paper is interested in agricultural fires. I aggregate and resample the land cover data which is at a resolution of 300 m to the fire data grid (which is at 1 km resolution); an indicator for agricultural land use is the main output from this process. All fires are then masked based on this indicator variable to find the subset of agricultural fires.

4.5 Meteorology

Wind data are used to construct exposure to smoke from agricultural fires for every origin-destination pixel pair. Details of the methodology follow in section 7.1 below. These data come from the ERA5 family of global gridded reanalysis datasets produced by the European Center for

[22] Further information on these products is available at https://firms.modaps.eosdis.nasa.gov

[23] Data is available at https://cds.climate.copernicus.eu/cdsapp/#!/dataset/satellite-land-cover

Medium Range Weather Forecasting (ECMWF).[24] Reanalysis data combine ground-level observations and satellite data with chemical transport models that represent physical and chemical processes in the atmosphere to produce reliable and complete coverage for the world. Since ground-level observations are particularly sparse in developing countries, these reanalysis data are widely used in the literature on climate and air pollution (Auffhammer et al. 2013). Hourly wind speed and direction data are taken from the ERA5-Land hourly dataset which is available at a resolution of 0.1°. These are combined with daily agricultural fires at the pixel level to construct the smoke exposure variable, as described in section 7.1 below. Finally, I also construct temporal averages for weather variables including rainfall, temperature and relative humidity from this data set to be used as controls in the regression analysis.

Summary statistics for various variables are provided in table A.1.

5 Pollution and worker migration in India

Air pollution has been documented to affect location choice decisions in various contexts including the US, UK and China (Banzhaf and Walsh 2008; Heblich et al. 2021; Chen et al. 2022). But there is no evidence that worker migration is affected by pollution levels in India. The environmental economics literature posits that since clean air is a normal good, demand for it should increase with income (Greenstone and Jack 2015). India is a much poorer country relative to the countries for which evidence of a migration response has been documented. Therefore, it is important in its own right to investigate whether such a migration response might be found in the context of a poor country such as India.

To answer this question, I estimate equation 1 on migration inflows into and out of districts (indexed by d) in India within two periods: 2006-2010 and 2001-2005. The variable Y_d can be one of three migration outcomes: count of migrants into the district ("in-migration"), count of migrants out of the district ("out-migration"), and net-migration (in-migration - out-migration). I use the natural log transformation for in- and out-migration since those outcomes are heavily skewed. The natural log transformation also allows for an elasticity interpretation for a 1% change anywhere across the entire range of pollution data. I use the inverse hyperbolic sine transformation for net-migration since those values can be negative. For large values of net-migration (above ~10), the coefficient can be interpreted as an elasticity (Bellemare and Wichman 2020).

$$\Delta Y_d = \beta \Delta log(PM_d) + \theta Y_d^{base} + \Delta W_d' \mu + Region\ FE_d + \Delta \epsilon_d \tag{1}$$

Identification strategy Equation 1 relates changes in the migration outcome between 2006-2010 and 2001-2005 to changes in average annual PM2.5 concentrations in those periods. The first dif-

[24] Data is available at https://cds.climate.copernicus.eu

ferences specification removes any fixed determinants of migration at the district level that also determine pollution, such as the presence of a coastline. I allow for separate trends in migration and pollution by controlling for a fixed effect for the three main geographic regions of India: the Indo-Gangetic plain in the north, the peninsular south and Himalayan far north and north-east. These fixed effects control for any omitted cultural or geographic factors that could cause migration patterns to trend differentially across these regions.

Even after controlling for fixed district characteristics, weather and differential regional trends, local changes in pollution over a decade are likely to be determined by economic growth patterns that are also positively (negatively) correlated with migrants inflows (outflows). Therefore, the OLS estimate is likely to be biased toward zero for both inflows and outflows. One solution to this endogeneity problem is to rely an instrumental variable that affects changes in pollution in district d but is plausibly exogenous to economic growth in d. The dramatic increase in agricultural fires or crop burning during the 2000s provides us with an instrument that disproportionately affects pollution in locations that are downwind of burning activity.

Smoke exposure instrument I build on the current approach in the literature that develops instrumental variables using upwind emissions (Freeman et al. 2017; Khanna et al. 2025). I start by calculating the smoke exposure of any location to upwind fires in other districts. The next step is to estimate the impact on PM2.5 of this smoke exposure. These two objects capture how crop burning control policies are likely to reduce downwind pollution.

The first step is to approximate the decay rate of smoke from crop burning emissions E_o in source district o when it gets to receptor district d. I construct a source-receptor smoke dispersal matrix for this purpose. Since agricultural fires are observed at a daily level, I leverage daily variation in wind patterns at the *origin* district to construct this matrix for every year y and $o \neq d$ as follows

$$\omega_{ody} = \left(\sum_{t=1/1/y}^{31/12/y} \theta_{odt} E_{ot} \right) \tag{2}$$

where

$$\theta_{odt} = \frac{wind_{odt}}{distance_{od}}$$

A schematic for the construction of θ_{odt} is shown in figure 2. The numerator $wind_{odt}$ is the daily average fraction of time that the wind at o blows towards d on day t. In order to calculate $wind_{odt}$, I start by assigning each hourly wind observation in o on day t into one of 36 bins of 10 degree span each, based on the wind direction that hour (true north is 0 degree as in the figure). I then construct the wind speed-weighted fraction of time the wind was blowing in each of these 36 bins

by aggregating hourly observations for day t. $wind_{odt}$ is calculated by summing up wind fractions for the bins within the 180-degree cone in the direction of d from o.

Daily smoke exposure from o to d is then calculated by multiplying θ_{od} with daily emissions E_{ot} at origin. Annual smoke exposure for year y from pixel o to d is the sum of these daily smoke exposures during the year. With this source-receptor matrix for each year in hand, I construct the smoke exposure instrument Ω_d^{IV} for destination d as the sum of exposures ω_{ody} from other origins o within 1000 km of the district. I use upwind districts within 1000km as that choice maximizes R2.

$$\Omega_{dy}^{IV} = \sum_{o \text{ within 1000km of } d} \omega_{ody} \tag{3}$$

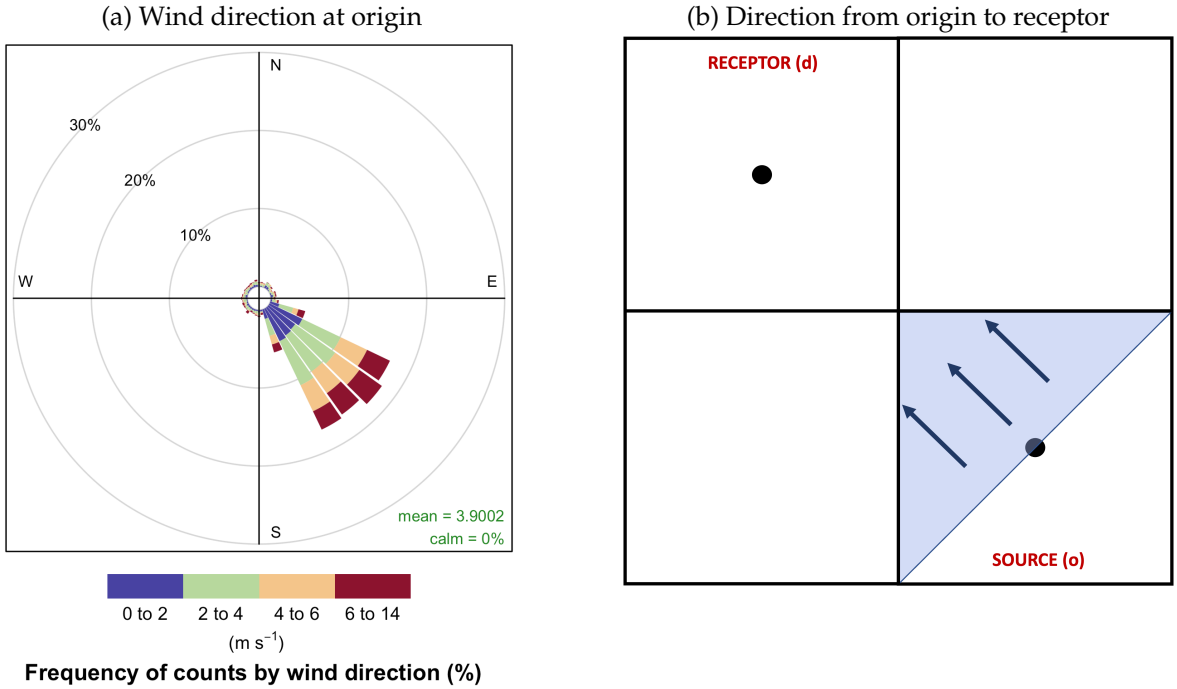


Figure 2: Schematic for construction of the weights $\frac{wind_{odt}}{distance_{od}}$ between source o and receptor d on day t. The windrose in panel (a) shows the daily fraction of time during the day when the wind blows from a given direction, with the radius of the arc indicating the fraction and the colors indicating wind speed. Panel (b) shows the angle between o and d. All bins in panel (a) that fall within the shaded area when superimposed on panel (b) are used to calculate the $wind_{odt}$. The average wind speed within each bin is used to weight the fraction of time in each angle bin.

I instrument for changes in pollution ΔPM_d with changes in smoke exposure $\Delta\Omega_d^{IV}$. The exclusion restriction is that the shift in fire activity in any upwind district s causes changes in migrant

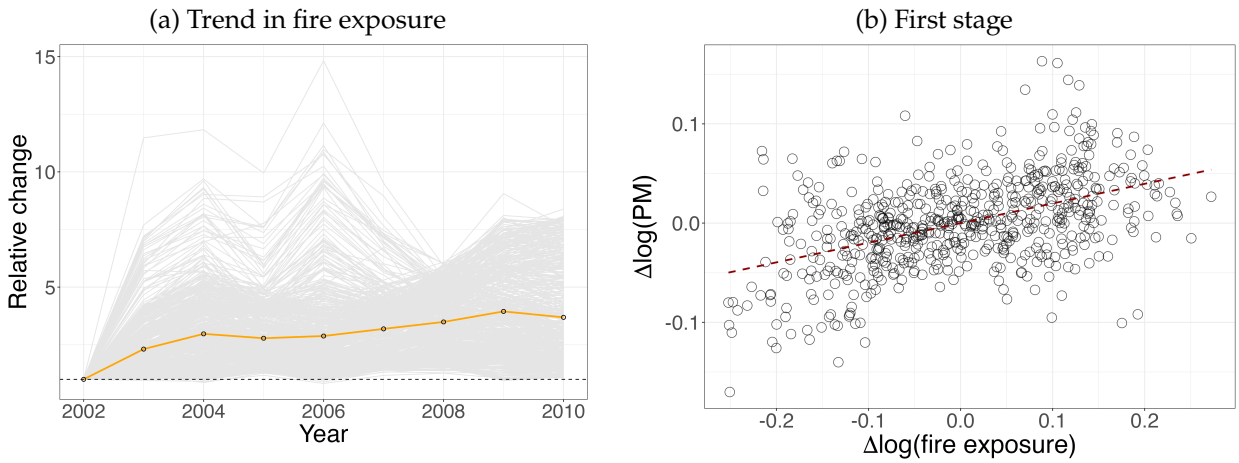


Figure 3: Details on the fire exposure instrument. Panel (a) shows the growth rate in fire exposure for all 617 districts over the 2000s (the value for the base year 2002 is normalized to 1). The sphagetti plot shows growth rates for individual districts whereas the orange line shows the average of the growth rate for all districts. Panel (b) presents the visually the first stage relationship between change in log fire exposure and change in log pollution between 2001-2005 and 2006-2010.

inflow/outflow in d only through its effect on pollution changes in d. I also control for changes in average weather between the two periods since the effect of smoke exposure from fires on pollution may be mediated by meteorological factors. The matrix ΔW_d includes change in rainfall, temperature, wind speed, total cloud fraction and relative humidity.

We may yet be worried that the exclusion restriction is violated if the trend in upwind fire activity in district s were correlated with residual drivers of economic growth and migration in district d. To allay these concerns, I also control for the baseline level of the migration variable in district d from *before* the sample period. Places that were more attractive to migrants before the sample period are likely to remain more attractive during the sample period. Therefore, controlling for this variable should make the exclusion restriction much more plausible.

Results Table A.2 displays the first stage results for three different outcome variables. In all three cases, a 1% change in smoke exposure increases pollution by 0.2%. These estimates are highly statistically significant, indicating a strong instrument. Formal tests of instrument strength are presented along with the 2SLS results. Panel (b) of figure 3 visualizes this first stage by plotting residual log(PM) on residual log(smoke exposure) after regressing these variables on the remaining set of predictors. As an indirect test of the exclusion restriction, table A.5 presents the correlation of the instrument with socioeconomic outcomes from *before* the baseline period. These variables include important covariates such as the literacy rate, female labor force participation and night lights that mediate or determine economic growth patterns. Each of these variables is significantly correlated with the instrument, implying that districts that were slightly poorer to begin with were more affected by crop burning in the 2000s. So, to be conservative, I will interpret the 2SLS results

as only suggestive of the true effect.

Table 1 presents the main empirical estimates for the impact of PM2.5 on migration for work among those aged 20-64 years. Column 1 reports the OLS estimate for the effect of a 1% increase in PM2.5 on net-migration. The effect of -1.68 inverse hyperbolic sine (ihs) units is small and not statistically significant. Column 2 reports the 2SLS estimate using the smoke exposure instrument. The coefficient estimate is statistically significant and the value is -28. Since the specification is linear-log (outcome being in ihs units), this implies that a 1 % increase in PM2.5 is associated with a significant reduction in net-migration by -0.28 inverse hyperbolic sine units. The mean value of net-migration is -173. This is large enough in absolute terms that we can interpret -0.28 as an elasticity. Column 3 shows that the OLS estimate for the natural log of migration inflows is -0.63. While it is biased toward zero, the estimate is still significant. The 2SLS estimate in column 4 increases in larger in size at -2.12 and is also more statistically significant. This indicates an elasticity of migration inflows to PM2.5 of -2.12%. Columns 5 and 6 report results using the natural log of migration outflows. The OLS estimate is not different from zero whereas the elasticity of migration outflows to PM2.5 using the 2SLS estimate is positive and meaningfully large at 0.86%. First-stage KP F-statistics are reported under the 2SLS columns for each of the three outcomes. These range between 126-131, confirming the strength of the instrument. Taken together, the results from this table show evidence that increases in pollution across Indian districts in the 2000s are correlated with lower migration inflows and higher outflows, with a strongly negative correlation with net-migration to more polluted districts.

Table 1: The effect of pollution on migration

Dependent variable:	Δ ihs(Net-migration)		Δ log(In-migration)		Δ log(Out-migration)	
	OLS	2SLS	OLS	2SLS	OLS	2SLS
	(1)	(2)	(3)	(4)	(5)	(6)
Δ log(PM2.5)	-1.68	-27.59	-0.63	-2.12	-0.13	0.86
	(3.24)	(7.38)	(0.28)	(0.52)	(0.18)	(0.41)
Region FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	624	624	626	626	624	624
KP F-stat		130.95		126.06		130.04
Mean (1000s)		-0.17		7.80		7.92
SD (1000s)		21.82		22.89		8.90

Notes: Estimates from regressions of first-differenced migration variables on pollution between 2002-2005 and 2006-2010. The outcome variables are winsorized at the 1% level to minimize the influence of outliers. Controls include baseline migration variable measured for the 1990s, region fixed effects and average weather. ihs refers to inverse hyperbolic sine. Robust standard errors are reported in parentheses. The mean and SD of the change in the level of the relevant migration are reported in the last two rows (in 1000s of workers).

Robustness checks Appendix Table A.3 examines an alternative specification using migration per capita instead of the log transformation. These results are consistent with the results in table 4.2, although the effect on in-migration becomes statistically insignificant.. Note that this is a linear-log specification, so the equivalent semi-elasticities (or the effect of a 1% increase in PM2.5 on per capita net-migration/in-migration/out-migration) can be obtained by multiplying the denominator by 100 (for a semi-elasticity in migrants per 100,000 residents).[25] The semi-elasticity estimates are 7 fewer in-migrants per 100,000 residents, 20 more out-migrants per 100,000 residents, and 28 fewer migrants on net per 100,000 residents. First stage F-stats continue to be between 126-129, indicating a strong instrument.

Our primary goal is to understand the effect of pollution on migration for the average district, not the average person, since we are interested in the effect of pollution regulation scenarios that are geographically targeted. So our preferred estimates are not population weighted. Appendix table A.4 compares results with and without population weighting. We see that the results with population weighting change little from table 1. The main difference is that the confidence interval for the 2SLS estimate on migration outflows barely includes zero with population weighting. However, the confidence intervals for all the 2SLS estimates overlap substantially in both tables.

Since the instrument is correlated with socioeconomic variables at baseline, we must be careful to not overly interpret these results as being causal. But, we can include these variables as additional controls in equation 1 to test if the estimates change substantially. One limitation of this exercise is that the sample size is almost 10% smaller as the socioeconomic variables are not available for all districts. Appendix table A.6 shows the results. We see that all coefficients shrink when controlling for these socioeconomic variables. But the result for migration inflows is still significant and large in magnitude at -1.53. It is reassuring that we find a strong effect on migration inflows even after including fixed effects and a whole range of controls.

Heterogeneity by educational attainment We might think that the effect of pollution on migration differs by the skill level of workers (as measured by their level of education). More educated workers usually earn more and may have access to more information about the health effects of pollution. They are also less likely to be credit-constrained compared to less educated workers. In fact, Khanna et al. (2025) find that the migration response to pollution is larger for skilled workers, i.e. those with a degree or higher. I test this hypothesis in India by estimating 1 on migrant flows for individuals without a college degree and those holding a college degree or higher.

Table 2 presents evidence that the pollution-driven migration effect does not differ by education attainment. Each of the columns in table 2 are 2SLS estimates, with the migration outcome being one of the ihs(net-migration) or log(in-migration) or log(out-migration). Within these migration, outcomes, a column header specifies whether the sample consists of those with a degree or those without. Focusing on the first two columns, a 100% increase in PM2.5 is associated with a decrease

[25] The coefficient of a linear-log specification provides the effect of a 100% increase in the predictor.

in net-migration by -12.04 and -25.28 *ihs* units for those with a degree and those without a degree, respectively. These coefficients are highly statistically significant and the first stage KP F-stat are 126 and 129. We can translate these coefficients into elasticities as follows: a 1% increase in PM2.5 is associated with reduced net-migration by -0.12% and -0.26% for those with a degree and those without respectively, with 95% confidence intervals ranging from [-0.23%, -0.014%] and [-0.4%, -0.09%].[26] Columns 3 and 4 show that the elasticity for migration inflows is -1.05 and -2.92 for those with a college degree and those without, respectively. Columns 5 and 6 show that the elasticity for migration outflows is smaller, at 0.05 and 0.9, with the coefficient for those with a college degree not being different from zero. Taken together, this table provides evidence that the effect of pollution on migration does not differ substantially by skill.

Table 2: The effect of pollution on migration by level of education

Dependent variable:	Δ ihs(net-migration)		Δ log(in-migration)		Δ log(out-migration)	
	w/ degree	w/o degree	w/ degree	w/o degree	w/ degree	w/o degree
	(1)	(2)	(3)	(4)	(5)	(6)
Δ log(PM2.5)	-12.04	-25.28	-1.05	-2.92	0.05	0.90
	(5.44)	(7.84)	(0.45)	(0.67)	(0.30)	(0.45)
Region FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	620	620	623	623	623	619
KP F-stat	125.83	129.11	124.04	125.03	126.57	131.23
Mean (1000s)	-0.04	-0.19	1.98	5.59	2.00	5.73
SD (1000s)	5.59	16.93	6.26	17.31	3.16	6.32

Notes: Estimates from 2SLS regressions of first-differenced migration variables on pollution between 2002-2005 and 2006-2010, by level of educational attainment. The outcome variables are winsorized at the 1% level to minimize the influence of outliers. Columns (1), (3) and (5) present estimates for those with a degree or above while coluns (2), (4) and (6) present estimates for those without a degree. Controls include baseline migration variable measured for the 1990s, region fixed effects and average weather. ihs refers to inverse hyperbolic sine. Robust standard errors are reported in parentheses. The mean and SD of the change in the level of the relevant migration are reported in the last two rows (in 1000s of workers).

