

Online Appendix for “The Productivity Consequences of Pollution-Induced Migration in China”, Gaurav Khanna, Wenquan Liang, Ahmed Mushfiq Mobarak and Ran Song

Table of Contents

A	Robustness of Empirical Specifications	A1
A.1	The Different Instruments and their Sources of Variation	A1
A.2	Alternative Model Specifications, Controls and Samples	A12
B	Air Pollution Data Disclosure in China	A20
C	Additional Tables and Figures	A21
D	Additional Counterfactual Results	A29
E	Additional Model Derivations	A36
E.1	Deriving Labor Supply and Welfare	A36
E.2	Elasticity of Capital, and Skill-biased Capital	A38
E.3	On Existence and Uniqueness of the Equilibrium	A39
E.4	Alternative Specification of Housing Market	A40
F	Data Appendix	A44

A Robustness of Empirical Specifications

We conduct a wide-range of meaningful robustness checks to evaluate the concreteness of the empirical relationship between air quality and migration. We explore threats to identification, different instruments, alternate model specifications and data.

A.1 The Different Instruments and their Sources of Variation

Here we explore threats to identification for our instruments. We test concerns of the endogenous placements of power plants, whereby policy makers may use the same function – the simultaneous interaction between wind directions, distance to a given city and coal consumption – to determine where to place new plants. We thus exclude any plants built within different distance radii around the city. We may still think that *newly* built plants are endogenously placed. Yet, our results are robust to relying on old power plants, and to make it even more conservative the cities near the newly built plants are in the ‘control’ group. Last, we show that the IV is not predicted by baseline city-level characteristics.

We also explore the variation underlying the thermal inversions IV. We fail to find meaningful predictors of future inversions, and as such conclude that such events are random. Finally, we explore the variation generated by China’s Huai river heating policy (Chen et al., 2013). Even though we fail to find substantial effects on out-migration rates, they do help predict differential in-migration by skill.

A.1.1 Instrumental Variable Estimates

In Table A1 we show the strength of the first stage relationships between our different instruments and our independent variable of interest.

In the first three columns of Table A2, we estimate a stacked regression analogue of Table 1, where we include an interaction term between pollution and high-skill indicators (and a control for the high-skilled indicator). Here, we instrument for pollution, and also for the interaction between pollution and high-skilled (using the interaction between the instrument and the high-skilled indicator). The results indicate that pollution increases emigration, and this effect is larger for the high-skilled.

In the next three columns of of Table A2, we look at different ways to reformulate the estimation sample. Instead of splitting up the sample into low and high skilled, we split it up into three categories: high school or below, those with some college education, and

Table A1: The First Stage Across Different Instruments

Panel A: City-level			Dependent variable: Log (PM2.5)						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Wind+Coal IV	0.0240*** (0.00211)	0.0113*** (0.00235)	0.00954*** (0.00236)						
Number of inversions				0.00222*** (0.000350)	0.00159*** (0.000329)	0.00201*** (0.000293)			
Strength of inversions							0.000730*** (0.000214)	0.000448** (0.000185)	0.000725*** (0.000168)
Observations	332	332	332	332	332	332	332	332	332
R-squared	0.297	0.40	0.458	0.212	0.441	0.538	0.097	0.389	0.483
City Controls	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Weather Controls	No	No	Yes	No	No	Yes	No	No	Yes
Panel B: Individual-level			Dependent variable: Log (PM2.5)						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Wind+Coal IV	0.0200*** (0.00191)	0.0136*** (0.00188)	0.0129*** (0.00188)						
Number of inversions				0.00212*** (0.000285)	0.00184*** (0.000254)	0.00227*** (0.000224)			
Strength of inversions							0.000822*** (0.000165)	0.000670*** (0.000151)	0.000944*** (0.000132)
Observations	761,548	761,548	761,548	761,548	761,548	761,548	761,548	761,548	761,548
R-squared	0.284	0.403	0.453	0.248	0.441	0.551	0.160	0.379	0.491
City Controls	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Demographics	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Weather Controls	No	No	Yes	No	No	Yes	No	No	Yes

Notes: City-level regressions of 332 cities in Panel A and individual-level regressions across 332 cities in Panel B. Panel A shows the first-stage results for city-level IV regressions and Panel B shows the first-stage results for individual-level IV regressions. Standard errors clustered at the hukou city level are reported in parentheses. City controls include distance to Shanghai, Tianjin and Shenzhen seaports. Weather controls include temperature, wind speed, sunshine duration and humidity. In Panels A and B columns 2-3, we also control for inverse distance weighted average of coal consumption for power plants within 500 km of a city.

those with college or above education. A steep education gradient is apparent, where the elasticity of migration with respect to PM2.5 is higher for higher levels of education.

Finally, Table A3 shows similar results using the combinations of the Wind IV and thermal inversions instruments, and the corresponding over-identification tests.

We further examine the association between pollution and *destination* choices. We regress city-level in-migration on local pollution concentration, and use our instruments to deal with the endogeneity of air quality in destination cities. In Table A4 we find that even for in-migration decisions, the response of high-skilled workers is greater than that of low-skilled workers. In other words, high-skilled workers are more likely to move to cities with clean air when they make location choices. Severe air pollution not only results in the outflow of high-skilled workers but also reduce their inflow.

Table A2: Stacked Regressions and Disaggregated Skill Levels

	Leave hukou city indicator					
	(1)	(2)	(3)	(4)	(5)	(6)
	Stacked Regressions (with High-skill interaction) Full sample Wind IV	Full sample Thermal Num	Full sample Thermal Strength	Disaggregated Education Levels High school or below	Some college Wind IV	College or above
Log(PM2.5)	0.0753** (0.0379)	0.0845*** (0.0203)	0.0704*** (0.0230)	0.0609 (0.0413)	0.114*** (0.0394)	0.175*** (0.0402)
Log(PM2.5) $\times$ High-skill	0.0589** (0.0247)	0.0810** (0.0376)	0.110** (0.0528)			
Observations	761,548	761,548	761,548	643,124	64,598	53,826
City Controls	Yes	Yes	Yes	Yes	Yes	Yes
Demographics	Yes	Yes	Yes	Yes	Yes	Yes
Weather Controls	No	Yes	Yes	No	No	No

Notes: Individual level regressions across 332 cities. Standard errors clustered at the hukou city level are reported in parentheses. We report Kleibergen-Paap rk Wald F statistic. Instrumental variables specification using the interaction between wind direction, distance to coal plant, and coal consumption at power plant (in column 1 and columns 4-6), and using the number of thermal inversions (in column 2) and the strength of thermal inversions (in column 3). We also control for a high-skilled indicator in column 1-3. City controls include distance to Shanghai, Tianjin and Shenzhen seaports. Demographics include age, age-squared, gender, marital status, and an urban hukou indicator. Weather controls include temperature, wind speed, sunshine duration and humidity. In column 1 and columns 4-6, we also control for inverse distance weighted average of coal consumption for power plants within 500 km of a city.

Table A3: Combined Instruments

	Dependent variable: Leave hukou city indicator					
	Wind and Coal +Number of inversions			Wind and Coal +Strength of inversions		
	Full sample (1)	Low edu (2)	High edu (3)	Full sample (4)	Low edu (5)	High edu (6)
Log (PM2.5)	0.0892*** (0.0187)	0.0836*** (0.0195)	0.124*** (0.0223)	0.0789*** (0.0202)	0.0717*** (0.0209)	0.126*** (0.0255)
Observations	761,548	643,124	118,424	761,548	643,124	118,424
Hansen J statistic	0.214	0.714	1.522	0.000	0.147	1.683
Hansen P value	0.643	0.398	0.217	0.995	0.701	0.195
City Controls	Yes	Yes	Yes	Yes	Yes	Yes
Demographics	Yes	Yes	Yes	Yes	Yes	Yes
Weather Controls	Yes	Yes	Yes	Yes	Yes	Yes
F-test of IVs	70.16	67.06	55.58	46.50	43.74	38.33

Notes: Individual level regressions across 332 cities. Standard errors clustered at the hukou city level are reported in parentheses. We report Kleibergen-Paap rk Wald F statistic. City controls include the log distance to Shanghai seaport, to Tianjin seaport, and to Shenzhen seaport, and inverse distance weighted average of coal consumption for power plants within 500 km of a city. Demographics include age, age-squared, gender, marital status, and an urban hukou indicator. Weather controls include temperature, wind speed, sunshine duration and humidity.