5.1 Why might pollution affect migration in India?

While these results provide evidence for a sorting response to pollution in India, they do not explain the mechanisms behind this finding. This paper hypothesizes two potential channels through which air pollution can affect migration decisions. The first channel is an effect on individual earnings that operates through the capitalization of well-documented harmful effects on physical worker productivity on market wages. While much of this literature focuses on the intensive margin of reduction in worker output per hour (Graff Zivin and Neidell 2012; Chang et al. 2019; Fu et al. 2021), there also may be an extensive margin reduction in the number of hours worked due to

[26] Since the absolute value of mean net-migration is >10 for both samples, we can interpret these values as elasticities.

air pollution (Hanna and Oliva 2015). The net result of these intensive and extensive margin effects is a reduction in total output and in the marginal product of labor. Firms may adjust to lower output and profits by attempting to lower the wages they pay. This is consistent with recent evidence on the negative effect of air pollution on worker income in the US and Mexico (Borgschulte et al. 2022; Hoffmann and Rud 2024). Figure 4 also presents correlations that are consistent with this hypothesis from India: real wages are lower in districts with higher pollution.

Secondly, clean air (or lower air pollution) can also be thought of as an amenity (for example, Chay and Greenstone 2005; Banzhaf and Walsh 2008). The literature on air pollution in developing countries suggests that workers have a low willingness to pay for clean air, including for reasonably cost-effective adaptations such as mask-wearing (Greenstone and Jack 2015; Baylis et al. 2023). This is consistent with the notion of clean air as a normal good, so that demand for it would be muted in a low-income country like India. But, the findings in this paper that districts with higher pollution receive lower migrants could also arise if workers avoid polluted cities due to preferences that are correlated with clean air. This second mechanism would be consistent with recent work documenting that the dis-amenity costs of living in cities lead some workers in India to forego gains of up to 35% of income by not migrating to cities (Imbert and Papp 2020b).

The effect of pollution on inflows of migrant workers is consistent with either mechanism. The quantitative model can help guide us toward a solution to the empirical challenge of separating the amenity and income channels through which pollution affects location choice of workers.

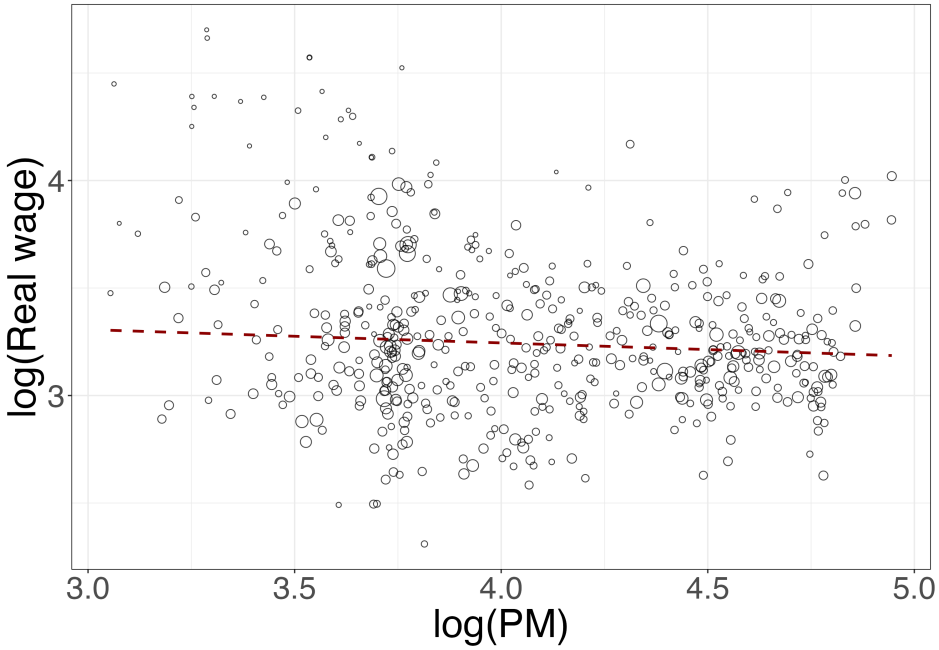


Figure 4: Wages are negatively correlated with pollution (2011). Dots represent average values for each district with dot size indicating the sum of weights provided by the National Sample Survey Office's Employment and Unemployment Survey, 2011-2012. A weighted linear regression fit is included. Inference is conducted in section 7.4.

6 Spatial equilibrium model with pollution dispersion

This section incorporates air quality spillovers into a canonical quantitative model of economic geography (Redding and Sturm 2008) to investigate how the movement of people and pollution interacts to determine income gains from spatially targeted pollution control measures. The reader is referred to Redding and Rossi-Hansberg (2017) for a survey of the economic geography literature on the development of these models.

6.1 Worker preferences

There are $\overline{L}_o$ workers in location o to begin with. Worker j has preferences over a consumption good C_d, amenities B_d and air quality (the inverse of air pollution level Z_d) as follows.

$$U_{jod} = \varepsilon_{jod} B_d Z_d^{\lambda} C_d exp(-m_{od})$$

B_d consists of a fixed component that can include climate and other amenities, as well as an endogenous component that varies in response to congestion. An example of such a congestion force is the cost of housing that depends on the housing supply elasticity. Z_d is the level of air pollution in location d. If workers have preferences over clean air, locations will be characterized by compensating differentials for pollution, with elasticity given by λ. If $\lambda < 0$ then pollution does indeed have amenity value for workers.

ε_{jod} is an idiosyncratic preference shifter that captures preferences for location d for worker j who originates from location o. ε_{jod} is i.i.d across workers and locations, and is drawn from a Frechet distribution given by the CDF $F(\varepsilon) = e^{-\varepsilon^{-\eta}}$. The parameter η controls the dispersion of these shocks. A small value of η implies that the probability of a large draw for ε is larger, implying that the worker is particularly attached to the origin location o and would not move even with large wage or amenity differentials between origin and destination. This can captures real world features such as strong local ties, for example. The parameter η can also be interpreted as the income elasticity of migration across districts.

In this quasi-dynamic model, worker originating in location o has made a decision on whether to stay or move to another destination location d when we observe them. But movement across locations is costly. This migration cost $exp(-m_{od})$ from origin o to destination d may represent physical costs of migration, salient differences in culture and language, and also policy differences such as access to welfare benefits that are attached to the location of birth. About 80% of migration in India is within the state, an entity that shares a common language and cultural features as well providing exclusive access to welfare benefits for residents.

Each worker provides one unit of labor and there is no labor-leisure trade-off. Therefore, income from labor in location d is given by wage w_d. The price of the consumption good is given by

P_d. Workers choose the location where they receive highest utility, subject to moving costs. If the indirect utility function for worker j is represented by V_{jod}, then the worker chooses d over d' if $V_{jod} > V_{jod'}$. Indirect utility function is given by

$$V_{jod} = \varepsilon_{jod} B_d Z_d^{\lambda} (\frac{w_d}{P_d}) exp(-m_{od}) \tag{4}$$

This formulation allows us to express the migration share π_{od} from o to d as follows, using the properties of the Frechet distribution. The derivation is provided in appendix section 12.5.

$$\pi_{od} = \frac{L_{od}}{\overline{L}_o} = \frac{[B_d Z_d^{\lambda} (\frac{w_d}{P_d}) exp(-m_{od})]^{\eta}}{\sum_{k=1}^{N} [B_k Z_k^{\lambda} (\frac{w_k}{P_k}) exp(-m_{ok})]^{\eta}} \tag{5}$$

All of the local income is derived from wages and is completely spent on demand for the consumption good. Therefore, total demand D_d is given by

$$D_d = w_d L_d$$

6.2 Production and general equlibrium

Each location d produces a homogeneous good Y_d using a linear technology with labor L_d and TFP A_d. Each worker supplies one unit of inelastic labor. TFP varies across locations and may be affected by pollution Z_d and agglomeration forces.

$$Y_d = A_d L_d$$

Markets are perfectly competitive. Therefore, the price of the consumption good equals marginal cost.

$$P_d = \frac{w_d}{A_d}$$

There is no goods trade, so the consumption good is produced and consumed locally. Output is then determined purely by demand D_d. Assuming the consumption good to be the numeraire ($P_d = 1$), the wage in each location is pinned down by

$$w_d = A_d \tag{6}$$

Model equilibrium is characterized by the following equation.

$$\begin{aligned}\pi_{od} = \frac{L_{od}}{\overline{L}_o} &= \frac{\left[B_d Z_d^{\lambda} w_d exp(-m_{od})\right]^{\eta}}{\sum_{k=1}^{N}\left[B_k Z_k^{\lambda} w_k exp(-m_{ok})\right]^{\eta}} \\ &= \frac{\left[B_d Z_d^{\lambda} A_d exp(-m_{od})\right]^{\eta}}{\sum_{k=1}^{N}\left[B_k Z_k^{\lambda} A_k exp(-m_{ok})\right]^{\eta}}\end{aligned} \tag{7}$$

The equilibrium allocation of labor in location d is given by

$$L_d = \sum_{o=1}^{N} \pi_{od} \overline{L}_o \tag{8}$$

Productivity A_d, amenity B_d and pollution Z_d endogenously adjust to reach equilibrium population. The next sections describe these adjustment mechanisms and associated elasticities. The population vector is the variable that adjusts until equilibrium is reached in the model, pinning down the values of all other endogenous variables that depend on it.

6.3 Productivity is endogenous due to agglomeration and pollution

TFP varies by location due to fixed exogenous factors like soil quality, presence of rivers, or availability of raw materials like mineral ores; agglomeration forces; and the effect of pollution on worker productivity. Equation 9 formalizes these ideas. $\overline{A}_d$ is exogenously determined productivity that does not depend on pollution or agglomeration.

$$A_d = \overline{A}_d Z_d^{\beta} L_d^{\phi} \tag{9}$$

If $\beta < 0$, productivity is negatively affected by pollution. This is the mechanism through which the health effects of air pollution on individual workers accumulate to affect the aggregate marginal product of labor in a location. It captures the complementarity of labor with other inputs such as capital and technology in the production process.

Productivity is also affected by agglomeration forces that increase when a greater number of workers populate a location (Combes and Gobillon 2015), and arises from any potential non-excludable innovation (Arrow 1962). The strength of these agglomeration forces is captured by ϕ.

6.4 Amenity is endogenous due to congestion

Non-pollution amenity value of a location depends on endogenous factors such as housing rental prices. The elasticity γ captures these factors, and we expect it to be negative. As more workers

move into a city, congestion forces such as rental rates rise, making the city slightly less desirable for the next migrant.

$$B_d = \overline{B}_d L_d^{\gamma} \tag{10}$$

6.5 Pollution is affected by local and upwind emissions

Pollution Z_d in location d is modeled as follows

$$Z_d = \overline{Z}_d \, L_d^{\psi} \, g(\Omega_d) \tag{11}$$

The component of pollution $\overline{Z}_d$ captures pollution from emissions sources within the district, including fixed sources such as naturally occurring forest fires. This component varies across districts and may also be understood as encompassing any control technologies that can reduce the pollution intensity of local emissions. The second component L_d^{ψ} captures pollution arising from economic output and activity. More people correlates with greater emissions from industry, higher vehicle-miles driven etc. We expect ψ to be positive. The third component $g(\Omega_d)$, is the contribution of all upwind emissions to pollution in location d. The function g(.) summarizes the pollution dispersion model, and I will describe its construction in section 7.1.

7 Identification and estimation of model parameters

Apart from the agglomeration and congestion elasticities, I estimate all the other parameters of the model. This section will present the estimation strategies for each of these parameters and discuss results. Summary statistics for the data used in the various parameter estimation exercises are provided in table A.1.

7.1 Pollution dispersion model ($g(\Omega_d)$)

In this section, I estimate the component of pollution Z_d that is explained by the dispersion model $g(\Omega_d)$. This model allows us to calculate pollution reduction due to crop burning under the *Rural* scenario. The main determinant of long-distance pollution dispersion from a given source is whether the emissions plume reaches a height where ground-level convective processes are unable to constrain the plume to remain close to the source location. Once captured by upper atmospheric winds, the plume can travel hundreds or even thousands of kilometers (Vallero 1973). As an example of this phenomenon, crop residue burning in rural areas is known to increase air pollution and cause public health issues in distant, downwind cities (Singh et al. 2021).

Although air pollution modelers have worked on understanding the contribution of crop residue burning on downwind pollution (Guttikunda et al. 2023), the tools used to conduct such studies are not amenable to the kind of policy analysis I perform due to the extremely high computational complexity demanded by full-scale Chemical Transport Models. A new generation of so-called Reduced Complexity Models such as InMap (see Tessum et al. 2017) have been developed for wide use in the US, but global versions are not in common use for countries as India. These tools also suffer from under-prediction of pollution from crop burning in India largely due to outdated emissions inventories (Thakrar et al. 2022). Therefore, in the next section, I develop a statistical method that facilitates a spatially explicit yet computationally feasible smoke dispersion model of where air quality would improve as a result of controlling crop residue burning.

Step 1: Calculate total smoke exposure This model builds on the the source-receptor smoke dispersal matrix summarized in equation 2. Unlike with the instrument in section 5, this pollution dispersion model intends to capture pollution from *all* smoke exposure, including own-district burning. This is a crucial difference with the use of upwind burning as an instrument; our objective is to construct a structural object that will be used in counterfactual analysis. Using the source-receptor smoke exposure matrix from equation 2, I construct total smoke exposure Ω_d for destination d as the sum of exposures ω_{ody} from all origins *o plus* the local annual emissions $E_{dy} = \sum_{t=1/1/y}^{31/12/y} E_{dt}$.

$$\Omega_{dy} = E_{dy} + (\sum_{o \neq d} \omega_{ody}) \tag{12}$$

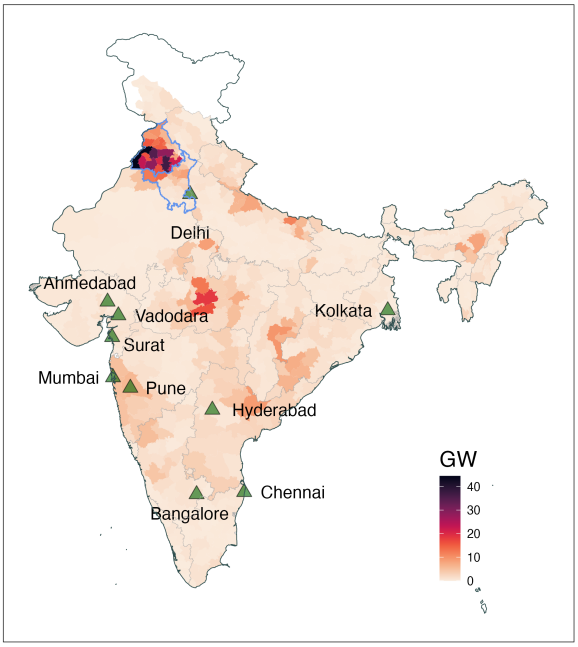
The vector Ω summarizes the smoke exposure of any given location to crop burning in *all* other locations, accounting for daily changes in fire activity and wind pattern at those locations. Panel (a) of figure 5 visualizes the locations where burning is more common. Panel (b) visualizes the smoke exposure Ω_d of districts.

Step 2: Estimate effect of smoke exposure on pollution Recall that equation 11 specifies the structural relationship between PM2.5 and smoke exposure Ω_d, with the transformation g(.) translating exposure to pollution. I specify this transformation to be linear in $log(\Omega_d)$. I assessed various specifications on best fit against a measure of true PM2.5 from agricultural fires provided by a scientific group based at IIT Delhi and Washington University of St. Louis (WUSTL) (McDuffie et al. 2021). Other specifications included a power law in level of Ω_d, a power law in $log(\Omega_d)$, and linear in Ω_d. The specification that performs best against this measure of true PM2.5 from fires is the linear-log. Further detail on the performance of these models is provided in appendix section 12.1.

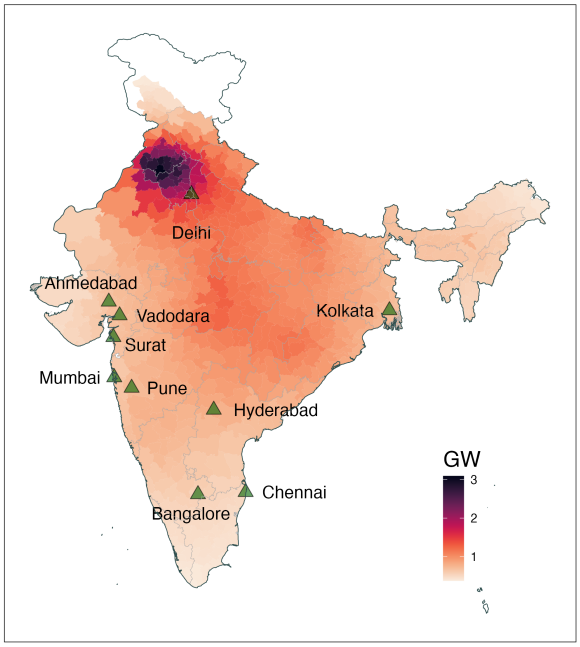
$$\text{PM from smoke}_d = g(\Omega_d) \equiv \tau_d * log(\Omega_d) \tag{13}$$

This contribution of annual smoke exposure to average PM2.5 concentration may be mediated by

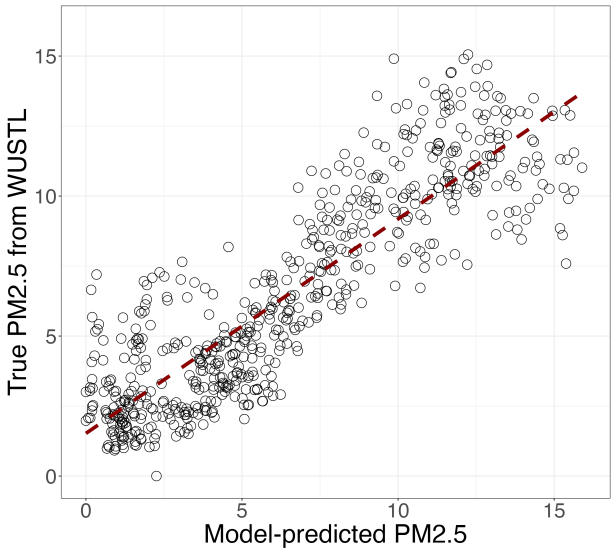
(a) Hotspots of burning activity



(b) Annual smoke exposure of districts



(c) Pollution dispersion model performance



(d) Predicted PM2.5 from crop burning

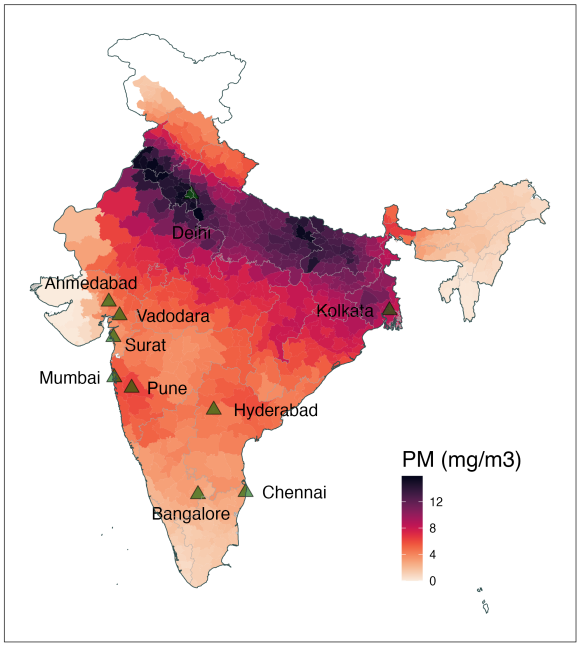


Figure 5: Example training data for pollution dispersion model and performance of the model. Panel (a) shows districts where burning activity was most widespread by plotting total annual fire radiative power for all fires within the district in 2016 (the first term in equation 12). Panel (b) shows the upwind component of total smoke exposure to all upwind fires (the second term in equation 12). Panel (c) plots predicted fire-driven PM2.5 in 2016 from the calibrated pollution dispersion model after calibration against fire-driven PM2.5 estimated using a Chemical Transport Model by scientists at the Washington University of St. Louis (see Steps 2 and 3 in this section for details on estimation and calibration). Panel (d) maps predicted fire-driven PM2.5 in 2016.

topography and meteorology. The smoke exposure variable does not account for either of these mediators. In order to capture the heterogeneity in this relationship across India, I estimate separate regressions for each district d on a sample that includes d and the set of its neighboring districts $s(d)$, with farther districts getting lower weights based on distance. This allows the relationship between smoke exposure and pollution to be estimated using localized variation in topography and meteorology, thus capturing spatial heterogeneity in this relationship. Equation 14 provides the estimating equation based on equation 13

$$PM_{s(d),y} = \tau_d log(\Omega_{s(d),y}) + D_{s(d)} + \epsilon_{s(d),y} \tag{14}$$

Let us remind ourselves again that if we had access to source-receptor PM2.5 contribution for smoke exposure, we would not need to conduct this exercise at all. Here I estimate this relationship using annual average PM2.5 from all sources estimated using satellite data from Hammer et al. (2020). Our goal is to isolate the variation in this outcome that is explained by total smoke exposure. I estimate equation 14 on panel data between 2002-2016 for the set of districts $s(d)$. The parameter τ_d is identified through within-district time series variation over 2002-2016 in smoke exposure. The district fixed effect aids with identification by removing the fixed component of pollution that may reflect higher industrialization or population, leaving the time series variation that includes the part of pollution that is caused by smoke exposure. The main identifying assumption is that pollution levels in d do not lead to abatement of crop burning locally or in upwind districts. This assumption is likely to be satisfied, given that most regulations on crop burning are not implemented (Jack et al. 2025).