Table A4: Pollution and In-migration

Panel A		Dependent variable: Share of In-migrants							
No weather	Wind+Coal IV			Number of inversions			Strength of inversions		
	Full sample (1)	Low edu (2)	High edu (3)	Full sample (4)	Low edu (5)	High edu (6)	Full sample (7)	Low edu (8)	High edu (9)
Log (PM2.5)	-1.838*** (0.523)	-1.834*** (0.514)	-2.242*** (0.708)	-1.204** (0.500)	-1.137** (0.475)	-1.645** (0.677)	-1.404** (0.643)	-1.275** (0.606)	-2.152** (0.903)
Observations	329	329	329	329	329	329	329	329	329
City Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Weather Controls	No	No	No	No	No	No	No	No	No
F-test of IVs	41.28	41.28	41.28	39.56	39.56	39.56	16.36	16.36	16.36

Panel B		Dependent variable: Share of In-migrants							
Weather controls	Wind+Coal IV			Number of inversions			Strength of inversions		
	Full sample (1)	Low edu (2)	High edu (3)	Full sample (4)	Low edu (5)	High edu (6)	Full sample (7)	Low edu (8)	High edu (9)
Log (PM2.5)	-1.841*** (0.512)	-1.833*** (0.494)	-2.273*** (0.718)	-1.204*** (0.373)	-1.181*** (0.357)	-1.451*** (0.503)	-1.359*** (0.421)	-1.301*** (0.401)	-1.794*** (0.583)
Observations	329	329	329	329	329	329	329	329	329
City Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Weather Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
F-test of IVs	33.74	33.74	33.74	91.64	91.64	91.64	47.63	47.63	47.63

Notes: City level regressions of 329 cities. We drop cities with missing in-migration rates. Dependent variable is the log share of in-migrants in city population. Independent variable is destination city PM2.5. Standard errors clustered at the hukou city level are reported in parentheses. We report Kleibergen-Paap rk Wald F statistic. Instrumental variables specification using the interaction between wind direction, distance to coal plant, and coal consumption at power plant, and using the number of thermal inversions and the strength of thermal inversions. City controls include distance to Shanghai, Tianjin and Shenzhen seaports, and inverse distance weighted average of coal consumption for power plants within 500 km of a city. Demographics include age, age-squared, gender, marital status, and an urban hukou indicator. In Panel B, weather controls include temperature, wind speed, sunshine duration and humidity. In Panels A and B columns 1-3, we also control for inverse distance weighted average of coal consumption for power plants within 500 km of a city.

A.1.2 Endogeneity Concerns over Instrumental Strategies

In this section we test certain identification assumptions. Our first instrument is the interaction between the three components: wind direction, distance and coal consumption. We may expect that policy makers take these components into account when placing large coal-fired plants near certain type of cities, and so the instrument may be correlated with unobservable characteristics of nearby cities. In Table A5, we exclude any plants built within 200km of a given city (first two columns), and then within 400km of the city (last two columns). Our results are similar to before, with an increase in precision.

Table A5: Different Distance Bins for Selection of Plants

	Dependent variable: Leave hukou city indicator					
	Distance 200-600km			Distance 400-800km		
	Full sample (1)	Low edu (2)	High edu (3)	Full sample (4)	Low edu (5)	High edu (6)
Log (PM2.5)	0.0947*** (0.0365)	0.0798** (0.0396)	0.163*** (0.0383)	0.0973*** (0.0304)	0.0834*** (0.0319)	0.168*** (0.0385)
Observations	761,548	643,124	118,424	761,548	643,124	118,424
City Controls	Yes	Yes	Yes	Yes	Yes	Yes
Demographics	Yes	Yes	Yes	Yes	Yes	Yes
F-test of IVs	56.73	50.95	43.11	59.73	53.08	58.44

Notes: Individual level regressions across 332 cities. Standard errors clustered at the hukou city level are reported in parentheses. We report Kleibergen-Paap rk Wald F statistic. Instrumental variables specification based on the interaction between wind direction, distance to coal plant, and coal consumption of power plant. City controls include distance to Shanghai, Tianjin and Shenzhen seaports. Demographics include age, age-squared, gender, marital status, and an urban hukou indicator.

We may expect that newly built plants are subject to more scrutiny as the conversation about air quality in China has recently escalated. In Table A6 we exclude newly built plants from the IV, and instead include cities with newly built ones in the ‘control’ group. Although this empirical strategy is more conservative, we find similar patterns and magnitudes.

One concern with the exogeneity of wind direction is that the Chinese government might select thermal plant locations in a way that pollution did not travel to populated or politically important cities. If that were the case, coal-fired plants are less likely to be located upwind of such cities. In Table A7, we present the number of large-scale thermal plants located upwind of five largest metropolitan areas in 2014 along with their total coal

Table A6: Excluding Newly Built Power Plants

	Dependent variable: Leave hukou city indicator							
	Plants > 5 yrs ago		Plants > 10 yrs ago		Plants > 15 yrs ago		Plants > 20 yrs ago	
	Low edu	High edu	Low edu	High edu	Low edu	High edu	Low edu	High edu
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Log (PM2.5)	0.0614 (0.0411)	0.137*** (0.0372)	0.0777* (0.0441)	0.146*** (0.0405)	0.0776* (0.0422)	0.149*** (0.0394)	0.1000** (0.0458)	0.156*** (0.039)
Observations	643,124	118,424	643,124	118,424	643,124	118,424	643,124	118,424
City Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Demographics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
F-test of IVs	44.50	40.42	33.20	34.36	35.40	42.87	30.93	38.28

Notes: Individual level regressions across 332 cities. Cities affected by new plants included in sample (i.e. in the ‘control’ regions) so as to generate conservative estimates. Standard errors clustered at the hukou city level are reported in parentheses. We report Kleibergen-Paap rk Wald F statistic. Instrumental variables specification using the interaction between wind direction, distance to coal plant, and coal consumption at power plant. City controls include the log distance to Shanghai seaport to Tianjin seaport and to Shenzhen seaport, and inverse distance weighted average of coal consumption for power plants within 500 km of a city. Demographics include age, age-squared, gender, marital status, and an urban hukou indicator.

consumption. Beijing and Tianjin are among the most populated and politically important cities in Northern China. The ratio of the upwind large thermal plants to the total number of large thermal plants is 22.4% and 50.9% for Beijing and Tianjin, respectively. The mean of the ratio for all Chinese cities is 37.8%, which is between the values in Beijing and Tianjin. This may suggest that the Chinese government does not necessarily locate the coal-fired plants away from populated or politically important cities.

To further test whether politicians avoid populated, politically important and rich cities when building new plants, we explore whether baseline city features predict newly built plants. In Table A8, we explore whether city-level characteristics in 2004 can predict (a) the ratio of upwind plants built after 2005, and (b) the IV based on plants built after 2005. We find no meaningful associations between these variables and possible predictors of a city’s influence, like baseline populations, GDP, total electricity consumption and industrial electricity consumption. In the following section, we also show that our results are robust to excluding big cities, major provincial capitals and coal producing regions (Table A13).

In Table A9, we construct placebo instruments, artificially changing the wind angle by 90, and then 180 degrees. The first two columns report the first stage results. As the angle is increased, the falsified instrument is less likely to predict PM2.5 or emigration.

Table A7: The Coal-fired Plants Located Upwind of Large metropolitans

City	Number of Upwind Plants	Ratio of Upwind Plants	Coal Consumption of Upwind Plants	Smallest Angle of Plants
Beijing	58	22.41%	2372	25
Tianjin	59	50.85%	2123	26
Shanghai	62	1.61%	2067	16
Guangzhou	30	16.67%	899	22
Shenzhen	25	44.00%	785	26
National mean	43.16	37.82%	511	25

Notes: The statistics are calculated using the large-scale thermal power plants located outside a given city and within 500km. Following [Freeman et al. \(2019\)](#), we define the upwind region as a section of a circular buffer drawn at a distance of 500km from a given city, and the angle between the section and the annual dominant wind direction of the city being at least 45 degrees.

Table A8: Baseline Economy and the Wind Direction IV

Dependent variable:	The ratio of upwind plants		Wind direction and coal plants IV	
	(1)	(2)	(3)	(4)
Baseline Population	-0.0106 (0.0309)	-0.0132 (0.0279)	0.412 (0.489)	0.226 (0.440)
Baseline GDP per capita	-0.0151 (0.0302)	-0.0231 (0.0289)	-0.132 (0.526)	-0.414 (0.522)
Baseline Industrial Elec cons	0.0225 (0.0224)		0.210 (0.347)	
Baseline Total Elec cons		0.0255 (0.0178)		0.414 (0.269)
Observations	281	281	281	281
City Controls	Yes	Yes	Yes	Yes

Notes: City level regressions for 281 cities. We drop cities with missing values in baseline characteristics. Dependent variables are based on plants built post 2005, independent variables are measured in the year 2004. Standard errors clustered at the hukou city level are reported in parentheses.