Step 3: Calibrate model using true fire-driven PM2.5 from Chemical Transport Model Note that the τ_d parameters are identified up to a constant ($s * \tau_d$) only since we use variation in total PM2.5. The power law models are better able to identify rescaled parameters because the log transformation of the RHS may absorb the constant s term in the fixed effect, but the outcome scale is also log in that case. The constant cannot be separated out at all in the linear-log model since PM2.5 is linear in $log(\Omega_d)$. In order to rescale the τ parameters for all these model, I calibrate the mean of the predicted PM2.5 according to the uncalibrated model to equal the mean of true fire-driven PM2.5 according to the WUSTL analysis. Specifically, I match these moments for 2016, the year for which WUSTL conducted their study. These rescaled τ_ds are used to make final PM2.5 predictions using the calibrated model.

Results Panel (c) of figure 5 documents the correlation with the WUSTL fire-driven PM2.5 data that uses a full-scale Chemical Transport Model. Visually there is a strong relationship between predicted and true fire-driven PM2.5, and appendix table A.7 shows that the R2 of this relationship is 0.77. The power-law models also perform decently but have less than half the predictive power of the linear-log relationship. Panel (d) of figure 5 plots these fire-driven predicted PM2.5 values. The

predicted fire-driven PM2.5 from the pollution dispersion model overlaps with overall PM2.5 levels in 1. As expected, PM2.5 from agricultural fires is higher over the northern states including Delhi and over the Indo-Gangetic belt, as suggested by various full-scale Chemical Transport Models (CTM). But recall that the major benefit of our approach over full-scale CTMs is that it can be used to conduct quick counterfactual scenarios.

Fire-driven pollution in the quantitative model In contrast to the empirical exercise to estimate τ_d, the smoke exposure object that I embed into structural equation 11 for the quantitative model is constructed using an estimated source-receptor matrix that depends on *average* wind patterns and emissions over 2001-2010. This allows us to characterize expected changes in pollution that migrants are likely to consider in their location choice decisions in response to the pollution control scenarios of interest. To differentiate the long-term average exposure object from the empirical object, I use a tilde in $\widetilde{\Omega}_d$.

$$g(\widetilde{\Omega}_{od}) \ = \tau_d \, log(\sum_{o \neq d} [\tilde{\theta}_{od} \widetilde{E_o}] + \widetilde{E_d}) \tag{15}$$

where

$$\tilde{\theta}_{od} = \frac{E[wind_{od}]}{distance_{od}} \tag{16}$$

The expected value of $wind$ in the numerator in equation 16 captures prevailing wind patterns at each origin; I construct these using 2001-2010 annual average wind patterns, weighting by the daily average fire count to capture seasonality in fire activity at source. Therefore, this method gives zero weight to wind patterns on days with no fire activity at the source location. Similarly, emissions E_o is also constructed using a 10-year average.

7.2 Income and pollution elasticities of migration (λ, η)

The equilibrium migration shares predicted by the quantitative model in equation 5 provide a means to separately identify η and λ - the income and pollution elasticities of migration respectively - using data on migration shares across Indian districts. I utilize migration data over the 10 years prior to the census enumeration period of February 2011. This covers the period substantial but uneven employment growth and job opportunities in cities pulled in migrant workers, thus creating strong variation in migration that is explained by equation 5.

In order to estimate this equation on migration shares data, we need to specify migration costs $exp(-m_{od})$. I assume that m_{od} can be parameterized such that migration costs are normalized ($exp(-m_{oo}) = 1$) and symmetric ($exp(-m_{od}) = exp(-m_{do})$).

$$m_{od} = \nu_1 ihs(dist_{od}) + \nu_2 \mathbb{1}(lang_{od}) + \nu_3 \mathbb{1}(state_{od})$$

Here $dist_{od}$ measures the physical distance between districts o and d, $\mathbb{1}(lang_{od})$ in an indicator for whether a different language is spoken in o and d and $\mathbb{1}(state_{od})$ in an indicator for whether district o and d belong to different states. The inclusion of language and state dummies are motivated by the literature on the determinants of migration in India (for example, see Kone et al. 2018).

Taking the natural log of equation 5 gives us the stochastic version of the migration equation that we take to the data

$$\begin{aligned}
log(\pi_{od}) = {} & \eta\, log(wage_d) && \text{[Real wage]} \\
& + \eta\lambda\, log(PM_d) && \text{[Pollution disamenity]} \\
& - \eta\nu_1 log(dist_{od}) - \eta\nu_2 \mathbb{1}(lang_{od}) - \eta\nu_3 \mathbb{1}(state_{od}) && \text{[Migration cost]} \\
& - V_o && \text{[Origin option value]} \\
& + \eta\, log(B_d) + e_{od} && \text{[Residual]}
\end{aligned}$$

where $V_o = log(\sum_{k=1}^{N} [B_k M_{ok} Z_k^{\lambda} w_k]^{\eta})$ is fixed within each origin, and the residual contains destination amenities and other idiosyncratic features that determine bilateral migration shares.[27]

Identification challenges for η and λ There are two main challenges: (1) the residual might contain destination-specific amenities as part of B_d that are correlated with wages or pollution, for eg. a coastal location that makes the location more desirable for some individuals while also reducing PM2.5 levels due to the sea breeze; and (2) the residual might contain origin-destination specific omitted factors in ϵ_{od} such as pre-existing migrant networks that can affect current migration patterns, even as those past networks may have been formed partly because the destination had higher past wages that also influence current wages. The solution I adopt to these identification problems is to instrument for both *log(PM)* and *log(wage)*.

Instrument for $log(wage)$ Past origin-destination networks from more than 10 years ago reflect historical connections that may also be correlated with current wage differentials across those district pairs. Through this past migrant network channel, wages in district j might affect wages in district d today. This motivates an instrument for $log(wage)$ that is specific to the origin-destination ($o-d$): wages in other origin districts $s \neq o$ affect wages in district d but are uncorrelated with the $o-d$ migration residual. This assumption is more plausible if we restrict j to the set of districts

[27] Pollution and wage are measured in 2010 and 2011 respectively. I use these annual values rather than multi-year averages so as to ensure instrument relevance. Since they are both strongly auto-correlated, we choose the year right before the census period to ensure maximum relevance. But wages are available only for the census year of 2011, not for the year before it.

in other states so as to avoid spurious correlation in wages between o and j that is driven by geography or state-level policy. In order to construct this $(o-d)$ specific wage instrument, I weight $(j-d)$ wages by the historical share of migration into d that arose from j.[28] Equation 17 formalizes the instrument.

$$\text{wageIV}_{o-d} = \sum_{j \ni S(j) \neq S(d) \& j \neq o} (\text{historical migration share}_{j-d}) * \text{ wage}_j \tag{17}$$

Here, $S(d)$ refers to the state to which district d belongs. Essentially, we sum up weighted wages from origin districts j that belong to all other states. The identification assumption is that *wageIV* affects migration from o to d only through its effect on wages in d. The main threat to this assumption is the existence of omitted determinants of wage in j that are also correlated with wage in d. Examples would be spillovers due to geographic proximity or similar labor policies. Because I construct this instrument by excluding districts from the same state, this threat is minimized.

Instrument for $log(PM)$ I instrument for *log(PM)* with the annual smoke exposure instrument developed in section 5. The identification assumption is that upwind burning affects migration from o to d only through the increase in smoke-driven pollution in d. In section 5, we saw that while this is a strong instrument there may be some residual concerns in the first-differenced specification. But the origin-destination specification that allows the inclusion of origin and state fixed effects as well as destination wages substantially weakens the identification assumptions, and therefore screens out most pathways through which the fire exposure instrument could violate the exclusion restriction.

First, origin fixed effects absorb all push-side determinants of migration that are constant across an origin district's choice set: local labour shocks, population size, caste composition, even origin-specific attitudes toward pollution. State fixed effects then restrict identification to within-state comparisons of destination districts that share broadly similar institutions, climate, and policy regimes. Together, these two sets of dummies mean that the instrument is now making more conservative comparisons across districts that are more alike in various respects such as same state-level agricultural policy, monsoon cycle and development trajectory, differing only in the idiosyncratic, wind-driven component of fire exposure.

Second, by controlling for destination wages we purge the main economic pull factor—expected earnings—that could affect migration flows through the effect of upwind burning (due to productivity effects). This allows us to isolate the amenity effect of pollution separate from the productivity channel. Thus, the remaining identifying variation for λ comes from quasi-random dispersion shocks in particulate matter that are plausibly orthogonal to other determinants of labor inflow.

[28] Historical migration data are for migration that occurred prior to 2001. These data are also available from the census.

Estimation procedure I estimate equation 7.2 using the Poisson Pseudo-Maximum Likelihood (PPML) method of Silva and Tenreyro (2006), as detailed in equation 18. This has several advantages over OLS. First, unlike OLS, it handles zero values for migration count. Second, it respects the adding up constraint for migration count that such gravity models imply (Fally 2015). Third, because of Jensen's inequality, the least squares estimator of $log(L_{od})$ on the right hand side variables is generally an inconsistent estimator for the model elasticities, whereas PPML is consistent (Santos Silva and Tenreyro 2022).[29]

$$
\begin{aligned}
L_{od} = exp(\ & \eta\, log(wage_d) && \text{[Real wage]} \\
& + \eta\lambda\, log(PM_d) && \text{[Pollution disamenity]} \\
& + \eta\nu_1 ihs(distance_{od}) + \eta\nu_2\, \mathbb{1}(Lang_{od}) + \eta\nu_3\, \mathbb{1}(State_{od}) && \text{[Migration cost]} \\
& - \overline{V}_o && \text{[Origin option value]} \\
& + \epsilon_{od}) && \text{[Residual]}
\end{aligned}
\tag{18}
$$

Although the instruments noted above help solve the endogeneity problem, an estimation challenge arises for nonlinear panel models like PPML with IV as the incidental parameters problem makes estimation inconsistent. I solve that estimation challenge by following the method described in Lin and Wooldridge (2019) who recommend adopting a control function approach that proceeds in two steps: (1) estimate the first stages using OLS and store the residuals from each of these first stages (2) include these residuals in the PPML estimation in a second stage, in addition to the original endogenous variables. The coefficients on the endogenous variables are now consistently identified, and the coefficients on the residuals provide a valid test for endogeneity of *log(PM)* and *log(wage)*. Inference is conducted using a panel bootstrap by repeating (1) and (2) on randomly drawn samples.

[29] $log(L_{od})$ as the explanatory variable is equivalent to $log(\pi_{od})$ because $\overline{L}_o$ is absorbed by the origin fixed effect.

7.2.1 PPML estimation results

Table 3: Estimates for income and pollution-as-disamenity elasticities of migration

	Poisson	First Stage: PM	First Stage: Wage	Poisson w/ CF
	L_{od}	$log(pm)$	$log(wage)$	L_{od}
	(1)	(2)	(3)	(4)
log(pm) [$\eta\lambda$]	-0.45			-1.1
	(0.23)			(0.38)
log(wage) [η]	1.59			4.36
	(0.22)			(0.66)
pm IV		0.36	-0.11	
		(0.02)	(0.02)	
wage IV		-0.03	0.22	
		(0.02)	(0.02)	
Residual log(pm)				0.82
				(0.4)
Residual log(wage)				-3.41
				(0.74)
Estimation Method	PPML	OLS	OLS	PPML
Observations	360000	360000	360000	360000
First Stage F-stat		83	30.8	
Origin FE and migration costs	Y	Y	Y	Y

Notes: PPML estimation results for migration elasticities using pairwise migration data between 2001-2010. Dependent variables are listed in rows. Column (1) is analogous to an OLS estimate, columns (2) and (3) are analogous to a first stage, and column (4) is analogous to a 2SLS estimation. All estimations control for ihs(distance), a language indicator, a state indicator and an origin fixed effect. CF in column (4) refers to Control Function, where the residuals of column (2) and (3) are included in the regression that estimates column (1) to give column (4). The sample consists of 600 districts. Clustered standard errors are reported in column (1) whereas cluster-bootstrapped errors in columns (2)-(4).

Table 3 presents the results from estimation of equation 18 using PPML. Column 1 presents results without correcting for the endogeneity of *log(wage)* and *log(PM)*. From a given origin district, a 1% higher PM2.5 level at a potential destination district is associated with 0.45% lower migration share to that district, whereas a 1% higher real wage at the destination district is associated with a 1.59% higher migration share to that district. The coefficient on *log(PM)* is statistically different from zero only at the 10% significance level. However, this coefficient suffers from omitted variable bias: districts with high amenities (that are in the residual) may attract more workers and

increase pollution, an effect that biases the coefficient toward zero since pollution and amenities are negatively correlated. The coefficient on *log(wage)* is positive as expected, but also suffers from similar endogeneity problems. Since most migration is within-state, districts that receive more migrants might actually be lower wage locations nationally, creating a negative correlation between $log(wage)$ and the residual. Further, because both $log(PM)$ and $log(wage)$ are likely include some measurement error, and this also downward biases the OLS estimate. The instruments help solve both the endogeneity and measurement error problems.

Columns 2 and 3 report results from first stage regressions on the two endogenous variables. There are no available weak instrument tests for the case of two endogenous instruments in a nonlinear model, as is the case of the PPML estimator (Lewis and Mertens 2022; Andrews et al. 2019). Nevertheless, I report separate F-stats for the two instruments. The F-statistics are 83 and 30.8 respectively for *log(PM)* and *log(wage)*, both comfortably above 10.[30]

The control function approach involves including residuals from the two first stage regressions into the PPML model with endogenous regressors. Column 4 reports results from this process. The coefficients on the residuals show that both the *log(PM)* and *log(wage)* are indeed endogenous in column 1, with a positive and negative selection effect respectively.

The income elasticity is given by the coefficient on *log(wage)*: 4.36 with a 95% CI of [3.07, 5.65]. The causal effect of a 1% increase in relative wage levels between origin and destination is to increase migration shares to the destination district by 4.36%. The point estimate for coefficient on *log(PM)* is -1.11 with a 95% CI of [-1.85, -0.37]. The causal effect of a 1% increase in relative PM2.5 levels between origin and destination is to reduce migration shares to the destination district by 1.1%. The amenity elasticity is -1.1/4.36=-0.25. The income elasticity comfortably dominates the amenity elasticity, as would be expected in a low-income setting such as India where the marginal willingness to pay for clean air has been estimated to be fairly low.[31]

7.3 Pollution elasticity of productivity (β)

Recall from equation 9 that the productivity elasticity of pollution measures how the health effects of pollution affect the aggregate marginal product of labor in a location. Motivated by equation 6, I use individual wage outcome data from the National Sample Survey (NSS) to estimate β. The NSS data provide wages measured for each individual based on their hours worked and total earnings in the past 7 days. I utilize wages for all individuals of working age for the survey year 2011-12. I construct real wages using state-level consumer price index.[32].

Equation 19 shows the Mincer formulation which is modified to include $log(PM)$ as an additional

[30] Weak instrument tests using Cragg-Donald and Kleibergen-Paap statistics are also not available for nonlinear panel models with instruments.

[31] The MWTP for clean air is measured assuming no market failures. In the presence of market failures such as in credit markets or simply abject poverty, the MWTP may be small despite clean air having amenity value.

[32] Available from the Ministry of Labor and Employment

explanatory variable. Note that I use annual pollution to understand how consistently higher pollution capitalizes into wages. This objects maps onto the parameter of interest β in eqn 9.

$$log(wage_{id}) = log(PM_{id}) + X'\beta + e_{id} \tag{19}$$

The controls X include standard variables such as age, education, rural/urban, male/female that can be found in a Mincer regression. I also include variables specific to the Indian context, viz. caste and religion, and the month of sample to control for seasonality.

The main identification challenge is that wages and pollution are jointly determined. I leverage the fire exposure instrument developed in section 5 as the solution to this challenge. Recall that the fire exposure of any district is the sum of burning in upwind districts that are within 1000 km weighted by wind and distance to the source district. It is aggregated to the calendar year to match with the annual average PM.

Table 4 shows the results. As earlier, the first stage is strong. Column 1 presents OLS estimates that are biased toward zero. The 2SLS estimate for the productivity elasticity in column 2 is -0.31. How does this compare with estimates from the literature? Chang et al. (2019) find a labor productivity elasticity of -0.09 to daily exogenous variation in PM2.5 levels for indoor pear-packers at a factory in the US while Graff Zivin and Neidell (2012) find an elasticity of -0.25 to daily Ozone (O_3) levels for outdoor fruit pickers at a farm in the US. These estimates are based on individual worker response to daily pollution exposure for indoor workers in manufacturing and outdoor workers in agriculture respectively. Evidence for productivity elasticities at extreme levels of pollution seen in developing countries is lacking, although Fu et al. (2021) calculate a larger elasticity of -0.44 for PM2.5 using nationally representative Chinese manufacturing data. The estimated elasticity of -0.31 is in the middle of these estimates.

Table 4: The effect of pollution on wages (ψ)

Dependent variable:	log(Real wage)		log(PM2.5)
	OLS	2SLS	First stage
	(1)	(2)	(3)
log(PM2.5)	-0.06	-0.31	
	(0.03)	(0.07)	
log(Burning exposure)			0.29
			(0.02)
Month FE	Yes	Yes	Yes
Observations	61,609	61,609	61,609
KP F-stat		8,616.04	
Number of districts	521	521	521

Notes: This table reports results from Mincer regressions of individual $log(wage)$ on a suite of determinants including age, education, urban dummy, female dummy, caste, religion and month of sample fixed effects. Table includes $log(PM)$ in the standard Mincer regression. Column (1) reports OLS estimates on $log(PM)$ and column (2) reports the 2SLS estimate using the fire/burning exposure instrument. Column (3) displays the first stage. Standard errors are clustered at the district level to account for within-district correlation in wages.

7.4 Output elasticity of pollution (ψ)

The output elasticity of pollution ψ governs how much pollution is generated from greater economic activity. Air pollution is a classic externality of production; an increase in industrial output is accompanied by higher emissions, commuting to work generates transport-driven pollution etc. As workers reallocate across locations, economic activity and pollution also reallocates with them. The equilibrium spatial allocation of production and pollution is jointly determined. This is a crucial mechanism through which pollution control affects aggregate productivity.

In order to estimate this elasticity, I rely on equation 11 that models pollution as a function of labor.[33] In section 7.1, we identified the component of pollution Z_d that is driven by upwind crop burning activity. Using these estimates, I calculate and store the residual of Z_d that is not explained by the smoke dispersion model, using the 10-year average emissions from equation 15 and expected

[33] Since output is linear in labor the labor elasticity of pollution is equivalent to the output elasticity.

value of distance-penalized wind from 16. This residual pollution Z_d^{res} is a power law function of labor in d.

$$Z_d^{res} \equiv \frac{Z_d}{\widetilde{\Omega}_{od}} = \overline{Z}_d \, L_d^{\psi}$$

Taking logs on this equation we get the following estimating equation

$$log(Z_d^{res}) = \psi log(L_d) + \epsilon_d^1 \tag{20}$$

where ϵ_d^1 contains $log(\overline{Z}_d)$.

A regression of the residual pollution on labor to identify ψ would lead to biased estimates since determinants of labor L_d are likely to be correlated with the error term ϵ_d^1. For example, unobserved productivity of a location is part of the residual, more productive places are more polluted, and they also attract more workers. A solution to resolve this endogeneity concern would be an instrument that moves labor supply without directly affecting the determinants of productivity and pollution. I exploit policy-driven variation in labor supply that comes from heterogeneous implementation of the National Rural Employment Guarantee Act (NREGA) that began in 2006.

Imbert and Papp (2020a) show that migration from districts that implemented the program decreased relative to those that did not implement it, because workers were able to access higher wages at the origin. For each district d, I construct a measure of exposure to NREGA implementation in other districts o by weighting the average number of person-days reported in the NREGA MIS system by the historical share of district o in migration to district d.[34] Equation 21 formalizes this instrument

$$\text{nregaIV}_d = \sum_{j \ni S(j) \neq S(d)} (\text{historical migration share}_{j-d}) * \text{ Avg NREGA person-days}_j \tag{21}$$

Identification assumption The identifying assumption is that the exposure of district d to NREGA provision in other districts o affects pollution in d only through its impact on labor L_d. Since NREGA is a poverty relief program, its provision is determined by negative productivity shocks such as heatwaves, droughts or floods. However, provision is also known to be mediated by political ideological of the ruling party at the state-level where provision decisions are made (Khera 2011). This creates variation in NREGA exposure across districts that is useful in identifying ψ. However, the identifying assumption would be violated if NREGA provision in o were correlated with determinants of productivity in d, for e.g., if negative productivity shocks were correlated in o and d. In order to allay this concern, I limit the set of origin districts to those that

[34] The NREGA MIS data were shared by Clement Imbert. I take the average of the number person-days between 2006-2010 to construct the instrument

fall in other states. This helps resolve this concern in two ways: (1) avoids correlation between productivity shocks in o and d due to geographic proximity (2) avoids correlation in NREGA provision and other policies that arise at the state-level.

Table 5 shows the results. Column (1) shows the OLS results which are biased. Column (3) shows the first stage: as expected, higher provision of NREGA in districts of other states that contributed migrants historically to district d reduces worker count in d. Using this as the instrument in column (2) reveals an output elasticity of pollution value of 0.69. The KP f-stat is 25.4, indicating a strong instrument. Figure A.2 shows the relationship between worker count and the NREGA exposure instrument. We see an extreme outlier dot at the extreme top-right corner, potentially having strong leverage.[35] In order to resolve this issue, I winsorize the instrument at the 1%/99%. Without performing this winsorization, the KP F-stat is 17 and the estimate is 0.76 (se=0.24), not significantly different from the estimates in table 5.

How does the estimate of 0.69 compare with estimates from the literature? Very few estimates of the output elasticity of pollution are available for developing countries in the literature. A related estimate comes from Fu et al. (2021) who calculate a larger elasticity of 1.43 for PM2.5 using firm-level Chinese manufacturing data. Since manufacturing is a much more important contributor to the Chinese economy (and to pollution) compared to the Indian economy, a smaller elasticity makes sense.

[35] This dot is Delhi. It has a large number of workers but also received migrants from poorer districts in other Northern states that had higher NREGA provision.

Table 5: The output elasticity of pollution (ψ)

Dependent variable:	log(Residual PM2.5)		log(Worker count)
	OLS	2SLS	First stage
	(1)	(2)	(3)
log(Worker count)	-0.17	0.69	
	(0.03)	(0.20)	
log(NREGA exposure)			-0.19
			(0.04)
Observations	615	615	615
KP F-stat		25.40	

Notes: Estimates of the output elasticity of pollution. Residual PM2.5 is calculated as the structural residual of equation 11 on annual PM2.5 for 2010, after accounting for predicted fire-driven PM2.5 ($g(\Omega)$) using the pollution dispersion model. Dependent variables are listed in rows. Since output is linear in labor, the explanatory variable is *log(Worker count)*. Column (1) shows results from an OLS estimate, column (2) shows the 2SLS estimate using the NREGA exposure instrument (see text for details). All regressions include an intercept that is not reported. Robust standard errors are reported in parentheses.

Table 6 summarizes the model parameters that are estimated in this paper.

8 Results

I now return to the question asked at the beginning of this paper. How do productivity gains from geographically targeted pollution control depend on labor reallocation? In this section, I answer this question through model counterfactuals that simulate our two pollution control scenarios. All other factors such as trade costs or preferences are left unchanged in the model counterfactuals. But in order to answer these questions through the quantitative model, I must take a stance on the the general equilibrium parameters that are not estimated in this paper. I source estimates from the literature for the main estimates and conduct robustness of the findings to alternative choices.

The agglomeration elasticity (ϕ) has been estimated by multiple studies before, with reviews in Rosenthal and Strange (2004) and Combes and Gobillon (2015). Estimates for the developed world seem to have converged on values between 0.01 and 0.02, but estimates of ϕ for developing countries are larger in magnitude and fewer in number. I benchmark ϕ to the value of 0.076 estimated by Chauvin et al. (2016) using wage data for Indian districts. They rely on historical population

density as the instrument for current density as is standard in this literature. I also conduct robustness around other choices. A higher value of ϕ would imply larger gains from reallocation toward denser cities.

The congestion elasticity (ψ) controls whether cities lose some of their amenity value when population rises. There are two principal sources of such congestion: a pure amenity value ψ_b that arises from competition for public goods such as parks, and an endogenous source arising from increases in housing rental rates ψ_r.[36] There are comparatively few estimates for these elasticities in the literature. The pure amenity component ψ_b is 0 in the US data according to Albouy (2008) whereas Combes and Gobillon (2015) estimate a value of -0.04. But, as Imbert and Papp (2020b) note, workers are willing to lose up to 35% of income by not migrating to cities due to urban disamenities that workers are unable to avoid; therefore the pure amenity component may be larger in India. As for the housing price elasticity, Bryan and Morten (2019) estimate a value of -0.08 for Indonesia. Given the prominence of informal housing in the developing world, estimates from Indonesia are a better fit for the Indian context. I set the congestion elasticity to -0.1 to account for the potentially larger pure dis-amenity value of congestion in India, and conduct robustness to other choices. A larger magnitude of the congestion elasticity would reduce the incentive of workers to migrate to cities, reducing productivity gains from abatement.

8.1 Solving the model

I implement the exact hat algebra method of Dekle et al. (2007) to solve for counterfactual *changes*. Solving for counterfactual changes eases calibration by eliminating many fixed components of the model such as productivities $\overline{A}$, amenities $\overline{B}$ and fixed pollution $\overline{Z}$ and fixed migration costs M. The second advantage of this method is that only changes in migration costs will affect the results, and it is therefore robust to any bias in the measurement of migration frictions that remains constant in the counterfactual. It is also robust to asymmetric migration costs $exp(-m_{od}) \neq exp(-m_{do})$, which could be the case if costs to leave more attractive origins are systematically higher.

Section 12.6 derives the system of equations that characterize the counterfactual equilibrium. Solving this system requires us to specify the parameters of the model ($g(.), \eta, \lambda, \beta, \phi, \psi, \gamma$), the initial values for π_{od} and $\overline{L}_o$, and the quantity κ_{od} that measures the importance of origin o for migration to d. See equation 27 for where κ comes in. Table 6 summarizes the model parameters.

[36] These congestion components are on top of the dis-amenity value of the pollution caused by an increase in the number of workers in a given location. This particular congestion elasticity is given by $\lambda\psi$.

Table 6: Model parameters and initial values

Parameter	Description	Value
$g(.)$	PM2.5 from smoke exposure	$\propto \log(\Omega_d)$
λ	Pollution-disamenity elasticity of migration	-0.25
η	Income elasticity of migration	4.36
β	Pollution elasticity of productivity	-0.31
ψ	Output elasticity of pollution	0.69
ϕ	Agglomeration elasticity	0.076
γ	Congestion elasticity	-0.1
π_{od}	Baseline pairwise migration shares	Migration data
κ_{od}	Baseline importance of migration from o for d	Migration data

Notes: Table reports quantitative model parameter values or data sources to calculate model quantities. Agglomeration elasticity (ϕ) and congestion elasticity (γ) are calibrated using estimates from literature. All other parameters and quantities are estimated. Please see text for detail on estimation or calibration

8.2 Productivity gains from targeted pollution control scenarios

Before delving into the results, let us recall the two pollution control scenarios The *Rural* scenario imposes a 10% reduction in emissions from burning in the states of Punjab and Haryana. This also lowers pollution in distant, downwind locations hundreds of kilometers away (McDuffie et al. 2021; Singh et al. 2021). I construct a pollution dispersion model to estimate spatially explicit benefits of controling crop burning in these two states. On the other hand, the *Urban* scenario reduces pollution almost exclusively within the 10 largest Indian cities through the control of localized sources such as vehicular emissions. These emissions are known to decay rapidly-even within a few hundred meters-to background levels (Liu et al. 2019). I provide evidence consistent with the literature in section 12.2. Since most vehicle-kilometers are driven on roads within the city, reducing these emissions will solely improve local air quality.

Scenarios are comparable on health benefits In order to make these scenarios comparable, I hold fixed the *population-exposure reduction* between them. Exposure reductions are used to compute health benefits from pollution control scenarios by national environmental agencies such as the US EPA. The standard methodology underlying these calculations relates the change in pollution exposure per person to changes in morbidity and mortality risk (for eg. the risk of death or disability from a stroke) using published epidemiological dose-response functions. Since I hold the total population-exposure fixed between these two scenarios, this methodology would score

these scenarios similarly on total health benefits.[37]

The normalization itself is done by calculating the total population x change in pollution (population-exposure) for the *Rural* scenario, while holding the spatial allocation of labor fixed. I then calculate the percentage change in pollution that when applied equally to the 10 largest cities would produce the same population-exposure reduction, keeping the population in those cities fixed. A uniform 3% reduction in PM2.5 levels in the 10 largest cities produces a similar population-exposure reduction to a 10% reduction in burning emissions in Punjab and Haryana. For every counterfactual *Rural* scenario I analyse, I re-estimate the uniform percentage reduction in PM2.5 for the 10 largest cities that would produce an equivalent population-exposure reduction. Figure 6 shows the change in PM2.5 levels across districts due to either scenario, with the 10 largest cities highlighted.

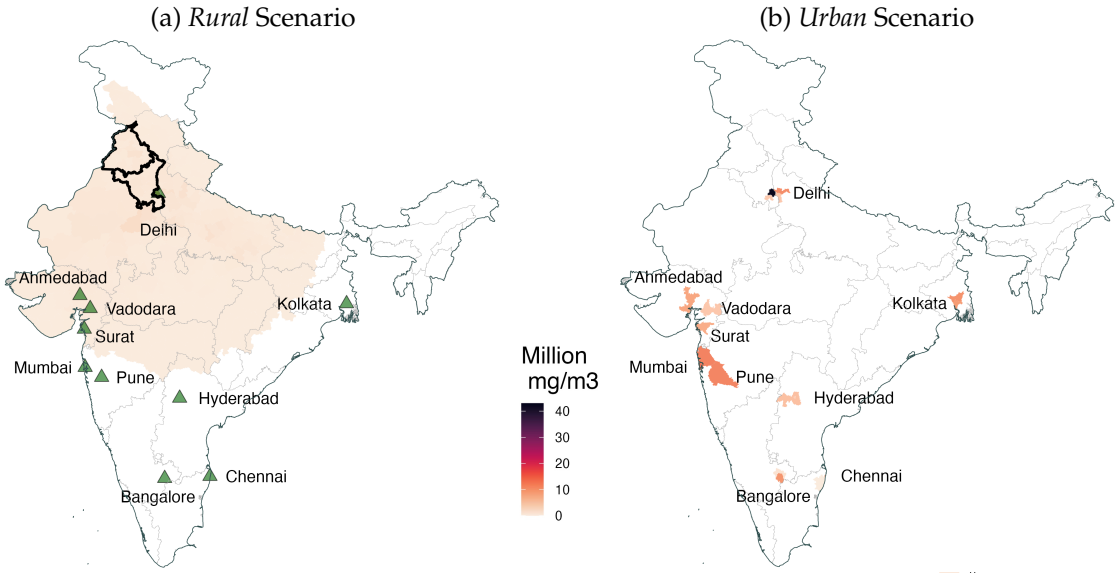


Figure 6: Population-exposure reduction from pollution control scenarios. Panel (a) maps the absolute value of the change in pollution x baseline population for the *Rural* scenario that imposes a 10% emissions reduction from crop burning in the states of Punjab and Haryana (highlighted in black). The pollution dispersion model is used to estimate change in pollution. The total population-exposure reduction from this scenario is calculated as the sum of change in pollution x population for all districts. Panel (b) maps the absolute value of change in pollution x population when the same population-exposure reduction is achieved through a uniform percentage reduction in the 10 largest cities of India (highlighted here).

8.2.1 Gains without labor reallocation

We start by analyzing how productivity gains from these two health-equivalent scenarios depend on the aggregate marginal product of labor in the places that are cleaned up. Here we hold the

[37] If the dose-response function is linear, these benefits will be identical. For some diseases, this function may be concave in pollution exposure. A concave dose-response would lead to larger health gains for places with lower baseline pollution; these are usually rural places. It is rare to find any convex dose-response functions, although ongoing epidemiological research may find one.

spatial allocation of labor fixed in the model counterfactuals. Column (1) of table 7 displays these results. The first row reports gains in national income due to the *Rural* scenario while the second row reports the same for the *Urban* scenario. A comparison of the two rows in column (1) shows GDP gains of 0.089% and 0.254% from the *Rural* and *Urban* scenarios respectively. Gains from the *Urban* scenario are almost 3 times larger compared to the *Rural* scenario when accounting only for the pre-existing differences in productivity of locations where pollution was reduced.

These results demonstrate a substantial difference in productivity gains even without accounting for labor reallocation. What they show is that since health is a complement to other factors of production that are more abundant in cities, reducing urban pollution yields larger economic gains. In other words, a given improvement in health produces larger productivity gains when realized in cities.

Table 7: Productivity gains from targeted pollution control

Policy scenario	Gain w/o reallocation	Gain from reallocation	Total gain
	(1)	(2)	(3)
Rural (Burning control in north India)	0.089%	0.023%	0.112%
Urban (emissions control in largest cities)	0.254%	0.427%	0.681%
Ratio (*Urban*/*Rural*)	2.85	18.56	6.08

Notes: Productivity gains under the *Rural* and *Urban* scenarios estimated using quantitative model counterfactuals. The *Rural* scenario imposes a 10% reduction in emissions from crop burning in north India whereas the *Urban* scenario achieves the same population-exposure change but through a uniform percentage reduction in pollution in the 10 largest cities. The numbers in rows 1 and 2 of this table are a percentage of national GDP in 2011 whereas the last row is a ratio. Column (1) reports gains without labor reallocation, reflecting the value of health improvements. Columns (2) and (3) allow spatial labor allocation to adjust across space. Column (2) reports additional gains that arise purely from labor reallocation while column (3) reports total gains in general equilibrium. The last row reports the ratio of gains for each column.

8.2.2 Gains with labor reallocation

Next, I analyze how aggregate productivity gains from these two scenarios depend on the spatial reallocation of labor that they induce. This reallocation occurs because changes in pollution levels change relative productivities and amenities across locations, causing some marginal workers to receive higher utility from moving to a now cleaner location. The strength of the labor reallocation mechanism is governed by the two migration elasticites that were estimated earlier.

Column (2) of table 7 displays income gains arising purely from spatial reallocation of labor whereas column (3) shows total gains from both the place-based productivity and labor reallocation mechanisms. From column (2), we see that the absolute gains from labor reallocation are

vastly different. Reallocation produces relatively small additional gains of 0.023% for the *Rural* scenario but a much larger 0.427% for the *Urban* scenario. Gains from reallocation are almost 19 times larger for the *Urban* scenario.

A comparison of the two rows in column (3) shows total GDP gains of 0.112% and 0.681% from the *Rural* and *Urban* scenarios respectively; total gains from the *Urban* scenario are ~6 times when accounting for the reallocation of labor. Reallocation accounts for ~63% of the total gains in the *Urban* scenario but only about 21% for the *Rural* scenario.

What explains these starkly different gains from labor reallocation? The classic literature on labor reallocation stresses the importance of moving workers to relatively more productive areas as the source of gains in average productivity for the economy (Lewis. 1954; Gollin 2014). Both of these scenarios induce reallocation across space by changing relative productivities and amenities. But the key difference is *where* labor reallocates. The *Rural* scenario reduces pollution predominantly in relatively less productive rural areas, whereas the *Urban* scenario reduces pollution in the most productive cities of India. Workers reallocate to much more productive cities when they are cleaned up, and this also amplifies existing agglomeration economies.

8.2.3 Role of geography and interaction with migration

The geography of the emissions source targeted by regulation is important in determining productivity gains. Here we see that targeting localized sources that largely increase pollution in cities produce larger productivity gains, both with and without labor reallocation. However, it is the interaction of this geography with labor reallocation that is critical to productivity gains. The concentrated reduction in pollution from the *Urban* scenario redirects workers to cities, unlocking much larger additional productivity growth of 0.427% compared to only 0.023% in the *Rural* scenario, an almost 19x difference. This tells us that understanding where pollution comes from is extremely important to quantifying productivity effects when accounting for migration in general equilibrium. Localized sources within cities create larger productivity wedges by directing workers away from productive cities, and the size of those migration-driven losses is far larger than what we would anticipate based on partial equilibrium health-based analysis.

8.2.4 Benefit-cost ratio for Rural scenario

While this analysis provides us with productivity gains for two policy scenarios that would be similarly rated on health grounds, a natural question is how these gains compare to abatement costs. I am able to estimate some back-of-the-envelope benefit-cost ratios for the *Rural* scenario using abatement cost estimates from Jack et al. (2025) who pay farmers not to burn crop residue in the state of Punjab in 2019. They estimate that an upfront payment of INR 2700-4050 avoids the burning of 10% of acreage under rice. I take the upper end number of INR 4050 to be conservative, and calculate the total cost involved in scaling this up to avoid 10% of emissions from rice crop

burning in Punjab and Haryana. I find a large benefit-cost ratio of **31:1**, which is purely based on estimated productivity gains. A detailed calculation is provided in appendix section 12.4.

This ratio would be larger if the health benefits were valued using techniques such as the Value of Statistical Life (VSL) and Quality-adjusted Life Years (QALY). But note that the health benefit valuation would produce similar or identical monetized benefits for both scenarios, whereas the productivity benefits under the *Urban* scenario are ~6x larger. Since there are few studies that produce marginal abatement cost (MAC) estimates for localized sources such as vehicles or cookstoves in India, I am unable to conduct a similar exercise for the *Urban* scenario. However, the literature suggests that pollution regulation is rarely binding in developing country contexts like India (Greenstone and Jack 2015), and therefore relatively low-cost ways to reduce pollution from vehicles and cookstoves should exist. In this vein, promising ways to reduce emissions from vehicular sources include scrapping the oldest and most polluting vehicles, improving testing under the mandatory Pollution Under Control (PUC) program or implementing congestion pricing.[38]

8.2.5 Robustness to alternative agglomeration and congestion elasticities

The agglomeration and congestion elasticities are the two parameters in the model that are calibrated using existing literature. All other parameters are identified using data from India and rigorous techniques of causal inference. As such, we might be worried about mis-specifying these two important elasticities. Agglomeration economies increase the benefits of scale: more workers in a given area increase productivity (Combes and Gobillon 2015). On the other hand, more workers also have negative consequences such as higher competition for housing, greater potential for communicable diseases to spread, or a simple crowding-out of rivaled public goods such as parks. These are all examples of congestion, and this can counteract the benefit of agglomeration economies. The agglomeration and congestion elasticities govern these mechanisms respectively. The relative importance of these elasticities could make a large difference to the ratio of gains from pollution control. Therefore, I conduct robustness around plausible values for these elasticities in this section.

Earlier in this section, I discussed existing estimates from literature. The agglomeration elasticity, in particular, is estiamted using data on Indian districts, so we may have more confidence in the estimate of 0.076 from Chauvin et al. (2016). However, there are essentially no estimates for the congestion elasticity from India. We set a congestion elasticity of -0.1 for the main estimates, which is larger than the -0.08 estimated for Indonesia by Bryan and Morten (2019). For both of these parameters, I consider combinations of absolute values in [0.01, 0.2], increasing by steps of 0.01. The congestion elasticity γ is negative while the agglomeration elasticity is positive. Figure 7 displays the ratio of gains from migration and total gains in GE from model counterfactuals for various

[38] This article explains why older vehicles might be responsible for most of the emissions load, and this article argues that pre-2005 vehicles were responsible for 70% of the emissions load in multiple cities in 2014.

combinations of these parameters through a heat map. I also highlight gains for the default values of the parameters in the heat map.