Table A9: Placebo Wind Directions

	Log (PM2.5)		Leave hukou city indicator			
	(1)	(2)	Coal IV Placebo (+90 degrees) Low edu	Coal IV Placebo (+90 degrees) High edu	Coal IV Placebo (+180 degrees) Low edu	Coal IV Placebo (+180 degrees) High edu
			(3)	(4)	(5)	(6)
Log (PM2.5)			-0.0480 (0.0539)	0.0430 (0.0542)	-0.115 (0.126)	0.00989 (0.0744)
Coal IV Placebo (wind direction+90 degrees)	0.00874** (0.00487)					
Coal IV Placebo (wind direction+180 degrees)		0.00102 (0.00171)				
Observations	332	332	643,124	118,424	643,124	118,424
City Controls	Yes	Yes	Yes	Yes	Yes	Yes
Demographics	N	N	Yes	Yes	Yes	Yes
F-test of IVs			6.482	9.715	0.903	1.599

Notes: Individual level regressions across 332 cities. Standard errors clustered at the hukou city level are reported in parentheses. We report Kleibergen-Paap rk Wald F statistic. Instrumental variables specification using the interaction between wind direction (falsified), distance to coal plant, and coal consumption at power plant. City controls include the log distance to Shanghai seaport, to Tianjin seaport, and to Shenzhen seaport, and inverse distance weighted average of coal consumption for power plants within 500 km of a city. Demographics include age, age-squared, gender, marital status, and an urban hukou indicator.

Finally, we turn our attention to the thermal inversions IV, used extensively by researchers in many different contexts (Arceo et al., 2016; Chen et al., 2022; Hicks et al., 2015; Jans et al., 2018), and as such, been scrutinized thoroughly. Nonetheless, we examine whether lagged pollution levels can predict future levels of the strength of thermal inversions in Table A10. We fail to find any such meaningful associations. Furthermore, we also find that lagged inversions do not predict future inversions, suggesting that levels of auto-correlation in inversions are low, and we may consider the data generating process underlying inversions to be close to random.

Table A10: Lagged Pollution and Thermal Inversions

	Dependent variable: Strength of inversions		
	(1)	(2)	(3)
Lagged Log(PM2.5)	-11.68 (12.88)	-11.79 (12.77)	-11.38 (12.80)
Lagged number of inversions			0.241 (0.181)
Lagged strength of inversions		0.0332 (0.0330)	-0.0373 (0.0774)
Observations	3,652	3,652	3,652
R-squared	0.961	0.961	0.961
City fixed effects	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes

Notes: Standard errors clustered at city level are reported in parentheses. City level regressions for 332 cities over 11 years (2005 to 2015). Specifications include city and year fixed effects.

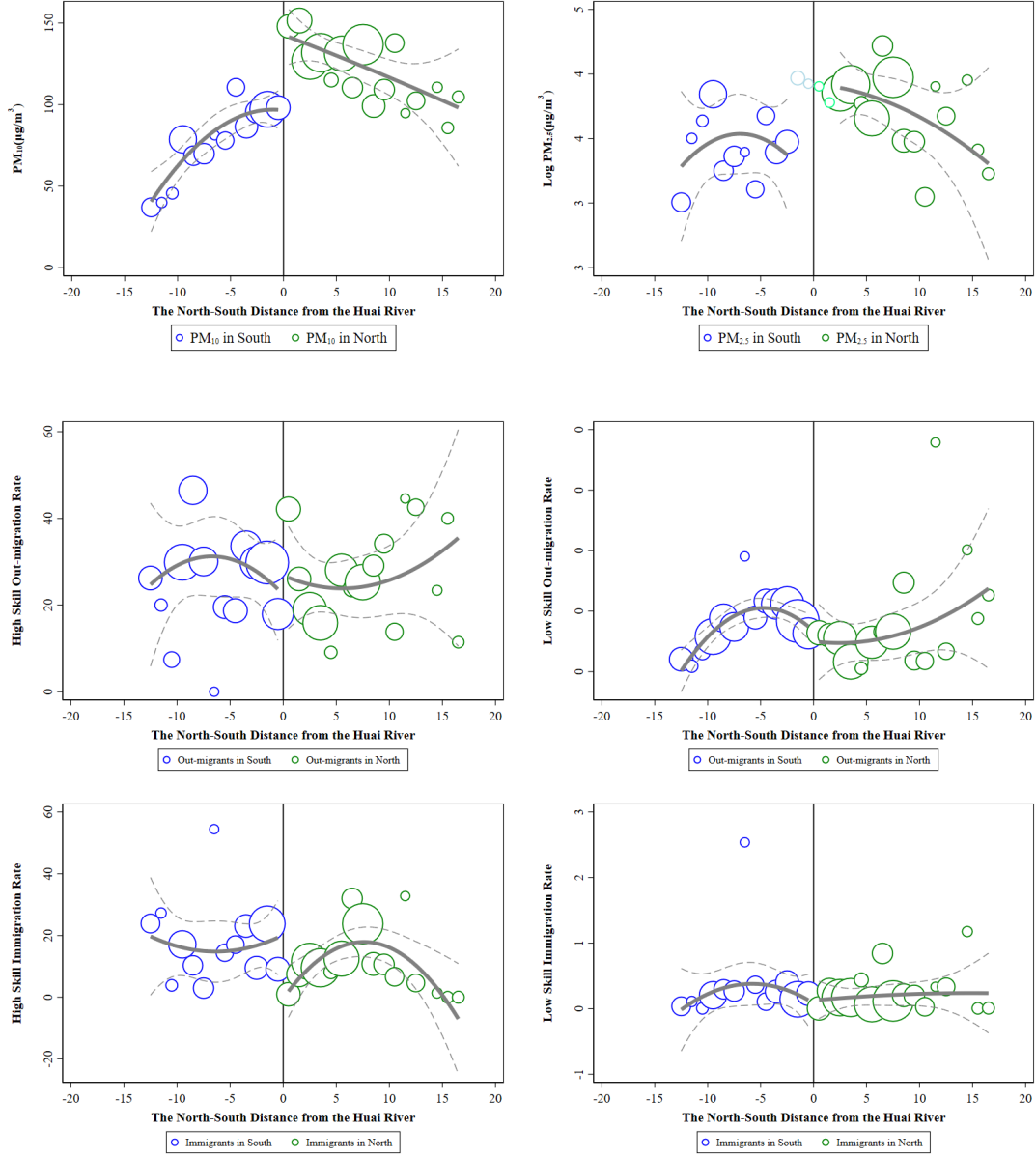
A.1.3 The Huai River Regression Discontinuity

Between 1950-1980 China established coal-based free heating systems to residences and offices north of the Huai River. This policy had long lasting effects, as even today the heating systems are different between the northern and southern parts of the country. The north relies on coal boilers releasing a large amount of pollutants. Chen et al. (2013) examine the effects of this policy on life expectancy using an RD where they compare cities just north of the river to those just south of it. We find that high-skilled workers are less likely to in-migrate into cities with more pollution (but find no significant impacts on out-migration).

Here, we leverage the same empirical setup to examine migration decisions. Figure A1 shows the RD graphs where the horizontal axis represents the distance between the city and

the Huai river. In our top row we show the discontinuity in PM10 and PM2.5 levels. In the bottom row, we look at in-migration rates, and find that only for the high-skill workforce, there is less in-migration in cities that have more pollution. This effect is statistically and economically meaningful in our RD regressions. We find no such differential response on in-migration for the low-skilled.

Figure A1: The Huai River RD



Notes: The top row shows the discontinuity in PM10 and PM2.5 at the Huai River. Second row shows the out-migration by skill level. Bottom row shows the in-migration rate by skill-level. Bubble sizes are baseline city populations. The PM10 data come from [Ebenstein et al. \(2017\)](#). We collect the data on latitude and longitude by hand ([Khanna et al., 2024](#)).

A.2 Alternative Model Specifications, Controls and Samples

Here we examine different model specifications, sample restrictions and control variables. First, we employ an individual-level longitudinal panel data, allowing us to track individuals over time and control for individual-level unobservables. Importantly, we use an alternative definition for migration status regardless of *hukou* location.

Then, we turn to the implications of cumulative pollution. We find that workers are more sensitive to cumulative pollution when they make location choices, compared to short-term pollution. Once again, the impacts of cumulative pollution are more pronounced for high-skilled workers than their low-skilled counterparts. Finally, we use alternative samples, alternative measures of air quality, and alternate sets of co-variate controls to examine the effects of pollution.