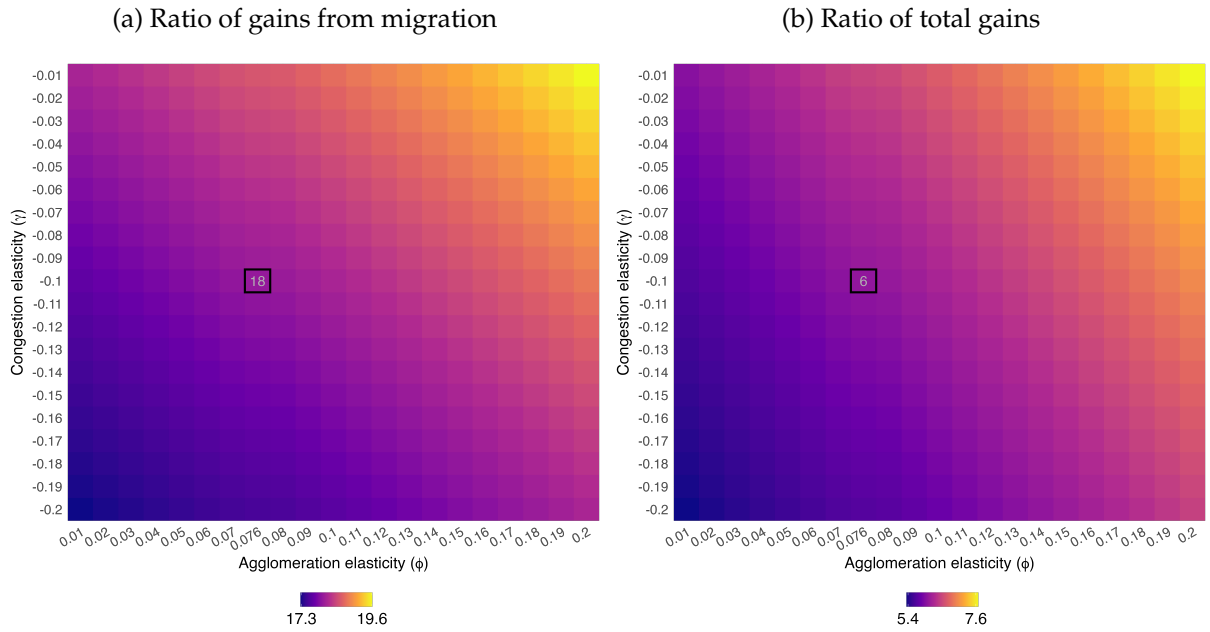


Figure 7: Robustness to congestion and agglomeration elasticities. This heatmap documents the change in the ratios reported in table 7 (last row labelled *Ratio*) using alternative combinations of the congestion and agglomeration elasticities. Panel (a) corresponds to column (2) in table 7 while panel (b) corresponds to column (3). The numbers within the boxes in black in the heatmap shows the ratios with the baseline values of the congestion and agglomeration elasticities.

The main takeaway from figure 7 is that the ratio of gains is stable for various parameter combinations. The largest changes from the baseline gains arise for extreme combinations of parameter values. For example, the largest ratio of gains arise when the agglomeration elasticity is large (0.2) and the congestion elasticity is small (-0.01). But even in this case, the ratio of gains from migration is ~19.6 while the ratio of total gains is ~7.6, compared to ratios of 18 and 6 under the default values. On the other extreme, when the agglomeration elasticity is small (0.01) and the congestion elasticity is large (-0.2), the ratio of gains shrinks to 17.3 and 5.4 respectively. Again, these ratios are not far from the baseline ratios. These results give us confidence that mis-specification of the agglomeration and congestion elasticities does not drive the large difference in gains for the *Urban* scenario relative to the *Rural* scenario that I document in this paper.

9 Conclusion

This paper studies how productivity gains from targeted pollution control depend on the worker migration response in general equilibrium. Targeted pollution control improves air quality, health, wages and amenities in some locations. In general equilibrium, this creates incentives for labor reallocation because some workers can now earn more utility by moving to cleaner locations. If

workers move to productive cities as a result, there will be larger gains in aggregate productivity. The geography of emissions sources (the location of the source and its long-distance pollution dispersion tendency) determines which places are cleaned up, and therefore how important worker migration is for productivity gains. In order to understand these interactions, I develop a spatial equilibrium framework of worker location choice and embed a pollution dispersion model within this framework.

I study two pollution control scenarios that are based on the wider policy debate on how to clean the urban air in the world's most polluted country, India. These scenarios include the control of crop residue burning in rural areas of north India that disperse smoke over long distances ('*Rural* scenario'), and the control of localized emissions (from sources such as vehicles and cookstoves) in the 10 largest cities that largely affect local air quality ('*Urban* scenario'). I develop a pollution dispersion model for smoke and estimate a pollution-smoke exposure relationship in order to quantify the pollution reduction from the *Rural* scenario. I hold fixed the total population exposure reduction from these two scenarios so that they would be rated similar on health impacts since standard regulatory guidance uses dose-response functions for pollution that are mostly linear. But they reduce pollution in very different places: both rural and urban areas for the *Rural* scenario, and largely urban areas for the *Urban* scenarios.

I then show that increases in district pollution reduce worker in-migration in India, using exogenous variation in pollution driven by shifts in upwind burning activity. This spatial labor reallocation may be caused by lower expected wages due to pollution or the disamenity value of pollution. In general equilibrium, the distribution of pollution and workers is determined jointly through sorting across districts. I take the labor supply equation across districts predicted by the quantitative model to data on pairwise migration across Indian districts to estimate the income and pollution-disamenity elasticities. Utilizing instrumental variables to correct for endogeneity, I find an income elasticity of 4.36 and an amenity elasticity of -0.25.

Then, using the quantitative model, I conduct counterfactuals to understand productivity gains from the two policy scenarios of interest. Productivity gains are almost 6 times larger for the *Urban* scenario, both because health improvements for workers in cities generate larger economic value, and because labor reallocation to productive cities reinforces this effect. Productivity gains purely from labor reallocation are 18 times larger for the *Urban* scenario in absolute terms. Reallocation accounts for only ~20% of the total income gains for the *Rural* scenario, but since the reallocation induced by that policy is toward less productive rural areas. On the other hand, under the *Urban* scenario, reallocation accounts for 63% of the total gains, because it induces reallocation to the largest cities and amplifies existing agglomeration economies.

Despite smaller gains from reallocation for the *Rural* scenarios that reduces emissions by 10% in northwestern India, I calculate a benefit-cost ratio of 31:1 using estimates of the marginal abatement cost per acre not burnt from Jack et al. (2025), and from my own estimates of how much crop area needs to be left unburnt to achieve 10% lower emissions. This suggests the failure of a Coasian

bargaining process wherein other states, particularly in North India, could compensate farmers in Punjab and Haryana for a costly reduction in fires. This failure may be down to lack of regulation of the pollution externality at the appropriate level (Banzhaf and Chupp 2012; Lipscomb and Mobarak 2017; Kahn et al. 2015), or to low levels of economic development where credit constraints and weak regulatory capacity are common (Jayachandran 2022; Besley and Persson 2009; Jack et al. 2025).

Finally, this paper underlines the importance of accounting for the interactions between worker migration and geography of emissions for productivity gains from pollution control. Emissions sources that increase pollution largely within cities lead to larger productivity losses because it keeps workers away from productive cities compared to upwind sources that also increases pollution in upwind rural areas. I demonstrate that it is important to understand these gains in general equilibrium since the size of migration-driven losses is far larger than the partial equilibrium health-based analysis that is common in the literature. While I focus on two mechanisms here, there may be other long-run general equilibrium adjustments that change the gains from pollution control. I leave those for future study.

10 Appendix Figures

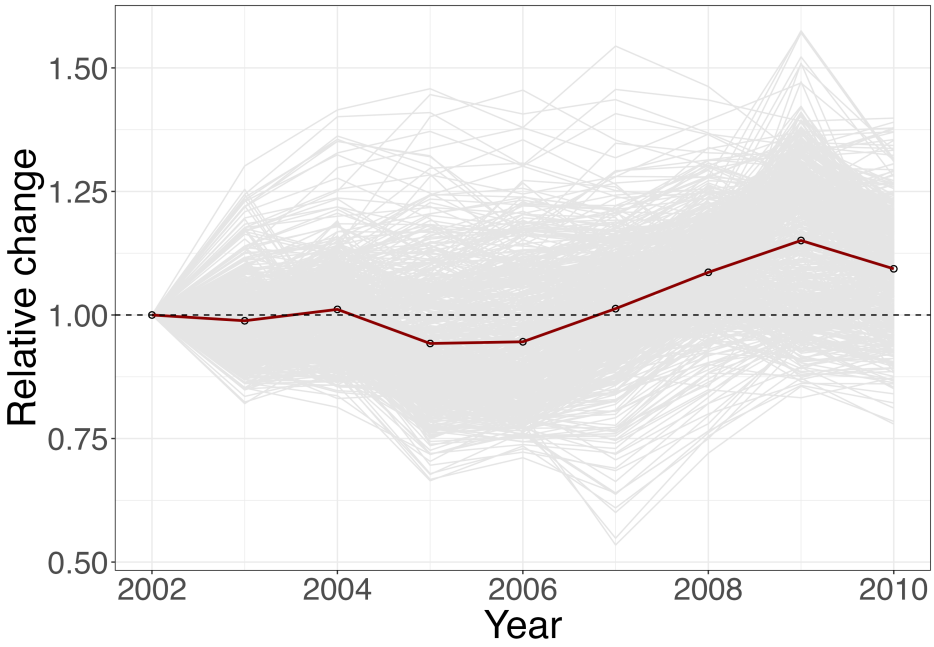


Figure A.1: Change in district PM2.5 during 2000s (the value for the base year 2001 is normalized to 1). The sphagetti plot shows growth rates for individual districts whereas the dark red line shows the average growth rates across all districts.

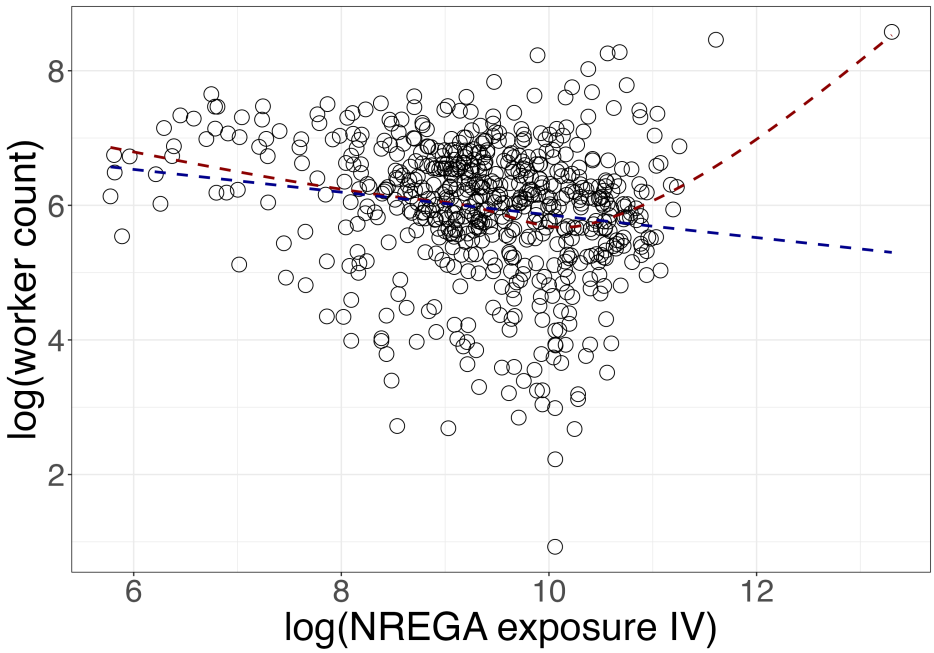


Figure A.2: First stage relationship between worker count and NREGA exposure IV. Each dot respresents a district. Worker count is calculated using census 2011. Refer to main text for construction of NREGA exposure instrument. Red and blue lines show nonparametric and linear fits respectively. The outlier at the top right is Delhi. The actual first stage is estimated after winsorizing the NREGA exposure IV at the 1% level.

11 Appendix Tables

Table A.1: Summary statistics

Variable	N	Mean	SD	Min	Max
Panel A: Effect of pollution on migration (change between 2006-2010 and 2001-2005)					
Δ ihs(Net-migration)	624	0	4.04	-13.92	15.15
Δ log(In-migration)	626	1.25	0.34	0.68	2.57
Δ log(Out-migration)	624	1.15	0.24	0.69	1.95
Δ log(PM2.5)	624	0.07	0.07	-0.17	0.22
Δ log(Fire exposure)	624	0.39	0.16	0.04	0.73
Panel B: Pollution dispersion model for smoke exposure (2002-2016) *(Reporting statistics from estimation results per district, or predictions for 2016)*					
Count of districts	617	451	111	155	601
Within R-squared	617	0.21	0.1	0	0.43
Calibrated τ_d	617	0.89	0.57	0	2.01
log(Fire exposure) in 2016	617	7.14	1.02	4.27	10.74
Predicted PM2.5 in 2016	617	6.51	4.28	0	15.9
Panel C: Estimation of migration elasticies η and λ (2010)					
Count of migrants	360000	40.4	725.76	0	127769
log(Real wage)	360000	5.25	0.4	4.22	7.42
log(Wage IV)	360000	3.35	0.9	0.32	9.71
log(PM2.5)	360000	3.92	0.45	2.49	4.81
log(Fire exposure)	360000	5.96	0.72	3.51	7.61
Panel D: Estimation of pollution elasticity of productivity β (2010)					
log(Real wage)	61609	3.5	0.85	-0.33	9.69
log(PM2.5)	61609	3.97	0.41	3.05	4.95
log(Fire exposure)	61609	5.93	0.73	3.97	7.68
Panel E: Estimation of output elasticity of pollution ψ (2010)					
log(Residual PM2.5)	615	2.43	0.66	1.48	5.96
log(worker count)	615	5.96	1.02	0.93	8.58
log(NREGA IV)	615	9.39	1.02	5.78	13.3

Notes: Summary statistics for various analysis samples. Panel (A) corresponds to the estimating equation 1 in section 5. Panel B corresponds to section 7.1. Each district gets a separate τ_d from panel estimates of log(PM2.5) on log(fire exposure) which are then calibrated using output from a Chemical Transport Model (see main text). The first three rows of panel B show the summary statistics for parameters from those estimation results, while the last two rows show summary statistics for the explanatory variable and the predicted fire-driven PM in 2016. Panel C corresponds to estimating equation 18 in section 7.2, panel D to equation 19 in section 7.3, and panel E to 20 in section 7.4.

Table A.2: First stage for the effect of pollution on migration

Dependent variable:	Δ log(PM2.5)		
	(1)	(2)	(3)
Δ log(Burning exposure)	0.20	0.20	0.20
	(0.02)	(0.02)	(0.02)
Region FE	Yes	Yes	Yes
Observations	625	626	625
Adjusted R^2	0.662	0.659	0.661
Second stage variable	Δ ihs(net-migration)	Δ log(in-migration)	Δ log(out-migration)

Notes: First stage estimates for regressions of first-differenced migration variables on pollution between 2002-2005 and 2006-2010. Controls include baseline migration variable measured for the 1990s, region fixed effects and average weather. Robust standard errors are reported in parentheses. The relevatn outcome variable is reported in the last row.

Table A.3: The effect of pollution on migration per capita

Dependent variable:	Δ Net-migration per 1000		Δ In-migration per 1000		Δ Out-migration per 1000	
	OLS	2SLS	OLS	2SLS	OLS	2SLS
	(1)	(2)	(3)	(4)	(5)	(6)
Δ log(PM2.5)	-3.84	-30.92	3.72	-9.99	8.27	19.90
	(6.46)	(7.87)	(6.03)	(6.40)	(1.64)	(4.19)
Region FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	548	548	549	549	548	548
KP F-stat		121.49		120.37		117.94
Mean		-0.35		4.60		4.94
SD		6.46		6.22		2.66

Notes: Estimates from regressions of first-differenced migration variables on pollution between 2002-2005 and 2006-2010. The outcome variables are migration per capita; they are winsorized at the 1% level to minimize the influence of outliers. Controls include baseline migration variable measured for the 1990s, region fixed effects and average weather. ihs refers to inverse hyperbolic sine. Robust standard errors are reported in parentheses. The mean and SD of the reported migration variable itself are reported in the last two rows.

Table A.4: The effect of pollution on migration with population weights

Dependent variable:	Δ ihs(Net-migration)		Δ log(In-migration)		Δ log(Out-migration)	
	OLS	2SLS	OLS	2SLS	OLS	2SLS
	(1)	(2)	(3)	(4)	(5)	(6)
Δ log(PM2.5)	-4.26	-25.88	-0.32	-1.69	0.41	0.75
	(4.23)	(7.91)	(0.31)	(0.46)	(0.20)	(0.39)
Region FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	548	548	549	549	548	548
KP F-stat		127.08		135.05		117.10
Mean (1000s)		-0.26		7.88		8.07
SD (1000s)		22.17		23.39		7.31

Notes: Estimates from regressions of first-differenced migration variables on pollution between 2002-2005 and 2006-2010, weighting each observation by the population of the district in 2010. The outcome variables are winsorized at the 1% level to minimize the influence of outliers. Controls include baseline migration variable measured for the 1990s, region fixed effects and average weather. ihs refers to inverse hyperbolic sine. Robust standard errors are reported in parentheses. The mean and SD of the change in the level of the relevant migration are reported in the last two rows (in 1000s of workers).

Table A.5: Correlation of the fire exposure with historical determinants of migration

Dependent variable:	% labor in ag	Literacy rate	Female labor force participation.	% irrigated land	Growth rate of night lights	Share of SC/ST
	(1)	(2)	(3)	(4)	(5)	(6)
Δ log(Burning exposure)	0.57	-0.33	0.08	-0.13	-2.05	0.16
	(0.05)	(0.04)	(0.03)	(0.02)	(0.18)	(0.06)
Region FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	549	549	549	549	544	549

Notes: Estimates from regressions of historical correlates of migration on the fire/burning exposure instrument. The outcome variables are in levels and are measured in or before 2000 (the migration sample period begins in 2001). The explanatory variable is the change in average log fire exposure measured between 2002-2005 and 2006-2010. Controls include baseline migration variable measured for the 1990s, region fixed effects and average weather. Robust standard errors are reported in parentheses.

Table A.6: The effect of pollution on migration controlling for historical determinants

Dependent variable:	Δ ihs(Net-migration)		Δ log(In-migration)		Δ log(Out-migration)	
	OLS	2SLS	OLS	2SLS	OLS	2SLS
	(1)	(2)	(3)	(4)	(5)	(6)
Δ log(PM2.5)	2.57	-2.62	-0.50	-1.53	0.10	-0.09
	(4.00)	(7.50)	(0.34)	(0.62)	(0.19)	(0.46)
Region FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	543	543	544	544	543	543
KP F-stat		133.03		130.33		139.38
Mean (1000s)		-0.17		7.80		7.92
SD (1000s)		21.82		22.89		8.90

Notes: Estimates from regressions of first-differenced migration variables on pollution between 2002-2005 and 2006-2010, controling for the level of historical correlates of migration measured before the migration sample period. The outcome variables are winsorized at the 1% level to minimize the influence of outliers. Controls include historical correlates of migration before 2000, baseline migration variable measured for the 1990s, region fixed effects and average weather. ihs refers to inverse hyperbolic sine. Robust standard errors are reported in parentheses. The mean and SD of the change in the level of the relevant migration are reported in the last two rows (in 1000s of workers).

12 Appendix

12.1 Performance of alternative specifications of pollution dispersion model

Table A.7 shows the performance of the three pollution dispersion models against a useful benchmark: PM2.5 from agricultural burning as estimated by a full-scale Chemical Transport Model provided by Washington University of St. Louis (WUSTL) in 2016 (McDuffie et al. 2021).[39] I assess the performance of predicted PM2.5 for 2016 from the three models I consider (linear-log, power law and power law in log) against these scientifically estimated PM2.5 data. I construct smoke exposure for 2016 using the meterology for that year, and then convert to PM2.5 using the estimated τ_ds. Then I run a standard linear regression of true PM2.5 against these predicted PM2.5 Table A.7 shows the results of this exercise.