A.2.1 Individual Longitudinal Panel and Alternative Definition of Migration

We employ an individual-level longitudinal panel and a different definition of migration status to explore the spatial sorting of Chinese workers. We use the China Labor-force Dynamic Survey (CLDS), which is a national social survey targeted at labor force dynamics in China. CLDS 2016 asks a retrospective history of locations for individuals, and we create an individual-level longitudinal panel between 2008 and 2016. Here, we define migration to be an indicator for whether or not an individual changed city location between years, regardless of whether they change their *hukou*. The strengths of the individual-level panel lie in that it allows us to account for individual-specific unobservables and track those who have moved multiple times and who have moved and returned home.

Table A11 presents the IV estimation of the relationship between pollution and the out-migration tendency of individuals, controlling for year- and individual- fixed effects. Including individual-fixed effects allow us to account for individual-level unobservables (such as taste for clean air, individual preferences for a specific city) that may be correlated with migration decisions. A 10% increase in PM2.5 raises out-migration rates by 2.19 percentage points for high-skilled workers.

In our main table, migrants are defined as those who are away from their *hukou* city, so we may miss those who move to a different city and obtain local *hukou* in the city. Thus, we may understate high-skilled workers' migration response to pollution, as migrants with high education attainment find it easier to obtain local *hukou* than those without. We include both non-*hukou* migration (change residential locations without changing *hukou*

Table A11: Individual Longitudinal Panel with Individual Fixed Effects

	Dependent variable: Leave city location indicator								
	Full sample (1)	Low edu (2)	High edu (3)	Full sample (4)	Low edu (5)	High edu (6)	Full sample (7)	Low edu (8)	High edu (9)
Log (PM2.5)	0.115** (0.0558)	0.0812 (0.0542)	0.219** (0.100)	0.0960* (0.0540)	0.0644 (0.0525)	0.193** (0.0964)	0.0774 (0.0539)	0.0385 (0.0538)	0.206** (0.0972)
Observations	122,841	104,184	18,657	122,841	104,184	18,657	122,841	104,184	18,657
Individual FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Weather Controls	No	No	No	Yes	Yes	Yes	No	No	No
Regional trend	No	No	No	No	No	No	Yes	Yes	Yes
Demographics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
F-test of IVs	51.62	43.38	55.60	53.91	43.99	59.93	50.67	42.48	56.85

Notes: Individual-level regressions across 277 cities and between 2008-16. The CLDS asks retrospective histories of residential locations, covering 277 cities. Standard errors clustered at the hukou city level are reported in parentheses. We report Kleibergen-Paap rk Wald F statistic. Instrumental variables specification using the number of thermal inversions. Demographics include age, age-squared. Weather controls include temperature, wind speed, sunshine duration and humidity. Region trend is a region-specific time trend.

locations) and *hukou* migration (change both residential and *hukou* location) in this definition of migration status. The effects on high-skill migration shown in Table A11 are, as expected, larger in magnitude than our baseline estimates.

We control for weather amenities in the next three columns. The coefficient estimates are similar. To account for the potential role played by differential migration patterns between coastal and inland China, we further add region-specific trends in the last three columns. Including region-specific trends do not affect our results.

A.2.2 Accumulated Pollution over Time

As migration decisions are long-lasting, we expect that people are more likely to response to accumulated pollution, compared to contemporaneous pollution shocks. While we measure out-migration in 2015, we wish to understand how migration decisions depend on the cumulative PM2.5 concentration over different time intervals. Since the coal-fired power plants are essentially leveraging the cross-sectional nature of the data, we use the occurrence of thermal inversions averaged over different time periods to deal with the endogeneity of cumulative air pollution.

In Table A12, we use specifications where PM2.5 are averaged over 5, 15 and 18 years, respectively, leading up to the migration decision. We find that the longer the time period of

Table A12: PM2.5 Measured over Different Time Intervals

	Dependent variable: Leave hukou city indicator								
	Full sample (1)	Low edu (2)	High edu (3)	Full sample (4)	Low edu (5)	High edu (6)	Full sample (7)	Low edu (8)	High edu (9)
Log (Mean PM2.5: 1998-2015)	0.116*** (0.0246)	0.107*** (0.0244)	0.178*** (0.0414)						
Log (Mean PM2.5: 2001-2015)				0.109*** (0.0235)	0.100*** (0.0235)	0.171*** (0.0396)			
Log (Mean PM2.5: 2011-2015)							0.0949*** (0.0227)	0.0883*** (0.0238)	0.134*** (0.0310)
Observations	761,548	643,124	118,424	761,548	643,124	118,424	761,548	643,124	118,424
City Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Demographics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Weather Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
F-test of IVs	60.74	59.53	40.59	65.97	64.41	43.95	88.00	84.45	66.93

Notes: Individual level regressions across 332 cities. Standard errors clustered at the hukou city level are reported in parentheses. We report Kleibergen-Paap rk Wald F statistic. Instrumental variables using the number of thermal inversions averaged over different time intervals. City controls include distance to Shanghai, Tianjin and Shenzhen seaports. Demographics include age, age-squared, gender, marital status, and an urban hukou indicator. Weather controls include temperature, humidity, sunshine duration and wind speed averaged over different time intervals.

PM2.5 exposure, the larger the response in small increments. As such, pollution exposure in a short time frame has a slightly smaller impact than the same amount of exposure spread out over a longer time period. Once again, we see similar empirical patterns as our baseline estimates. The effects of cumulative pollution on emigration are also more pronounced for the skilled than the unskilled.

A.2.3 Different Samples and Alternative Measure of Air Quality

We use different samples to explore the relationship between pollution and out-migration. We then use an alternative measure of air quality as our independent variable. First, we examine whether big cities, high polluting cities, or province capitals are driving our results. We exclude such cities one at a time in Panel A and columns 1-6 in Panel B of Table A13. This may help allay concerns about the influence of major cities or capitals in pollution policy, placement of plants, or being outliers in terms of pollutants and/or skilled jobs.

Next, we may expect that coal-fired plants are more likely to be located in coal producing areas, affecting the underlying industrial structure, and raising concerns about other unobservable associations with migration rates. Shanxi is the largest coal producing province in China, which we exclude in our estimation. As reported in columns 7-9 in Panel B of Table A13, the results are slightly more precisely estimated.

Table A13: Without Big Cities, High Polluters and Coal Producing Regions

Panel A:	Exclude Beijing			Exclude Tianjin			Exclude Shijiazhuang		
	Full sample (1)	Low edu (2)	High edu (3)	Full sample (4)	Low edu (5)	High edu (6)	Full sample (7)	Low edu (8)	High edu (9)
Log (PM2.5)	0.0762** (0.0380)	0.0611 (0.0414)	0.131*** (0.0363)	0.138** (0.0544)	0.125** (0.0606)	0.177*** (0.0474)	0.0782** (0.0383)	0.0623 (0.0416)	0.138*** (0.0382)
Observations	752,993	638,529	114,464	745,903	632,324	113,579	758,412	640,631	117,781
City Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Demographics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
F-test of IVs	52.66	46.47	42.14	31.33	27.59	29.11	51.84	46.03	40.10

Panel B:	Exclude Shenyang			Exclude Zhengzhou			Exclude Cities in Shanxi Province		
	Full sample (1)	Low edu (2)	High edu (3)	Full sample (4)	Low edu (5)	High edu (6)	Full sample (7)	Low edu (8)	High edu (9)
Log (PM2.5)	0.0766** (0.0378)	0.0606 (0.0411)	0.136*** (0.0370)	0.0813** (0.0387)	0.0640 (0.0420)	0.153*** (0.0369)	0.0911*** (0.0342)	0.0763** (0.0367)	0.146*** (0.0338)
Observations	758,589	641,096	117,493	754,535	637,980	116,555	722,306	609,605	112,701
City Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Demographics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
F-test of IVs	53.49	47.36	42.45	51.04	45.28	39.94	85.42	78.30	68.94

Notes: Individual level regressions across 331 cities in Panel A and column 1-6 in Panel B (we exclude one big city at a time) and across 321 cities in column 7-9 in Panel B (we drop cities in Shanxi). Standard errors clustered at the hukou city level are reported in parentheses. We report Kleibergen-Paap rk Wald F statistic. Instrumental variables specification using the interaction between wind direction, distance to coal plant, and coal consumption at power plant. City controls include distance to Shanghai, Tianjin and Shenzhen seaports, and inverse distance weighted average of coal consumption for power plants within 500 km of a city. Demographics include age, age-squared, gender, marital status, and urban hukou indicator.

In Table A14, we split up the sample by age groups and marital status. We see that the implications of pollution on emigration are stronger for younger workers with higher education. Panel B further shows that married individuals are more responsive, perhaps as they may have young children that may be affected by poor air quality.