Table A.7: Performance of pollution dispersion model

Dependent variable:	PM2.5 from agri burning (WUSTL)			
	Linear-log	Power law	Power law in log	Linear
	(1)	(2)	(3)	(4)
Model-predicted PM2.5	0.77	1.06	0.66	0.25
	(0.02)	(0.06)	(0.04)	(0.02)
Observations	617	617	617	617
Adjusted R^2	0.77	0.39	0.36	0.25

Notes: Table shows regression estimates of true PM2.5 from agricultural fires in 2016 (as measured by the Washigton University of St. Louis) on the predicted fire-driven PM2.5 from the pollution dispersion model with various specifications of $g(\Omega_d)$. The specification is listed in a row. Robust standard errors are reported in the parentheses. Both outcome and explanatory variables are in the same units of μg/m3.

It can be seen immediately that the linear-log model in column (1) performs best by far. The other models also capture meaningful variation in true fire-driven PM2.5, but are ultimately able to explain only between 25-39% of the variation. On the other hand the linear-log model is able to explain 77% of the true fire-driven PM2.5. Therefore, I select the linear-log model as the most appropriate choice to understand the counterfactual reduction in pollution from the control scenarios.

12.2 Pollution decay from vehicular emissions (λ, η)

The *Urban* scenario assumes a uniform reduction in pollution from localized sources within the 10 largest cities of India. But is this a realistic scenario, and targeting which emissions sources is most likely to lead to such an outcome? City-level source apportionment studies in India suggest that vehicular emissions and burning of solid fuels like charcoal in cookstoves contribute up to 50%

[39] Annual PM2.5 data for the year 2016 have been summarized over census 2011 districts and can be accessed at https://urbanemissions.info/india-air-quality/india-satpm25/

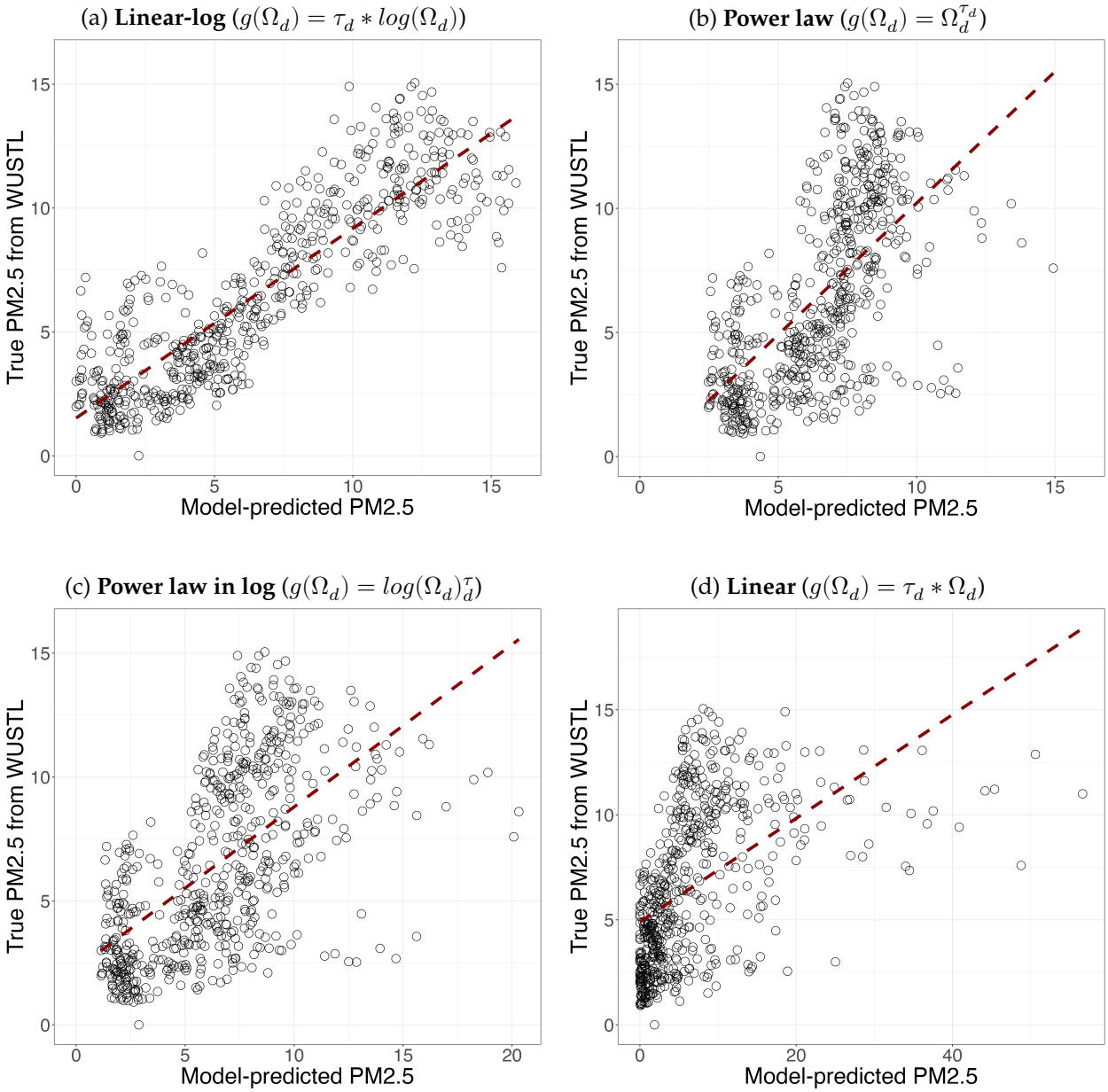


Figure A.3: Performance of pollution dispersion model. This figure plots true PM2.5 from agricultural fires in 2016 (as measured by the Washigton University of St. Louis) against the predicted fire-driven PM2.5 from the pollution dispersion model with various specifications of $g(\Omega_d)$. Each panel represents a different specification for $g(\Omega_d)$. Both x and y-variables are in the same units of μg/m3. The x-axis of panels (c) and (d) are wider than the y-axis whereas the two axes in panels (a) and (d) are equally-sized.

of PM2.5 concentrations (McDuffie et al. 2021). Reducing pollution from these sources through solutions such congestion pricing or adoption of cleaner cookstoves without changing population would affect pollution Z_d via $\overline{Z}_d$, since these reduce the emissions intensity of activity.

Various studies on the decay rates of vehicular pollutants finds that these emissions rarely travel more than a couple of kilometers (Liu et al. 2019). In this section, I ask whether similar decay patterns imply that vehicular emissions in Indian cities tend to remain localized, affecting pollution largely within the city.

I construct a dataset of hourly pollution and meteorology at monitoring stations for 8 of the 10 largest cities. These include Ahmedabad, Chennai, Delhi, Hyderabad, Bangalore, Kolkata, Mumbai and Pune. The data can be downloaded from the CPCB's website.[40] The stations were manually matched with coordinates available on the website. I merge these station-level data with shapefiles for primary roads from OpenStreetMap, as provided by UrbanEmissions.info.[41] Each of these primary roads contains multiple segments, and I construct the distance of each segment to all monitoring stations with the city. The resulting data set contains data on road segment x monitoring station x time. Apart from PM2.5 readings, the monitoring stations also gather data on wind speed and direction.

Closely following pollution decay models in the scientific literature, I specify the following model

$$PM_{cmt} = PM_{c,s(t)}^{AtRoad} \, f(\delta; d_{m,s(t)}) + PM_{ct}^{background} \tag{22}$$

where c, m, s, t index city, monitoring station, road segment and time of measurement respectively. PM refers to pollution and d refers to distance from road segment to monitoring station.

Unlike the granular pollution decay estimation in scientific studies such as in Liu et al. (2019), I am unable to collect data over time on pollution at a specific road segment and for every few meters away from that segment. Instead, I rely on variation in how emissions are directed away from a road segment by the meteorology at the time. For a given monitoring station, there might be two road segments at different angles from the station. Based on the direction in which wind is blowing and its speed, the station will best capture pollution from one of these road segments. Therefore, I rely on wind speed and direction at the time of measurement, along with distance, to determine which road segment is most important for pollution readings at a give monitor. This is why distance $d_m, s(t)$ is a time varying object; for each monitoring station m the segment that most affects pollution at time t can change along with the distance d. Specifically, I find the major road segment $s(t)$ from which the wind is blowing toward m at time t. I also weight this distance by wind speed measured at the monitoring station m. If the wind is stronger, emissions are more likely to be carried a greater distance. This approach has one advantage over the granular pollution decay exercises carried out for one segment at a time: I am able to characterize decay of pollution from major roads in 8 Indian cities using real-world data.

However, we lack measurements of PM^{AtRoad} and $PM^{background}$ because we are relying on publicly available data. To estimate equation (1) without these variables, we write down a simplified nonparametric model on binned distance $d_{m,s(t)}$ with city x hour-of-day fixed effects to absorb variation in $PM_{c,s(t)}^{AtRoad}$ and $PM_{ct}^{background}$.

[40] https://airquality.cpcb.gov.in/ccr/#/caaqm-dashboard-all/caaqm-landing
[41] https://urbanemissions.info/india-air-quality/india-ncap-airshed-osmroads

$$\text{Relative PM}_{cmt} = \sum_{b}(\delta_b * d_{m,s(t)}) + FE_{c,h(t)} + e_{cmt} \tag{23}$$

with the omitted distance bin being the one closest to the road. I transform PM_{cmt} to Relative PM_{cmt} by dividing by the average PM for this omitted bin. This allows us to interpret the coefficients as percentage changes relative to the average value within the bin nearest the road. The city x hour-of-day fixed effect controls for differences in the level of emissions and the background pollution level across cities by time of day as well. The identifying variation in pollution at station m comes from changes in wind direction and speed that trace out changes in in $d_m, s(t)$, and determine which road segment contributes most to this pollution.

Table A.8 shows summary statistics for the data on which estimation is conducted. The maximum reading of pollution the instrument can make is 1000 mg/m3, but there are only 125 instances of that. Most stations are located within 10 km of a road. Figure A.4 shows the distribution of stations by distance from road.

Table A.8: Summary statistics: road-pollution decay estimation

Variable	N	Mean	SD	Min	Max
PM2.5 (μmg/m3)	3628606	68.5	78	0.01	1000
Distance (km)	3628606	4.89	8.2	0	100

Notes: This table shows summary statistics for PM2.5 and windspeed-weighted, wind-based distance from the nearest road segment. The distance variable is computed for each hour based on the nearest road segment from which wind is blowing toward each monitoring station, weighted by the wind speed measured at the station at the time. The sample consists of hourly observations between the years 2017-2024 for 118 monitoring stations in 9 of the 10 largest cities in India.

Equation 23 estimates the average percentage change in pollution by distance bins away from major roads. Figure A.5 plots the effect of distance from road. Pollution falls by 30% of the level at the road within 5 km. The shape of the curve is close to being an exponential as most vehicular pollution studies find (Liu et al. 2019), although they model pollution within the 2-3 kilometers only.

The main takeaway from this section is that since the largest fraction of vehicle-kilometers are driven on roads within the city, the *Urban* scenario will largely lead to localized pollution reduction within these cities.

12.3 Causes of crop burning

These two states of Punjab and Haryana are characterized by a rice-wheat cultivation system. In these rice-wheat systems, rice is cultivated during the monsoon or "Kharif" season (June-November) while the wheat crop is cultivated in the winter or "Rabi" season (December-April).

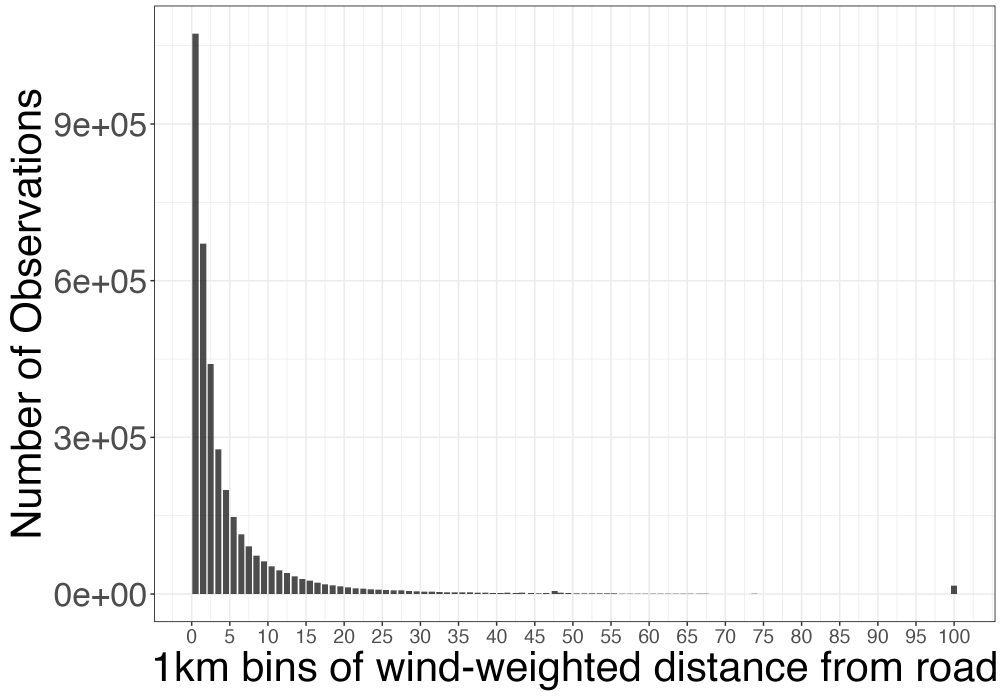


Figure A.4: Number of observations by 1km distance bins away from road

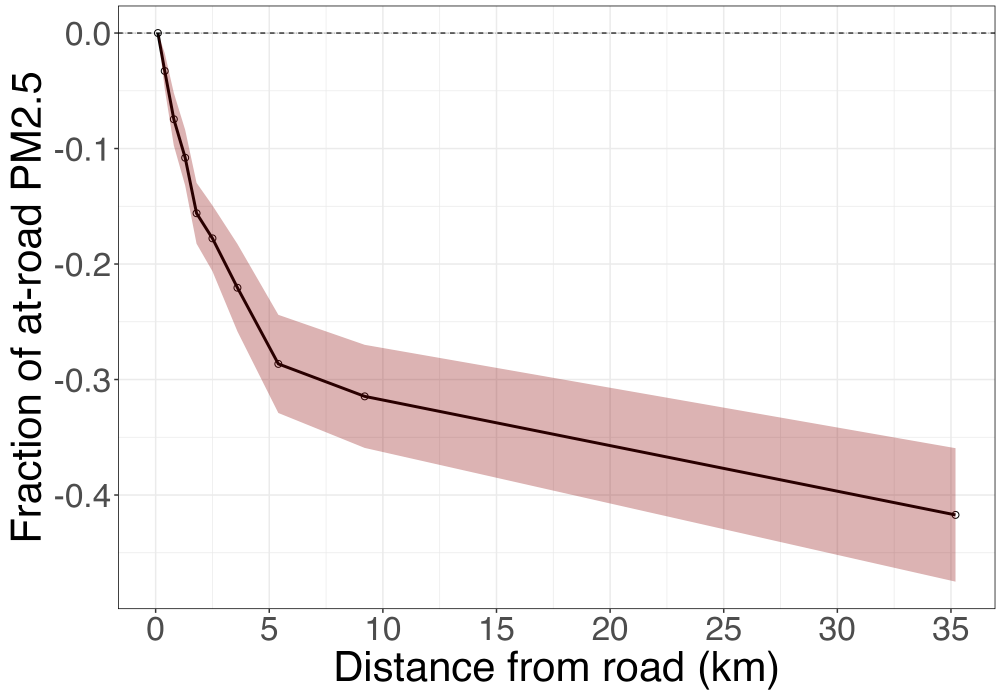


Figure A.5: Nonparametric binned regression estimates for decay in PM2.5 by distance from road. The omitted bin is the one closest to the road (<280 meters). There are 10 bins with equal number of observations within each bin. The outcome variable is normalized PM2.5 (dividing by average PM2.5 in the omitted bin) so that the coefficients on distance bins can be interpreted as percentage changes relative to pollution at the road. The mean distance within each distance bin is plotted on the x-axis. Standard errors are clustered at city x hour-of-day to account for hourly commute patterns separately for each city.)

The rice crop harvesting process leaves a residue in the field that must be removed before planting of wheat in early Rabi season. Wheat must also be planted in the first weeks of winter in order to get optimal yields. Fires are a cheap technology that can be used to remove this residue. The short duration between the rice harvest in late October and the optimal wheat planting window in early November further incentivizes farmers to burn the residue.

The rice-wheat system has its roots in the Green Revolution of the 1960s. Until then, North-Western India was a primarily wheat-growing region with little rice consumption or production locally. The advent of the Green Revolution brought with it many institutional innovations from the Indian State that increased agricultural productivity substantially across India. In the states of Punjab and Haryana, this took the form of massive subsidies for tubewells which could be used to access shallow groundwater to irrigate fields that did not have access to the pre-existing large canals systems built by the colonial British empire. This newfound access to groundwater allowed farmers to diversify their crop portfolio during the monsoon months by allowing the cultivation of water-intensive rice crop. The state of Punjab contributed less than 1% of India's rice in 1961; by the late 1990s this figure was up to 10%, even as total rice output across India also increased substantially. The use of fires to clear rice residue started in the 1990s. The earliest observations of fires from the NASA FIRMS database starting in 2002 clearly demonstrate that northwestern India already had a disproportionate share of fires in Indian agriculture.

12.4 Cost-benefit calculation for crop burning scenario

In this section, I calculate a back-of-the-envelope benefit-cost ratio for the crop burning policy, assuming that it can be achieved through a Payment for Ecosystem Services policy that would directly pay farmers not to burn. Jack et al. (2025) conduct an RCT where they show that such a policy can indeed lead to substantial reduction in burning activity. In particular, the treatment arm where they provide a portion of money upfront to alleviate credit constraints reduces burning by 10%. They also provide an abatement cost estimate of INR 2700-4050 per acre of rice planted. In this section, I use these abatement cost estimates to calculate back-of-the-envelope a benefit-cost ratio for the *Rural* scenario.

This scenario imposes a 10% reduction in emissions from crop burning in Punjab and Haryana. In order to calculate back-of-the-envelope abatement costs for the *Rural* scenario, we need to convert the 10% emissions reduction into an estimate of the crop area that is not burnt. Once we have that, we can use the per-acre marginal abatement costs to calculate total abatement costs. To back this out, I estimate the relationship between acreage under rice cultivation and the emissions from burning.

In equation 24, I model emissions E_d as a function of area under rice cultivation, where f_d^E captures how rice cultivation translates into burning. I limit this estimation to districts in Punjab and Haryana, since there is much less burning in the rest of India due to contextual and institutional differences, as described in 12.3. This allows us to estimate acres avoided using data from the same state where Jack et al. (2025) conducted their study.

$$E_d = f_d^E(RiceArea_d) \tag{24}$$

Figure A.6 documents a strong relationship between agricultural fire emissions E_d (as measured through total fire radiative power across pixels in each district) and the area under rice crop, both

from 2010. It also shows that the functional form of $f_d^E(.)$ is likely linear. I also include the linear regression estimate of 7.4 in the figure. This estimate comes from a simple linear model that includes state fixed effects to capture any level differences in technology across the two states that may also drive the rate of burning. The important point here is that the relationship is linear.

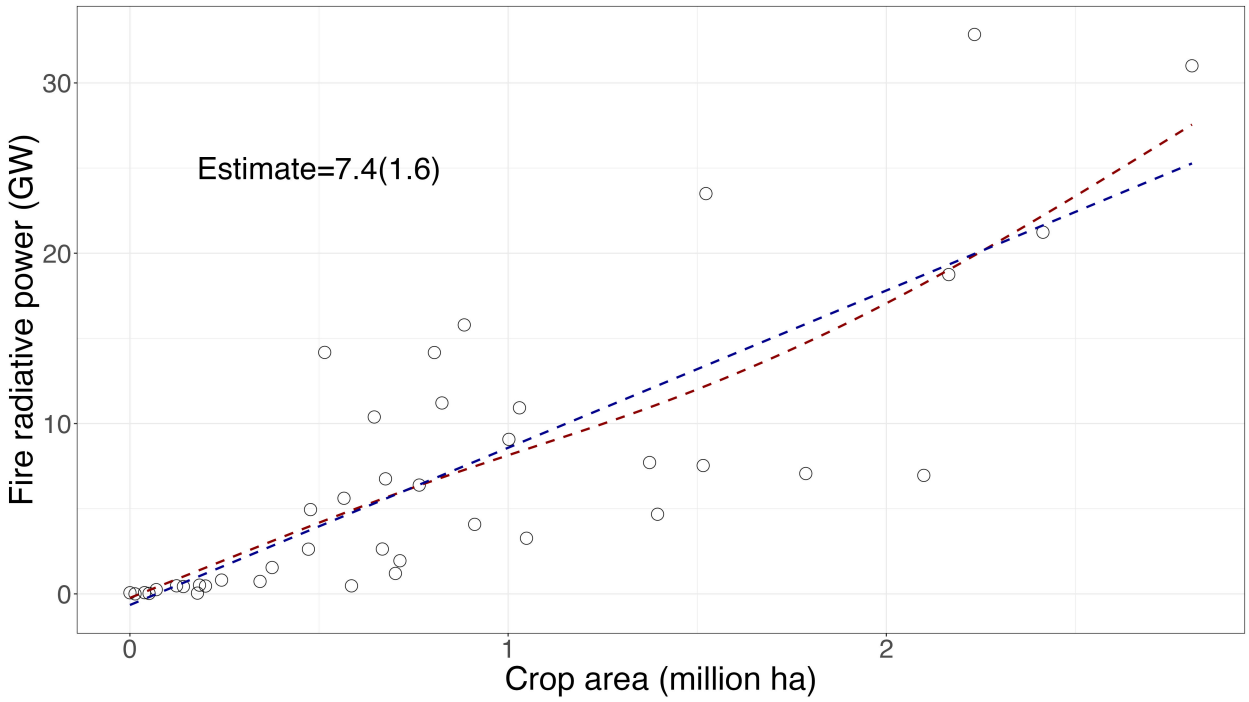


Figure A.6: Correlation between burning emissions and rice acreage in Punjab and Haryana. Each dot represents a district in Punjab or Haryana in 2010. The y-axis is the total fire radiative power of all fires within the district in 2010. The x-axis is the area under rice crop provided by ICRISAT. The lines show linear and nonparametric fits in blue and red respectively. The linear regression estimate with robust standard errors in parentheses is shown in the upper left of the figure.

Lets take the upper end abatement cost estimate of INR 4050 per acre from Jack et al. (2025) to be conservative. Since benefits from pollution control are expressed as a percentage of national GDP in 2011, an easy way to calculate the benefit-cost ratio would be to estimate total abatement costs in the same percentage of 2011 national GDP. In order to do this, I take the following steps:

1. INR 4050 in 2019 equals INR 3488 in 2015.[42]
2. Total acres under rice cultivation in 2011 was 11.12 million acres. Avoided acreage burnt using 10% = 1.112 million acres.[43]
3. Total abatement cost = INR 3488 x 1.112 x 1e6 = INR 3.879 billion (measured in 2015 INR)
4. Total national GDP in 2011 (measured in 2015 USD) = 1.62 trillion.[44]

[42] Multiply by 0.8613 = (147.8/171.6) using CPI data at https://data.worldbank.org/indicator/FP.CPI.TOTL?locations=IN)

[43] Total rice acreage of 11.12 million acres in Punjab and Haryana come from author's own calculations.

[44] From https://data.worldbank.org/indicator/NY.GDP.MKTP.KD?locations=IN

5. Convert total GDP in 2011 to 2015 INR by multiplying with 66.3329. This equals INR 107.47 trillion (measured in 2015 INR).[45]

6. Calculate total abatement cost as percentage of GDP: $\frac{3.879}{(107.47*1000)}$*100 = **0.0036%**

This provides us with a total abatement cost of 0.0036% of national GDP for the *Rural* scenario. Total benefits in GE equal 0.112%, leading to a benefit-cost ratio of $\frac{0.112}{0.0036}$ = **31:1**.

12.5 Derivation of migration shares

Each individual j born in origin district o draws an idiosyncratic utility shock ε_{jod} for living in destination d. Eq. (4) specifies the *indirect utility* of j, conditional on wages, prices and pollution, from choosing destination d as

$$V_{jod} \;=\; \underbrace{\varepsilon_{jod}}_{\text{taste shock}} \;\times\; \underbrace{B_d\, Z_d^{\lambda}}_{\text{amenity and pollution}} \;\times\; \underbrace{\frac{w_d}{P_d}}_{\text{real wage}} \;\times\; \underbrace{e^{-m_{od}}}_{\text{migration cost term}}, \qquad \varepsilon_{jod} \overset{\text{i.i.d.}}{\sim} \text{Fréchet}(\eta),$$

where B_d is the fixed amenity level in district d, Z_d denotes pollution in d, and $\lambda < 0$ is the (elasticity) parameter capturing how pollution reduces amenity. $\frac{w_d}{P_d}$ is the real wage in district d. m_{od} is the bilateral cost (in utility-units) of moving from origin o to destination d; hence $e^{-m_{od}}$ is a multiplicative cost-penalty.

Finally, each ε_{jod} is drawn i.i.d. from a Fréchet distribution with shape parameter $\eta > 0$:

$$\Pr(\varepsilon_{jod} < u) \;=\; \exp(-u^{-\eta}), \quad u > 0.$$

Define

$$\Psi_{od} \;\equiv\; B_d\, Z_d^{\lambda}\, \frac{w_d}{P_d}\, e^{-m_{od}}.$$

Then we can write

$$V_{jod} \;=\; \Psi_{od}\, \varepsilon_{jod}.$$

Step 1. CDF of a single V_{jod}.

Since

$$\Pr(\varepsilon_{jod} < u) \;=\; \exp(-u^{-\eta}),$$

it follows that for any $x > 0$,

$$\Pr(V_{jod} < x) \;=\; \Pr(\Psi_{od}\, \varepsilon_{jod} < x) \;=\; \Pr\Big(\varepsilon_{jod} < \tfrac{x}{\Psi_{od}}\Big) \;=\; \exp\Big(-\big(\tfrac{x}{\Psi_{od}}\big)^{-\eta}\Big) \;=\; \exp\Big(-x^{-\eta}\, \Psi_{od}^{\,\eta}\Big).$$

[45] Historical USD-INR conversion rate here: https://en.wikipedia.org/wiki/Exchange_rate_history_of_the_Indian_rupee

Hence each V_{jod} itself is Fréchet-distributed with "scale" Ψ_{od} and the same shape η. Equivalently,

$$F_{V_{jod}}(x) \;=\; \exp[-x^{-\eta}\,\Psi_{od}^{\eta}], \quad x>0.$$

Step 2. Joint distribution of $\max_k V_{jok}$. An individual j born in o will compare V_{jod} for $d = 1,\dots,N$ and choose the d that delivers the highest value. The probability that $\max_k V_{jok} < x$ (i.e. every destination k yields utility below x) is

$$\Pr(\max_k V_{jok} < x) \;=\; \prod_{k=1}^{N} \Pr(V_{jok} < x) \;=\; \prod_{k=1}^{N} \exp[-x^{-\eta}\,\Psi_{ok}^{\eta}] \;=\; \exp\Big[-x^{-\eta}\sum_{k=1}^{N}\Psi_{ok}^{\eta}\Big].$$

Define

$$\Phi_o \;\equiv\; \sum_{k=1}^{N}\Psi_{ok}^{\eta}.$$

Then

$$\Pr(\max_k V_{jok} < x) \;=\; \exp[-x^{-\eta}\,\Phi_o].$$

Step 3. PDF of individual V_{jod}. Differentiate the CDF of V_{jod} w.r.t. x to get its density:

$$f_{V_{jod}}(x) \;=\; \frac{\mathrm{d}}{\mathrm{d}x}\Big[\exp(-x^{-\eta}\,\Psi_{od}^{\eta})\Big] \;=\; \exp(-x^{-\eta}\,\Psi_{od}^{\eta}) \;\times\; [\eta\,x^{-\eta-1}\,\Psi_{od}^{\eta}].$$

Thus

$$f_{V_{jod}}(x) \;=\; \eta\,x^{-\eta-1}\,\Psi_{od}^{\eta}\,\exp(-x^{-\eta}\,\Psi_{od}^{\eta}).$$

Step 4. Probability that destination d is chosen. By definition, the migration-share $\pi_{od} = \Pr(V_{jod} \geq V_{jok}\ \forall\, k \neq d)$. Equivalently, one can write

$$\pi_{od} \;=\; \int_0^{\infty} \Pr(V_{jok} < x\ \forall\, k \neq d)\ \mathrm{d}\Pr(V_{jod} \leq x).$$

But

$$\Pr(V_{jok} < x\ \forall\, k \neq d) \;=\; \frac{\Pr(\max_k V_{jok} < x)}{\Pr(V_{jod} < x)} \;=\; \frac{\exp[-x^{-\eta}\,\Phi_o]}{\exp[-x^{-\eta}\,\Psi_{od}^{\eta}]} \;=\; \exp\Big[-x^{-\eta}\,(\Phi_o - \Psi_{od}^{\eta})\Big].$$

Hence

$$\pi_{od} = \int_0^\infty \Big[\exp\big(-x^{-\eta}(\Phi_o - \Psi_{od}^\eta)\big)\Big]\, \mathrm{d}\Big[F_{V_{jod}}(x)\Big]$$

$$= \int_0^\infty \exp\big(-x^{-\eta}(\Phi_o - \Psi_{od}^\eta)\big)\, \big[\eta\, x^{-\eta-1}\, \Psi_{od}^\eta\big]\, \exp\big(-x^{-\eta}\, \Psi_{od}^\eta\big)\, \mathrm{d}x$$

$$= \int_0^\infty \eta\, x^{-\eta-1}\, \Psi_{od}^\eta\, \exp\big[-x^{-\eta}(\Phi_o - \Psi_{od}^\eta + \Psi_{od}^\eta)\big]\, \mathrm{d}x \;=\; \int_0^\infty \eta\, x^{-\eta-1}\, \Psi_{od}^\eta\, \exp\big[-x^{-\eta}\, \Phi_o\big]\, \mathrm{d}x.$$

Step 5. Evaluating the integral. Set

$$t \;=\; \Phi_o\, x^{-\eta} \;\;\Longrightarrow\;\; \mathrm{d}t \;=\; -\eta\, \Phi_o\, x^{-\eta-1}\, \mathrm{d}x \;\;\Longrightarrow\;\; x^{-\eta-1}\, \mathrm{d}x \;=\; -\frac{1}{\eta\, \Phi_o}\, \mathrm{d}t.$$

When $x \to 0^+$, $t \to +\infty$; when $x \to +\infty$, $t \to 0^+$. Hence

$$\pi_{od} = \int_{x=0}^{\infty} \eta\, \underbrace{x^{-\eta-1}\, \mathrm{d}x}_{=(-\,\mathrm{d}t)/(\eta\,\Phi_o)}\, \Psi_{od}^\eta\, \exp\big(-\Phi_o\, x^{-\eta}\big)$$

$$= \Psi_{od}^\eta \int_{t=+\infty}^{0} \eta \left(-\tfrac{\mathrm{d}t}{\eta\, \Phi_o}\right) \exp(-t) \;=\; \frac{\Psi_{od}^\eta}{\Phi_o} \int_{t=0}^{\infty} e^{-t}\, \mathrm{d}t \;=\; \frac{\Psi_{od}^\eta}{\Phi_o}\,.$$

Since $\int_0^\infty e^{-t}\, \mathrm{d}t = 1$, we obtain the closed-form:

$$\boxed{\pi_{od} \;=\; \frac{\Psi_{od}^\eta}{\sum_{k=1}^N \Psi_{ok}^\eta} \;=\; \frac{\big[B_d\, Z_d^\lambda\, (w_d/P_d)\, e^{-\,m_{od}}\big]^\eta}{\displaystyle\sum_{k=1}^{N} \big[B_k\, Z_k^\lambda\, (w_k/P_k)\, e^{-\,m_{ok}}\big]^\eta}\,,}$$

which matches the migration-share expression in eq. (5) of the main text.

12.6 Hat-algebra for labor-share equilibrium

1. Baseline setup. We have N origins $o \in \{1, \dots, N\}$ and N possible destination districts $d \in \{1, \dots, N\}$. Each origin o supplies N_o workers (think of N_o as the population at o, which does not change from baseline to counterfactual). In the baseline equilibrium:

$$\pi_{od} \;=\; \Pr(V_{jod} \geq V_{jok}\ \forall\, k) \;=\; \frac{\Psi_{od}^{\,\eta}}{\displaystyle\sum_{k=1}^{N} \Psi_{ok}^{\,\eta}}, \quad \Psi_{od} \;=\; B_d\, Z_d^\lambda\, \frac{w_d}{P_d}\, e^{-\,m_{od}}.$$

where B_d is the fixed amenity level in district d, Z_d denotes pollution in d, and $\lambda < 0$ is the

(elasticity) parameter capturing how pollution reduces amenity. $\frac{w_d}{P_d}$ is the real wage in district d. m_{od} is the bilateral cost (in utility-units) of moving from origin o to destination d; hence $e^{-m_{od}}$ is a multiplicative cost-penalty, and $\eta > 0$ is the Fréchet shape parameter.

Given these choice probabilities, the total number of workers (labor) actually living in d is

$$L_d \;=\; \sum_{o=1}^{N} N_o \, \pi_{od}.$$

Finally, define the baseline *labor share* in d as

$$\ell_d \;=\; \frac{L_d}{\sum_{j=1}^{N} L_j}.$$

In the counterfactuals we hold total labor $\sum_j L_j = \sum_o N_o$ constant, so $\ell_d \propto L_d$.

2. Counterfactual ("CF") definitions and hats. In a counterfactual experiment, some—or all—of the following may change:

$$B_d \to B_d^{\text{CF}}, \quad Z_d \to Z_d^{\text{CF}}, \quad w_d \to w_d^{\text{CF}}, \quad P_d \to P_d^{\text{CF}}, \quad m_{od} \to m_{od}^{\text{CF}}.$$

Nothing happens to the origin populations N_o; those remain fixed. We write

$$\hat{B}_d \;\equiv\; \frac{B_d^{\text{CF}}}{B_d}, \quad \hat{Z}_d \;\equiv\; \frac{Z_d^{\text{CF}}}{Z_d}, \quad \widehat{\left(\tfrac{w_d}{P_d}\right)} \;\equiv\; \frac{(w_d^{\text{CF}}/P_d^{\text{CF}})}{(w_d/P_d)}, \quad \hat{m}_{od} \;\equiv\; m_{od}^{\text{CF}} - m_{od}.$$

Define also

$$\Psi_{od}^{\text{CF}} \;=\; B_d^{\text{CF}} \, (Z_d^{\text{CF}})^{\lambda} \, \frac{w_d^{\text{CF}}}{P_d^{\text{CF}}} \, e^{-m_{od}^{\text{CF}}},$$

so that the *counterfactual* "attractiveness" is

$$\Psi_{od}^{\text{CF}} \;=\; \underbrace{B_d \, Z_d^{\lambda} \, \frac{w_d}{P_d} \, e^{-m_{od}}}_{=\Psi_{od}} \;\times\; \underbrace{\left[\hat{B}_d \, \hat{Z}_d^{\lambda} \, \widehat{\left(\tfrac{w_d}{P_d}\right)} \, e^{-\hat{m}_{od}}\right]}_{\equiv\, \hat{\Psi}_{od}}.$$

In other words,

$$\boxed{\hat{\Psi}_{od} \;=\; \frac{\Psi_{od}^{\text{CF}}}{\Psi_{od}} \;=\; \hat{B}_d \, \hat{Z}_d^{\lambda} \, \widehat{\left(\tfrac{w_d}{P_d}\right)} \, e^{-\hat{m}_{od}}.} \tag{25}$$

3. Counterfactual choice probabilities. In the CF, migration share between origin–destination

pair (o, d) is

$$\pi_{od}^{\text{CF}} \;=\; \frac{\left(\Psi_{od}^{\text{CF}}\right)^{\eta}}{\sum_{k=1}^{N}\left(\Psi_{ok}^{\text{CF}}\right)^{\eta}}.$$

But $\Psi_{od}^{\text{CF}} = \Psi_{od}\,\hat{\Psi}_{od}$. Therefore

$$\pi_{od}^{\text{CF}} = \frac{\left(\Psi_{od}\,\hat{\Psi}_{od}\right)^{\eta}}{\sum_{k=1}^{N}\left(\Psi_{ok}\,\hat{\Psi}_{ok}\right)^{\eta}} = \frac{\Psi_{od}^{\eta}\,\hat{\Psi}_{od}^{\eta}}{\sum_{k=1}^{N}\Psi_{ok}^{\eta}\,\hat{\Psi}_{ok}^{\eta}}.$$

Divide numerator and denominator by Ψ_{od}^{η}. In the baseline, $\pi_{od} = \Psi_{od}^{\eta}/\sum_{k}\Psi_{ok}^{\eta}$. Hence

$$\begin{aligned}
\pi_{od}^{\text{CF}} &= \frac{(\Psi_{od}\,\hat{\Psi}_{od})^{\eta}}{\displaystyle\sum_{k=1}^{N}(\Psi_{ok}\,\hat{\Psi}_{ok})^{\eta}} \\
&= \frac{\Psi_{od}^{\eta}\,\hat{\Psi}_{od}^{\eta}\;/\;\Psi_{od}^{\eta}}{\displaystyle\sum_{k=1}^{N}\Psi_{ok}^{\eta}\,\hat{\Psi}_{ok}^{\eta}\;/\;\Psi_{od}^{\eta}} \;=\; \frac{\hat{\Psi}_{od}^{\eta}}{\displaystyle\sum_{k=1}^{N}\left(\tfrac{\Psi_{ok}^{\eta}}{\Psi_{od}^{\eta}}\right)\hat{\Psi}_{ok}^{\eta}} \\
\left(\text{use } \tfrac{\Psi_{ok}^{\eta}}{\Psi_{od}^{\eta}} = \tfrac{\pi_{ok}}{\pi_{od}}\right) &= \frac{\hat{\Psi}_{od}^{\eta}}{\displaystyle\sum_{k=1}^{N}\left(\tfrac{\pi_{ok}}{\pi_{od}}\right)\hat{\Psi}_{ok}^{\eta}} \\
&= \pi_{od}\,\frac{\hat{\Psi}_{od}^{\eta}}{\displaystyle\sum_{k=1}^{N}\pi_{ok}\,\hat{\Psi}_{ok}^{\eta}}
\end{aligned}$$

Define

$$\hat{\Phi}_{o} \;\equiv\; \sum_{k=1}^{N}\pi_{ok}\,\hat{\Psi}_{ok}^{\eta},$$

so that

$$\boxed{\pi_{od}^{\text{CF}} \;=\; \pi_{od}\,\frac{\hat{\Psi}_{od}^{\eta}}{\hat{\Phi}_{o}} \quad\Longrightarrow\quad \hat{\pi}_{od} \;\equiv\; \frac{\pi_{od}^{\text{CF}}}{\pi_{od}} \;=\; \frac{\hat{\Psi}_{od}^{\eta}}{\hat{\Phi}_{o}}.}$$

4. Counterfactual labor in each destination. Baseline labor in d is

$$L_{d} \;=\; \sum_{o=1}^{N} N_{o}\,\pi_{od}.$$

In the counterfactual,

$$L_d^{\text{CF}} \;=\; \sum_{o=1}^{N} N_o\, \pi_{od}^{\text{CF}} \;=\; \sum_{o=1}^{N} N_o \left[\pi_{od}\, \hat{\pi}_{od}\right] \;=\; \sum_{o=1}^{N} [N_o\, \pi_{od}]\; \hat{\pi}_{od}.$$

But the baseline weight of origin o in destination d is

$$\kappa_{od} \;\equiv\; \frac{N_o\, \pi_{od}}{L_d}.$$

Hence

$$L_d^{\text{CF}} \;=\; L_d \sum_{o=1}^{N} \kappa_{od}\, \hat{\pi}_{od}.$$

Therefore the *hat* of destination d's labor is

$$\boxed{\hat{L}_d \;=\; \frac{L_d^{\text{CF}}}{L_d} \;=\; \sum_{o=1}^{N} \kappa_{od}\, \hat{\pi}_{od} \;=\; \sum_{o=1}^{N} \kappa_{od}\, \frac{\hat{\Psi}_{od}^{\,\eta}}{\hat{\Phi}_o}.} \tag{26}$$

5. Counterfactual labor share ℓ_d. By definition, $\ell_d = L_d / \sum_j L_j$. Since total labor $\sum_j L_j = \sum_o N_o$ is fixed, we have

$$\hat{\ell}_d \;=\; \frac{\ell_d^{\text{CF}}}{\ell_d} \;=\; \frac{L_d^{\text{CF}} / (\sum_j N_j)}{L_d / (\sum_j N_j)} \;=\; \frac{L_d^{\text{CF}}}{L_d} \;=\; \hat{L}_d.$$

Hence $\hat{\ell}_d = \hat{L}_d$. Substituting from (26),

$$\boxed{\hat{\ell}_d \;=\; \sum_{o=1}^{N} \kappa_{od}\, \frac{\hat{\Psi}_{od}^{\,\eta}}{\sum_{k=1}^{N} \pi_{ok}\, \hat{\Psi}_{ok}^{\,\eta}},} \quad \kappa_{od} \;=\; \frac{N_o\, \pi_{od}}{L_d}. \tag{27}$$

This boxed system of expressions above defines the counterfactual equilibrium in "hat-algebra": it is *exact* (no linearization, no approximation), and it shows how each counterfactual change in $B_d, Z_d, w_d/P_d, m_{od}$ enters directly into $\hat{\ell}_d$.