Finally, we turn our attention to an alternative measure of air quality. As sources of pollution affect not just PM2.5 but also other pollutants, we may pick up the combined impact of many pollutants. Air Quality Index (AQI) is an overall indicator for air pollution concentration calculated using multiple atmospheric pollutants including SO₂, NO₂, PM₁₀, PM_{2.5}, O₃ and CO. Furthermore, the AQI is officially reported and widely disseminated. Table A15 shows similar results using the AQI as our independent variable of interest.

Table A14: By Age Groups

	Dependent variable: Leave hukou city indicator							
	Age<35		Age>=35		Age<35&Married		Age<35&Unmarried	
	Low edu (1)	High edu (2)	Low edu (3)	High edu (4)	Low edu (5)	High edu (6)	Low edu (7)	High edu (8)
Log (PM2.5)	0.0656 (0.0611)	0.190*** (0.0493)	0.0567 (0.0359)	0.0673*** (0.0223)	0.0931 (0.0719)	0.251*** (0.0699)	0.00454 (0.0427)	0.138*** (0.0363)
Observations	199,958	70,643	443,166	47,781	133,002	33,590	66,956	37,053
City Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Demographics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
F-test of IVs	33.93	40.62	49.92	40.29	34.38	36.57	30.91	42.88

Notes: Individual level regressions across 332 cities. Standard errors clustered at the hukou city level are reported in parentheses. We report Kleibergen-Paap rk Wald F statistic. Instrumental variables specification using the interaction between wind direction, distance to coal plant, and coal consumption at power plant. City controls include distance to Shanghai, Tianjin and Shenzhen seaports, and inverse distance weighted average of coal consumption for power plants within 500 km of a city. Demographics include age, age-squared, gender, marital status, and an urban hukou indicator.

Table A15: Air Quality Index

	Dependent variable: Leave hukou city indicator		
	Full sample	Low edu	High edu
	(1)	(2)	(3)
Log (AQI)	0.0745 (0.0628)	0.0531 (0.0650)	0.170** (0.0705)
Observations	708,482	595,958	112,524
City Controls	Yes	Yes	Yes
Demographics	Yes	Yes	Yes
F-test of IVs	36.58	35.26	30.60

Notes: Individual level regressions across 238 cities that report AQI in 2014. Independent variable is Log (Annual mean Air Quality Index)([AQI, 2014](#)). Standard errors clustered at the hukou city level are reported in parentheses. We report Kleibergen-Paap rk Wald F statistic. Instrumental variables specification using the interaction between wind direction, distance to coal plant, and coal consumption at power plant. City controls include the log distance to Shanghai seaport, to Tianjin seaport, and to Shenzhen seaport, and inverse distance weighted average of coal consumption for power plants within 500 km of a city. Demographics include age, age-squared, gender, marital status, and an urban hukou indicator.

A.2.4 Additional Controls

In this section, we include various sets of controls that may confound the association between local air pollution and migration decisions. In [Table A16](#) we examine robustness to additional geological and meteorological controls when using the thermal inversions in-

struments. These include measures of rainfall, average slope (elevation) of the city, fraction of the region covered by water bodies, and whether the city is coastal or not.

In Table A18 we test the validity of the wind IV to other economic and industrial controls. In the first few columns of Table A18 Panel A, we account for other determinants of the demand for skilled work at baseline. We do this to check whether the potential for skilled work just so happens to be in places that are correlated with skill-specific migration. We find our estimates display similar patterns as before if we add controls for the teacher-student ratio, the number of hospitals and doctors per capita in 2000.

Table A16: Additional Weather and Geological Controls: For Thermal Inversions

	No. of Inversions			Strength of Inversions		
	Full Sample (1)	Low edu (2)	High edu (3)	Full Sample (4)	Low edu (5)	High edu (6)
Log(PM 2.5)	0.117*** (0.0443)	0.115** (0.0469)	0.155*** (0.0483)	0.0880* (0.0514)	0.0842 (0.0541)	0.151** (0.0599)
Observations	685,699	575,693	110,006	685,699	575,693	110,006
City Controls	Yes	Yes	Yes	Yes	Yes	Yes
Weather Controls	Yes	Yes	Yes	Yes	Yes	Yes
Demographics	Yes	Yes	Yes	Yes	Yes	Yes
F-Test of ivs	37.10	34.22	35.23	13.80	13.06	13.13

Notes: Individual level regressions across 261 cities (for which we have measures of the controls). Standard errors clustered at the hukou city level are reported in parentheses. We report Kleibergen-Paap rk Wald F statistic. Instrumental variables specification using thermal inversions. City controls include distance to Shanghai, Tianjin and Shenzhen seaports. Weather controls include temperature, wind speed, sunshine duration and humidity. Demographics include age, age-squared, gender, marital status, and an urban hukou indicator. In addition, we control for annual rainfall, average slope (elevation) of the city, fraction of the region area under water coverage, and whether on the coast. Data on average slope and the fraction of area under water coverage come from [Xu et al. \(2018\)](#) and [USGS \(2010\)](#). We manually collect the data on whether a city is on the coast ([Khanna et al., 2024](#)).

In the next three columns of Table A18 Panel A, we include controls for local economic production in 2000. Our economic controls include baseline measures of GDP per capita, as well as the ratio of product values of services and manufacturing as proxies for the industrial structure. Our results are not meaningfully affected by these controls.

Fine particle concentration tends to be correlated with local industrial pollutant emissions. To account for the potential role played local industrial emissions, we add industrial SO2 emission, waste water emission and dust emission as covariates in last three columns of Table A18 Panel A. The inclusion of these industrial emissions does little to affect the impacts of PM2.5 concentration.

Table A17: Additional Weather and Geological Controls: For House Price Regressions

	Log(House Prices) (1)	Log(House Prices) (2)	Log(House Prices) (3)
$\text{Log}(\text{Population})_d$	0.259** (0.121)	0.314** (0.147)	0.204* (0.119)
$\text{Log}(L_{hd}/L_{ud})$	0.423*** (0.0583)	0.407*** (0.0690)	0.456*** (0.0556)
Observations	121	121	117
Weather Control	No	Yes	Yes
Geological Control	No	No	Yes
F-Test of ivs	12.66	13.14	11.73

Notes: City-level regressions, and robust standard errors. We report Kleibergen-Paap rk Wald F statistic. Instrumental variables specifications, where we instrument for log population and log skill ratio as described in the text and Table 5. City controls include distance to Shanghai, Tianjin and Shenzhen seaports. Weather controls include temperature, wind speed, sunshine duration and humidity. Geographic controls include average slope (elevation) of the city, fraction of the region area under water coverage, and whether on the coast.

In the first three columns of Panel B, we start with all three sets of controls that we have in Panel A entered simultaneously. After which, we address the concern that power plants may be built near cities that require more electricity. Even though those plants supply electricity to vast areas including many remote provinces (Freeman et al., 2019), we examine this concern by controlling for city-level electricity consumption. We include total (industrial, commercial and residential) electricity consumption in the next three columns of Table A18 Panel B, and add industrial electricity consumption in the last three columns. Adding city-level electricity consumption barely changes our empirical patterns.

Table A18: Additional Controls: Skills, Economic and Industrial Emissions

Panel A									
Additional Controls:	Dependent variable: Leave hukou city indicator								
	Baseline skill controls			Baseline economy controls			Industrial emissions controls		
	Full Sample	Low edu	High edu	Full Sample	Low edu	High edu	Full Sample	Low edu	High edu
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Log (PM2.5)	0.0677* (0.0406)	0.0542 (0.0436)	0.124*** (0.0412)	0.0420 (0.0397)	0.0263 (0.0424)	0.0997** (0.0434)	0.0532 (0.0404)	0.0382 (0.0436)	0.118*** (0.0357)
Observations	674,032	565,239	108,793	674,032	565,239	108,793	674,032	565,239	108,793
City Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Demographics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
F-test of IVs	53.82	51.26	32.84	50.63	48.07	33.28	58.99	57.31	36.04
Panel B									
Additional Controls:	Dependent variable: Leave hukou city indicator								
	All 3 sets of controls			All 3 sets + Total elec cons			All 3 + Industry elec cons		
	Full Sample	Low edu	High edu	Full Sample	Low edu	High edu	Full Sample	Low edu	High edu
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Log (PM2.5)	0.0765** (0.0340)	0.0683* (0.0359)	0.111*** (0.0330)	0.0669* (0.0364)	0.0600 (0.0387)	0.0984*** (0.0342)	0.0658* (0.0353)	0.0586 (0.0375)	0.0996*** (0.0334)
Observations	674,032	565,239	108,793	650,828	544,681	106,147	650,828	544,681	106,147
City Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Demographics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
F-test of IVs	57.86	57.16	33.74	52.91	51.93	31.44	54.04	53.46	31.21

Notes: Individual level regressions across 261 cities in Panel A, and columns 1-3 of Panel B (cities for which we have measures of the controls). Individual-level regressions across the 253 cities for which we have electricity consumption data and the full set of city controls, in columns 4-9 of Panel B. Standard errors clustered at the hukou city level are reported in parentheses. We report Kleibergen-Paap rk Wald F statistic. Instrumental variables specification using the interaction between wind direction, distance to coal plant, and coal consumption at power plant. Baseline skill controls include teacher student ratio, log hospitals per capita and log doctors per capita in 2000. Baseline economy controls include log GDP per capita and industrial structure (the product value at service sector / manufacture sector). Industrial emissions controls include log industrial SO2 emission, log industrial waste water emission and log industrial dust emission. City controls include distance to Shanghai, Tianjin and Shenzhen seaports, and inverse distance weighted average of coal consumption for power plants within 500 km of a city. Demographics include age, age-squared, gender, marital status, and an urban hukou indicator. Total electricity consumption includes industrial, residential and commercial consumption.

B Air Pollution Data Disclosure in China

Despite hazardous levels of pollution, Chinese citizens used to have limited or no access to information about local air quality. In 2008 the US embassy published PM2.5 data from five Chinese cities, leading to a public attention on air quality. In response to public demand for the publication of PM2.5 data, the Chinese government started to disclose real time PM2.5 data in most Chinese cities from 2012. Information on real time PM2.5 was made available in all Chinese cities by January 1, 2015.

Table B1: PM2.5 Data Disclosure and Outmigration

Dependent variable:	Leave city location indicator					
	Full sample (1)	Low edu (2)	High edu (3)	Full sample (4)	Low edu (5)	High edu (6)
$\text{Log}(\text{PM2.5})_{ot} \times \text{Data disclosed indicator}_{ot}$	0.000902 (0.00147)	-0.000585 (0.00139)	0.0106** (0.00495)	0.000997 (0.00143)	-0.000482 (0.00130)	0.0105** (0.00488)
Observations	122,841	104,184	18,657	122,841	104,184	18,657
Individual FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Demographics	Yes	Yes	Yes	Yes	Yes	Yes
Regional trend	No	No	No	Yes	Yes	Yes

Notes: Individual-level regressions across 277 cities. The CLDS lists the history of residential locations, which include 277 cities. Standard errors clustered at the city level. We report Kleibergen-Paap rk Wald F statistic. Demographics include age, age-squared. We control for the log of PM2.5 and an indicator for PM2.5 data disclosure. Data disclosure indicator=1 if PM2.5 data was officially made available in a given city-year, =0 otherwise.

The disclosure of pollution information had an important effect on household avoidance behavior. The sales of indoor air filtration more than doubled in response to the PM2.5 data disclosure in 2012. Table B1 shows that PM2.5 data disclosure had significant impacts on the out-migration elasticity of the high-skilled.

C Additional Tables and Figures

Table C1: Examples of the Points-based *Hukou* Policy Across Chinese Cities

City	Beijing	Shanghai	Guangzhou	Shenzhen
Total hukou points needed	Varies	72	60	100
Education	Doctoral degree: 37 points Master degree: 26 points Bachelor degree: 15 points Some college: 10.5 points	Doctoral degree: 27 points Master degrees: 24 points Bachelor degree: 21 points	Above college: 60 points Some college: 40 points High school: 40 points	Doctoral degree: 100 points Master degrees: 90 points Bachelor degree: 80 points College: 60 points
Skills		College English Test 6-8: 8 points College English Test 4: 7 points	Junior workers: 10 points Middle-level workers: 30 points High-level workers: 50 points	Junior workers: 20 points Middle-level workers: 40 points High-level workers: 70 points Senior technical worker: 100 points Junior professional: 70 points Middle professional: 90 points Senior professional: 100 points

Note: Table shows a few examples of *hukou* requirements for city workers. Data come from the official government websites of the Beijing, Shanghai, Guangzhou and Shenzhen. ‘Total *hukou* points needed’ show the number of points required to obtain local *hukou*, whereas the lower rows show how different qualifications add to the individual’s points.

Table C2: Preferences for Environmental Issues by Education Levels

	Claim environmental issue is terrible (1)	Discuss environmental issues (2)	Donation for environment protection (3)	Concern over environmental issue (4)	Appeal on Environmental issue (5)	Government environmental activity (6)	Non-government environmental activity (7)
High school	0.130*** (0.0116)	0.135*** (0.0173)	0.0592*** (0.0115)	0.121*** (0.0153)	0.0252*** (0.00841)	0.121*** (0.0138)	0.0691*** (0.0116)
Some college or above	0.173*** (0.0118)	0.233*** (0.0190)	0.152*** (0.0154)	0.176*** (0.0191)	0.0571*** (0.0120)	0.248*** (0.0168)	0.157*** (0.0143)
P-value	0.001	0.000	0.000	0.000	0.010	0.000	0.000
t-value	-3.20	-6.10	-5.65	-3.58	-2.62	-8.16	-6.25
Baseline average	0.497	0.376	0.106	0.389	0.062	0.133	0.100
City FE	Y	Y	Y	Y	Y	Y	Y
Demographics	Y	Y	Y	Y	Y	Y	Y
N	29169	11331	11331	11331	11331	11331	11331
adj. R2	0.085	0.202	0.182	0.202	0.186	0.196	0.203

Note: Individual level regressions across 113 cities in column 1 and 127 cities in column 2-7. Demographic controls include age, age-squared, gender, marital status and an indicator for urban *hukou*. The data source is the China Household Panel Survey (CFPS, 2016) in column 1 and the Chinese General Social Survey (CGSS, 2013) in column 2-7. In the CFPS, there is a survey question: In your opinion, how terrible the environment issue is in China. (0=totally not terrible; 2,...,10=very terrible). Based on this question, we define environmental attitude dummy: D=1, if the answer is 6-10; =0, if the answer is 0-5. In the CGSS, there is a survey question: whether you participate in the following activity. 1=never, 2=occasionally; 3=often. We define an indicator: D=1 if the answer=2,3; D=0 if the answer=1. P-value: the p-values of test of Some college or above=High school; t-value: t-values of test of Some college or above=High school. Standard errors clustered at the city level are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table C3: First-stage Relationships for Additional Model-based Parameter Estimation

Panel A: For Labor Supply		Log (Unskilled Real Wage)		Log (Skilled Real Wage)	
		(1)	(2)	(3)	(4)
NTR IV		0.161*** (0.0161)	0.162*** (0.0153)		
WID IV				0.00222*** (0.000123)	0.00387*** (0.000111)
Pollution IV	No. of Inversions		Strength of Inversions	No. of Inversions	Strength of Inversions
Observations	13,570		13,570	13,570	13,570
R-squared	0.132		0.132	0.136	0.153
Panel B:		For Agglomeration		For House Prices & Pollution	
		Log (Number of Skilled Workers)		Log(Skill Ratio)	Log(Population)
		(1)	(2)	(3)	(4)
$\Delta \text{College graduates}_{2001-5}$		0.400*** (0.0386)	0.395*** (0.0390)		
$\Delta \text{College graduates}_{2001-5} / \text{Baseline } L_{it,2001}$				0.835*** (0.130)	0.127 (0.0915)
Predicted migration flow				0.00158 (0.00148)	0.00425*** (0.00102)
Pollution IV	No. of Inversions		Strength of Inversions		
Observations	121		121	121	121
R-squared	0.672		0.665	0.573	0.546

Notes: We control for number of inversions in Column 1 and 3, and strength of inversions in Column 2 and 4. All regressions weighted by the population in 2000. Panel A: We control for hukou index, the interaction of hukou index to migration status indicator, the inverse hyperbolic sine of distance from origin to destination cities. We also control for temperature, humidity, sunshine duration, and wind speed. Standard errors clustered at the city level are reported in parentheses. Panel B: We control for distance to seaport, region fixed effects, city area and weather amenities (temperature, humidity, sunshine duration and wind speed). Robust standard errors are reported in parentheses. We report Kleibergen-Paap rk Wald F statistic.