Finally, recall from (25) that

$$\hat{\Psi}_{od} \;=\; \hat{B}_d\, \hat{Z}_d^{\lambda}\, \widehat{\left(\tfrac{w_d}{P_d}\right)}\, e^{-\hat{m}_{od}}. \tag{28}$$

Substituting (28) into (27) delivers the *exact* mapping from any shock vector $\{\hat{B}_d,\, \hat{Z}_d,\, \widehat{(w_d/P_d)},\, \hat{m}_{od}\}$ to each destination's new labor-share $\hat{\ell}_d$. In particular, if we have each baseline pair $\{\pi_{od}, \kappa_{od}\}$,

then

$$\hat{\ell}_d \;=\; \sum_{o=1}^{N}\Big(\tfrac{N_o\,\pi_{od}}{L_d}\Big)\,\frac{\big[\hat{B}_d\,\hat{Z}_d^{\lambda}\,(\widehat{w_d/P_d})\,e^{-\hat{m}_{od}}\big]^{\eta}}{\sum_{k=1}^{N}\pi_{ok}\,\big[\hat{B}_k\,\hat{Z}_k^{\lambda}\,(\widehat{w_k/P_k})\,e^{-\hat{m}_{ok}}\big]^{\eta}}.$$

I solve in this "hat" form, which means that shocks to pollution due to city-specific abatements scenario change $\hat{Z}_d$, triggering a new labor share equilibrium. This sytem can also be used to assess the effect of other shocks such as a wage shift that changes $(\widehat{w_d/P_d})$, or a migration-cost subsidy that changes $\hat{m}_{od}$). It allows us to immediately compute new $\{\hat{\ell}_d\}_{d=1}^{N}$ without re-solving the entire non-linear system from scratch.

13 Bibliography

Albouy, David. 2008. "Are Big Cities Bad Places to Live? Estimating Quality of Life Across Metropolitan Areas." Working {Paper}. Working Paper Series. National Bureau of Economic Research. https://doi.org/10.3386/w14472.

Andrews, Isaiah, James H. Stock, and Liyang Sun. 2019. "Weak Instruments in Instrumental Variables Regression: Theory and Practice." Annual Review of Economics 11 (1): 727–53. https://doi.org/10.1146/annurev-economics-080218-025643.

Arrow, Kenneth. 1962. "Economic Welfare and the Allocation of Resources for Invention." In The Rate and Direction of Inventive Activity: Economic and Social Factors, 609–26. Princeton University Press.

Au, Chun-Chung, and J. Vernon Henderson. 2006. "Are Chinese Cities Too Small?" The Review of Economic Studies 73 (3): 549–76. https://www.jstor.org/stable/20185020.

Auffhammer, Maximilian, Solomon M. Hsiang, Wolfram Schlenker, and Adam Sobel. 2013. "Using Weather Data and Climate Model Output in Economic Analyses of Climate Change." Review of Environmental Economics and Policy 7 (2): 181–98. https://doi.org/10.1093/reep/ret016.

Banzhaf, H. Spencer, and B. Andrew Chupp. 2012. "Fiscal Federalism and Interjurisdictional Externalities: New Results and an Application to US Air Pollution." Journal of Public Economics 96 (5): 449–64. https://doi.org/10.1016/j.jpubeco.2012.01.001.

Banzhaf, H. Spencer, and Randall P. Walsh. 2008. "Do People Vote with Their Feet? An Empirical Test of Tiebout." American Economic Review 98 (3): 843–63. https://doi.org/10.1257/aer.98.3.843.

Bayer, Patrick, Nathaniel Keohane, and Christopher Timmins. 2009. "Migration and Hedonic Valuation: The Case of Air Quality." Journal of Environmental Economics and Management 58 (1): 1–14. https://doi.org/10.1016/j.jeem.2008.08.004.

Baylis, Patrick, Kenneth Lee, Harshil Sahai, and Michael Greenstone. 2023. "Is the Demand for Clean Air Too Low? Experimental Evidence from Delhi." Unpublished Manuscript, November.

Behrer, A. Patrick. 2023. Man or Machine? Environmental Consequences of Wage Driven Mechanization in Indian Agriculture. Policy Research Working Papers. The World Bank. https://doi.org/10.1596/1813-9450-10376.

Bellemare, Marc F., and Casey J. Wichman. 2020. "Elasticities and the Inverse Hyperbolic Sine Transformation." Oxford Bulletin of Economics and Statistics 82 (1): 50–61. https://doi.org/10.1111/obes.12325.

Berkouwer, Susanna B., and Joshua T. Dean. 2022. "Credit, Attention, and Externalities in the Adoption of Energy Efficient Technologies by Low-Income Households." American Economic Review 112 (10): 3291–3330. https://doi.org/10.1257/aer.20210766.

Besley, Timothy, and Torsten Persson. 2009. "The Origins of State Capacity: Property Rights, Taxation, and Politics." American Economic Review 99 (4): 1218–44. https://doi.org/10.1257/aer.99.4.1218.

Borgschulte, Mark, David Molitor, and Eric Zou. 2022. "Air Pollution and the Labor Market: Evidence from Wildfire Smoke." Working {Paper}. Working Paper Series. National Bureau of Economic Research. https://doi.org/10.3386/w29952.

Bryan, Gharad, and Melanie Morten. 2019. "The Aggregate Productivity Effects of Internal Migration: Evidence from Indonesia." Journal of Political Economy 127 (5): 2229–68. https://doi.org/10.1086/701810.

Chang, Tom Y., Joshua Graff Zivin, Tal Gross, and Matthew Neidell. 2016. "Particulate Pollution and the Productivity of Pear Packers." American Economic Journal: Economic Policy 8 (3):

141–69. https://doi.org/10.1257/pol.20150085.
———. 2019. "The Effect of Pollution on Worker Productivity: Evidence from Call Center Workers in China." American Economic Journal: Applied Economics 11 (1): 151–72. https://doi.org/10.1257/app.20160436.
Chauvin, Juan Pablo, Edward Glaeser, Yueran Ma, and Kristina Tobio. 2016. "What Is Different About Urbanization in Rich and Poor Countries? Cities in Brazil, China, India and the United States." Working {Paper} 22002. National Bureau of Economic Research. https://doi.org/10.3386/w22002.
Chay, Kenneth Y., and Michael Greenstone. 2005. "Does Air Quality Matter? Evidence from the Housing Market." Journal of Political Economy 113 (2): 376–424. https://doi.org/10.1086/427462.
Chen, Shuai, Paulina Oliva, and Peng Zhang. 2022. "The Effect of Air Pollution on Migration: Evidence from China." Journal of Development Economics 156 (May): 102833. https://doi.org/10.1016/j.jdeveco.2022.102833.
Combes, Pierre-Philippe, and Laurent Gobillon. 2015. "Chapter 5 - The Empirics of Agglomeration Economies." In Handbook of Regional and Urban Economics, edited by Gilles Duranton, J. Vernon Henderson, and William C. Strange, 5:247–348. Handbook of Regional and Urban Economics. Elsevier. https://doi.org/10.1016/B978-0-444-59517-1.00005-2.
Currie, Janet, and Reed Walker. 2019. "What Do Economists Have to Say about the Clean Air Act 50 Years After the Establishment of the Environmental Protection Agency?" Journal of Economic Perspectives 33 (4): 3–26. https://doi.org/10.1257/jep.33.4.3.
Dekle, Robert, Jonathan Eaton, and Samuel Kortum. 2007. "Unbalanced Trade." American Economic Review 97 (2): 351–55. https://doi.org/10.1257/aer.97.2.351.
Deryugina, Tatyana, Garth Heutel, Nolan H. Miller, David Molitor, and Julian Reif. 2019. "The Mortality and Medical Costs of Air Pollution: Evidence from Changes in Wind Direction." American Economic Review 109 (12): 4178–4219. https://doi.org/10.1257/aer.20180279.
Fally, Thibault. 2015. "Structural Gravity and Fixed Effects." Journal of International Economics 97 (1): 76–85. https://doi.org/10.1016/j.jinteco.2015.05.005.
Freeman, Richard, Wenquan Liang, Ran Song, and Christopher Timmins. 2017. "Willingness to Pay for Clean Air in China." Working paper 24157. National Bureau of Economic Research. https://doi.org/10.3386/w24157.
Fu, Shihe, V Brian Viard, and Peng Zhang. 2021. "Air Pollution and Manufacturing Firm Productivity: Nationwide Estimates for China." The Economic Journal 131 (640): 3241–73. https://doi.org/10.1093/ej/ueab033.
Ganguly, Tanushree, Kurinji L. Selvaraj, and Sarath K. Guttikunda. 2020. "National Clean Air Programme (NCAP) for Indian Cities: Review and Outlook of Clean Air Action Plans." Atmospheric Environment: X 8 (December): 100096. https://doi.org/10.1016/j.aeaoa.2020.100096.
Ghanem, Dalia, and Junjie Zhang. 2014. "'Effortless Perfection:' Do Chinese Cities Manipulate Air Pollution Data?" Journal of Environmental Economics and Management 68 (2): 203–25. https://doi.org/10.1016/j.jeem.2014.05.003.
Gollin, Douglas. 2014. "The Lewis Model: A 60-Year Retrospective." Journal of Economic Perspectives 28 (3): 71–88. https://doi.org/10.1257/jep.28.3.71.
Government of India, Ministry of Finance. 2013a. Demand for Grants 2012-13. New Delhi: Department of Economic Affairs, Economic Division. https://doe.gov.in/files/detailed-demands-documents/detaildg270312.pdf.
———. 2013b. Economic Survey 2012–13. New Delhi: Department of Economic Affairs, Economic Division.

Graff Zivin, Joshua, and Matthew Neidell. 2012. "The Impact of Pollution on Worker Productivity." American Economic Review 102 (7): 3652–73. https://doi.org/10.1257/aer.102.7.3652.

Greenstone, Michael. 2022. "AQLI India Fact Sheet."

Greenstone, Michael, and Rema Hanna. 2014. "Environmental Regulations, Air and Water Pollution, and Infant Mortality in India." American Economic Review 104 (10): 3038–72. https://doi.org/10.1257/aer.104.10.3038.

Greenstone, Michael, Guojun He, Ruixue Jia, and Tong Liu. 2022. "Can Technology Solve the Principal-Agent Problem? Evidence from China's War on Air Pollution." American Economic Review: Insights, April. https://doi.org/10.1257/aeri.20200373.

Greenstone, Michael, and B. Kelsey Jack. 2015. "Envirodevonomics: A Research Agenda for an Emerging Field." Journal of Economic Literature 53 (1): 5–42. https://doi.org/10.1257/jel.53.1.5.

Greenstone, Michael, Kenneth Lee, and Harshil Sahai. 2021. "Indoor Air Quality, Information, and Socioeconomic Status: Evidence from Delhi." AEA Papers and Proceedings 111 (May): 420–24. https://doi.org/10.1257/pandp.20211006.

Guttikunda, Sarath K., Sai Krishna Dammalapati, Gautam Pradhan, Bhargav Krishna, Hiren T. Jethva, and Puja Jawahar. 2023. "What Is Polluting Delhi's Air? A Review from 1990 to 2022." Sustainability 15 (5): 4209. https://doi.org/10.3390/su15054209.

Hammer, Melanie S., Aaron van Donkelaar, Chi Li, Alexei Lyapustin, Andrew M. Sayer, N. Christina Hsu, Robert C. Levy, et al. 2020. "Global Estimates and Long-Term Trends of Fine Particulate Matter Concentrations (1998–2018)." Environmental Science & Technology 54 (13): 7879–90. https://doi.org/10.1021/acs.est.0c01764.

Hanna, Rema, and Paulina Oliva. 2015. "The Effect of Pollution on Labor Supply: Evidence from a Natural Experiment in Mexico City." Journal of Public Economics 122 (February): 68–79. https://doi.org/10.1016/j.jpubeco.2014.10.004.

Harris, John R., and Michael P. Todaro. 1970. "Migration, Unemployment and Development: A Two-Sector Analysis." The American Economic Review 60 (1): 126–42. https://www.jstor.org/stable/1807860.

Heblich, Stephan, Alex Trew, and Yanos Zylberberg. 2021. "East-Side Story: Historical Pollution and Persistent Neighborhood Sorting." Journal of Political Economy 129 (5): 1508–52. https://doi.org/10.1086/713101.

Heo, Seonmin Will, Koichiro Ito, and Rao Kotamarthi. 2025. "International Spillover Effects of Air Pollution: Evidence from Mortality and Health Data." The Review of Economics and Statistics, March, 1–45. https://doi.org/10.1162/rest_a_01581.

Hoffmann, Bridget, and Juan Pablo Rud. 2024. "The Unequal Effects of Pollution on Labor Supply." Econometrica 92 (4): 1063–96. https://doi.org/10.3982/ECTA20484.

Imbert, Clément, and John Papp. 2020a. "Short-Term Migration, Rural Public Works, and Urban Labor Markets: Evidence from India." Journal of the European Economic Association 18 (2): 927–63. https://doi.org/10.1093/jeea/jvz009.

———. 2020b. "Costs and Benefits of Rural-Urban Migration: Evidence from India." Journal of Development Economics 146 (September): 102473. https://doi.org/10.1016/j.jdeveco.2020.102473.

Jack, B. Kelsey, Seema Jayachandran, Namrata Kala, and Rohini Pande. 2025. "Money (Not) to Burn: Payments for Ecosystem Services to Reduce Crop Residue Burning." American Economic Review: Insights 7 (1): 39–55. https://doi.org/10.1257/aeri.20230431.

Jayachandran, Seema. 2022. "How Economic Development Influences the Environment." Annuul Review of Economics 14 (September): 1–30. https://doi.org/doi.org/10.1146.

Kahn, Matthew E., Pei Li, and Daxuan Zhao. 2015. "Water Pollution Progress at Borders: The Role

of Changes in China's Political Promotion Incentives." American Economic Journal: Economic Policy 7 (4): 223–42. https://doi.org/10.1257/pol.20130367.

Khanna, Gaurav, Wenquan Liang, Ahmed Mushfiq Mobarak, and Ran Song. 2025. "The Productivity Consequences of Pollution-Induced Migration in China." American Economic Journal: Applied Economics 17 (2): 184–224. https://doi.org/10.1257/app.20220655.

Khera, Reetika, ed. 2011. The Battle for Employment Guarantee. Oxford, New York: Oxford University Press.

Kone, Zovanga L, Maggie Y Liu, Aaditya Mattoo, Caglar Ozden, and Siddharth Sharma. 2018. "Internal Borders and Migration in India*." Journal of Economic Geography 18 (4): 729–59. https://doi.org/10.1093/jeg/lbx045.

Lewis, Daniel J, and Karel Mertens. 2022. "A Robust Test for Weak Instruments with Multiple Endogenous Regressors - FEDERAL RESERVE BANK of NEW YORK." https://www.newyorkfed.org/research/staff_reports/sr1020.html.

Lewis., Arthur. 1954. "Economic Development with Unlimited Supplies of Labour." The Manchester School 22 (2): 139–91. https://doi.org/10.1111/j.1467-9957.1954.tb00021.x.

Lin, Wei, and Jeffrey M. Wooldridge. 2019. "Chapter 2 - Testing and Correcting for Endogeneity in Nonlinear Unobserved Effects Models." In Panel Data Econometrics, edited by Mike Tsionas, 21–43. Academic Press. https://doi.org/10.1016/B978-0-12-814367-4.00002-2.

Lipscomb, Molly, and Ahmed Mushfiq Mobarak. 2017. "Decentralization and Pollution Spillovers: Evidence from the Re-Drawing of County Borders in Brazil*." The Review of Economic Studies 84 (1): 464–502. https://doi.org/10.1093/restud/rdw023.

Liu, Shi V., Fu-lin Chen, and Jianping Xue. 2019. "A Meta-Analysis of Selected Near-Road Air Pollutants Based on Concentration Decay Rates." Heliyon 5 (8): e02236. https://doi.org/10.1016/j.heliyon.2019.e02236.

McDuffie, Erin E., Randall V. Martin, Joseph V. Spadaro, Richard Burnett, Steven J. Smith, Patrick O'Rourke, Melanie S. Hammer, et al. 2021. "Source Sector and Fuel Contributions to Ambient PM2.5 and Attributable Mortality Across Multiple Spatial Scales." Nature Communications 12 (1): 3594. https://doi.org/10.1038/s41467-021-23853-y.

Redding, Stephen, and Esteban Rossi-Hansberg. 2017. "Quantitative Spatial Economics." Annual Review of Economics 9 (1): 21–58. https://doi.org/10.1146/annurev-economics-063016-103713.

Redding, Stephen, and Daniel M. Sturm. 2008. "The Costs of Remoteness: Evidence from German Division and Reunification." American Economic Review 98 (5): 1766–97. https://doi.org/10.1257/aer.98.5.1766.

Roback, Jennifer. 1982. "Wages, Rents, and the Quality of Life." Journal of Political Economy 90 (6): 1257–78. https://www.jstor.org/stable/1830947.

Rosenthal, Stuart S., and William C. Strange. 2004. "Chapter 49 - Evidence on the Nature and Sources of Agglomeration Economies." In Handbook of Regional and Urban Economics, edited by J. Vernon Henderson and Jacques-François Thisse, 4:2119–71. Cities and Geography. Elsevier. https://doi.org/10.1016/S1574-0080(04)80006-3.

Santos Silva, J. M. C., and Silvana Tenreyro. 2022. "The Log of Gravity at 15." Portuguese Economic Journal 21 (3): 423–37. https://doi.org/10.1007/s10258-021-00203-w.

Silva, J. M. C. Santos, and Silvana Tenreyro. 2006. "The Log of Gravity." The Review of Economics and Statistics 88 (4): 641–58. https://doi.org/10.1162/rest.88.4.641.

Singh, Nimish, Shivang Agarwal, Sumit Sharma, Satoru Chatani, and Veerabhadran Ramanathan. 2021. "Air Pollution Over India: Causal Factors for the High Pollution with Implications for Mitigation." ACS Earth and Space Chemistry 5 (12): 3297–3312. https://doi.org/10.1021/acsearthspacechem.1c00170.

Tessum, Christopher W., Jason D. Hill, and Julian D. Marshall. 2017. "InMAP: A Model for Air Pollution Interventions." PLOS ONE 12 (4): e0176131. https://doi.org/10.1371/journal.pone.0176131.

Thakrar, Sumil K., Christopher W. Tessum, Joshua S. Apte, Srinidhi Balasubramanian, Dylan B. Millet, Spyros N. Pandis, Julian D. Marshall, and Jason D. Hill. 2022. "Global, High-Resolution, Reduced-Complexity Air Quality Modeling for PM2.5 Using InMAP (Intervention Model for Air Pollution)." PLOS ONE 17 (5): e0268714. https://doi.org/10.1371/journal.pone.0268714.

Tiwari, Anshuman. 2025. "The Effect of Air Pollution on GDP: Evidence from a Natural Experiment in India." {SSRN} {Scholarly} {Paper}. Rochester, NY: Social Science Research Network. https://doi.org/10.2139/ssrn.5243794.

US EPA, OAR. 2015. "Setting Emissions Standards Based on Technology Performance." Overviews and {Factsheets}. https://www.epa.gov/clean-air-act-overview/setting-emissions-standards-based-technology-performance.

Vallero, Daniel. 1973. "Fundamentals of Air Pollution - 1st Edition." https://shop.elsevier.com/books/fundamentals-of-air-pollution/stern/978-0-12-666560-4.