Table C4: Regional Balance Tests for NTR IV and WID IV

Dependent Variable	NTR IV	NTR IV	WID IV	WID IV
Log Population	0.00491 (0.00704)	0.00449 (0.00756)	0.135 (1.154)	0.0493 (1.264)
Log Non-agricultural population	-0.000809 (0.0066)	-0.000879 (0.00694)	1.259 (1.228)	1.242 (1.243)
Log Share of workers in secondary sector	-0.00642 (0.00912)	-0.0054 (0.00978)	0.969 (1.084)	0.594 (1.122)
Log Share of workers in tertiary sector	-0.0116 (0.00853)	-0.0122 (0.00981)	-0.714 (1.261)	-0.515 (1.372)
City GDP	9.42E-11 (6.58E-11)	1.00E-10 (6.63E-11)	1.47E-08 (1.11E-08)	1.20E-08 (1.15E-08)
Log PM2.5	-0.000941 (0.00619)	-0.000316 (0.00696)	0.206 (0.968)	0.506 (1.126)
Log Area	-0.00554 (0.0036)	-0.00492 (0.00406)	-0.38 (0.515)	-0.253 (0.565)
Export licenses		0.00405 (0.0678)		4.973 (8.106)
Contract intensity		-0.0229 (0.103)		10.12 (10.84)
Log Distance to big cities	-2.65E-05 (2.37E-05)	-2.95E-05 (2.56E-05)	-0.00308 (0.00424)	-0.00269 (0.00428)
P-value for baseline attributes	0.594	0.801	0.418	0.434
Observations	108	105	108	105
Adj R-squared	0.162	0.139	0.389	0.389
Weather Controls	Yes	Yes	Yes	Yes

Notes: We use the sample cities for our bilateral migration regressions in Table 4 , and combine data on city attributes from China City Statistical Year Books, data on contract intensity from Nunn (2007) and data on export licenses from Bai et al. (2017). Robust standard errors are reported in parentheses. $*p < 0.10$, $**p < 0.05$, $***p < 0.01$

Table C5: Industry-Level Balance Test for WID IV and NTR IV

Dependent Variable	Sector-specific NTR Gap		Sector-specific WID	
Labor over value-added	0.00795 (0.0313)	-0.00227 (0.0352)	1.216 (2.971)	-3.187 (6.726)
Capital over value-added	-0.000431 (0.00064)	-0.000359 (0.000524)	-0.0182 (0.0399)	0.0749 (0.0736)
Contract intensity	0.098 (0.0727)	0.0156 (0.0718)	-8.52 (14.25)	-7.533 (13.36)
Wage		-0.619 (0.531)		-4.665 -65.63
Log Input Value		-0.472 (0.392)		33.79 (29.8)
Log Output Value		0.468 (0.394)		-33.52 (29.84)
Log Non-processing export		0.00904 (0.0142)		-4.89 (3.429)
Log Processing export		0.0183 (0.0125)		0.0278 (1.338)
Observations	111	111	52	52
Adj R-squared	-0.006	0.124	-0.051	0.109

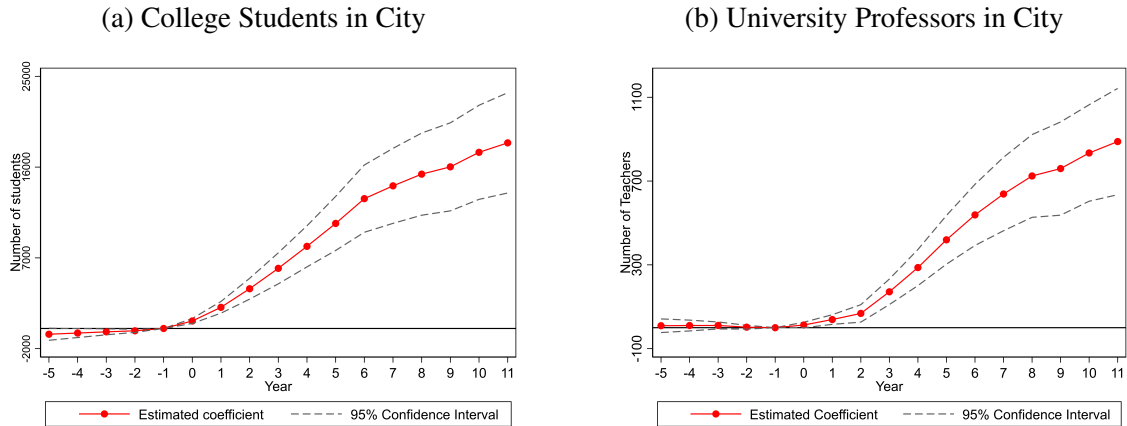
Notes: Data on contract intensity from [Nunn \(2007\)](#). Data on all other independent variables come from the Annual Survey of Industrial Production (ASIP). Robust standard errors are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Figure C1: The Distribution in Pollution and Wages Across Cities



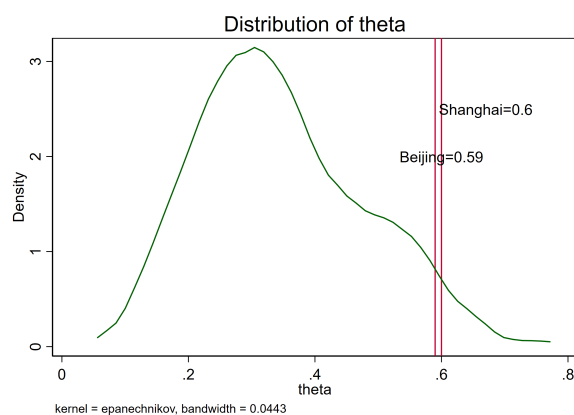
Notes: Distributions across cities for different years. Wage distribution drawn from the City Statistical Yearbooks. PM2.5 data from the Global Annual PM2.5 grids.

Figure C2: Event Study of the Number of College Students by Baseline Propensity to Expand



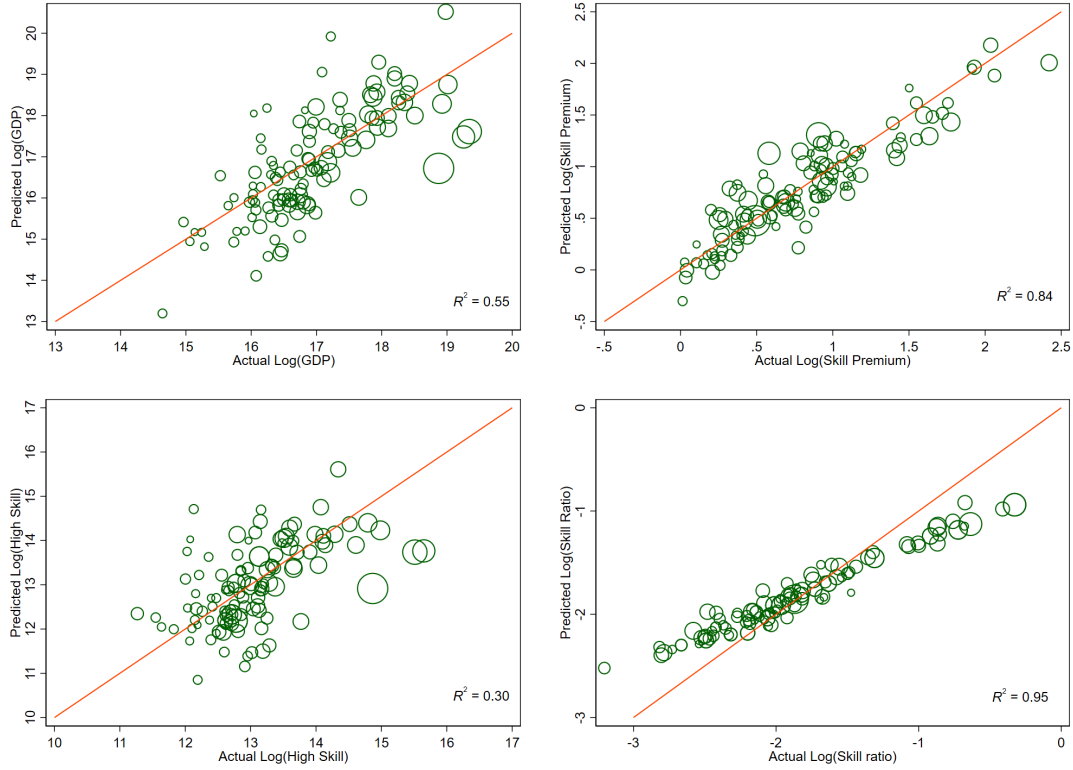
Notes: We test for pre-trends and dynamics of the college expansion policy in an event study framework. City-by-year level regressions from 1994 to 2010. Outcome is number of new college students, relative to one year preceding the expansion (1998). In Figure C2a, we run the regression: $Students_{dt} = \beta_0 + \sum_{t=1994}^{2010} \beta_t (1_{t=1998} \times Treat_{d,1990}) + \gamma_t + \delta_c + \epsilon_{dt}$. In Figure C2b the dependent variable is the number of professors in the city. We plot β_t . $Treat_{d,1990}$ is the number of university students in 1990. Horizontal axis is normalized to the year preceding the expansion (1998).

Figure C3: Distribution of θ_{hd} across cities, from Equation 18



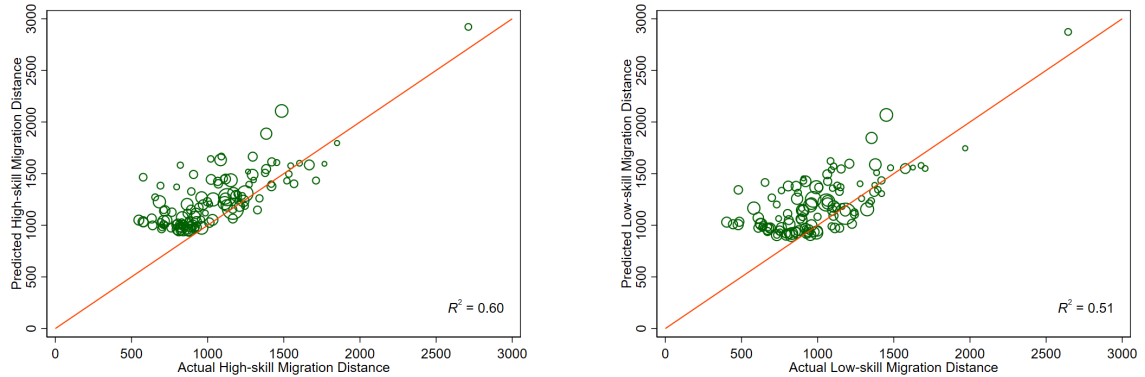
Graph describes the distribution of θ_{hd} across cities. Equation 18 describes how we estimate θ_{hd} .

Figure C4: Model Fit in 2015



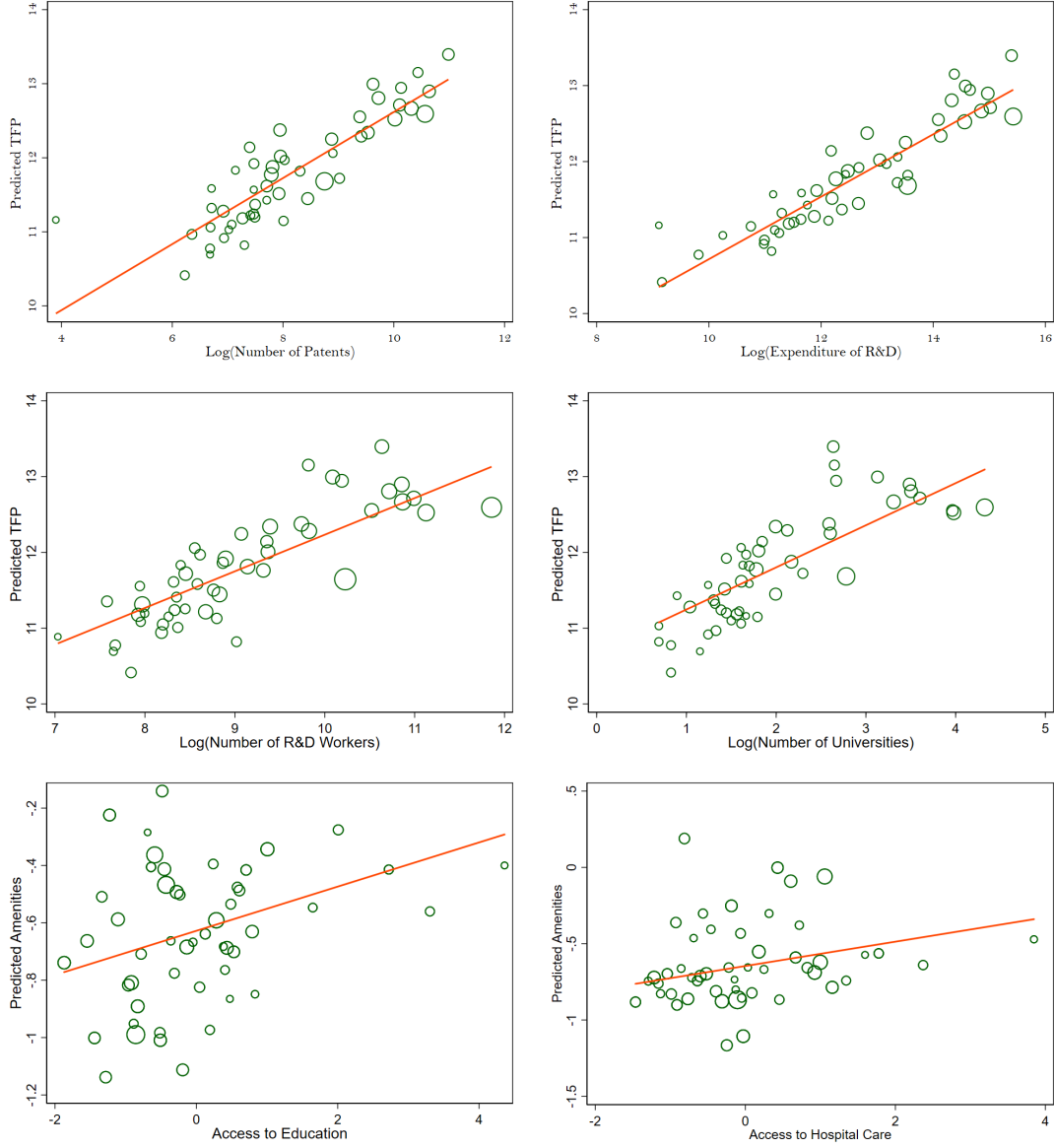
Notes: We plot the actual and predicted relationship between our main variables, where the predictions are based on model-estimated parameters. Bubbles are weighted by populations in the year 2000. We exclude the residual ξ_{sod} in our model predictions.

Figure C5: Out of Sample Test of Migration Distance



Notes: We plot the relationship between predicted and actual migration distance for migrants who move out from their *hukou* regions, where the predictions are based on model-estimated parameters. Bubbles are weighted by populations in the year 2000. We exclude the residual ξ_{sod} in our model predictions.

Figure C6: Correlates of Predicted TFP and Predicted Amenities



Notes: We plot the relationship between predicted TFP and various innovation measures in the top four figures, and plot the relationship between predicted amenities and access to education and between predicted amenities and access to hospital care in the bottom two figures. Access to education is measured by the first principal component of the number of schools per capita and the number of teachers per capita. Access to hospital care is measured by the first principal component of the number of hospitals per capita and the number of doctors per capita. The raw data for innovation measures, access to education and access to hospital care come from China Statistical Yearbooks (CCSY, 2016, 2018). Bubbles are weighted by populations in the year 2000. We exclude the residual ξ_{sod} in our model predictions.

D Additional Counterfactual Results

Table D1: Additional Counterfactual Results: Reducing Pollution in Beijing

	Panel A: GDP per Worker in Nearby Cities (within 300km from Beijing) (%)			
	Overall Change	Health Channel	Relocation	Relocation+Agglom
	(1)	(2)	(3)	(4)
Reduce steady state PM2.5	-0.072	0.000	-0.025	-0.072
Relax skilled hukou	-0.061	0.000	-0.023	-0.061
Relax unskilled hukou	-0.008	0.000	0.005	-0.008
Reduce PM2.5 & relax skilled hukou	-0.161	0.000	-0.060	-0.161
Reduce PM2.5 & relax unskilled hukou	-0.086	0.000	-0.022	-0.086
	Panel B: GDP per Worker in Beijing Under No Externalities or Pure Relocation (%)			
	Overall Change (no externalities)	Health Channel (no externalities)	Relocation (no externalities)	Relocation (pure relocation)
	(1)	(2)	(3)	(4)
Reduce steady state PM2.5	10.448	5.907	4.288	4.102
Relax skilled hukou	4.920	0.000	4.920	4.920
Relax unskilled hukou	-4.649	0.000	-4.649	-4.649
Reduce PM2.5 & relax skilled hukou	15.187	5.907	8.763	8.604
Reduce PM2.5 & relax unskilled hukou	5.797	5.907	-0.104	-0.304

Notes: In panel A, we replicate our exercises in Table 7 and present the effects on nearby cities. Panel A column 1 shows the gain to overall GDP per worker. Panel A column 2 shows the component purely explained by the health-productivity channel. Panel A column 3 shows the pure relocation channel, and Panel A column 4 also incorporates agglomeration as a consequence of relocation. In panel B columns 1-3, we do additional exercises, shutting down agglomeration and congestion effects. Panel B Column 1 shows the gain to overall GDP per worker. Panel B column 2 shows the component purely explained by the health-productivity channel. Panel B column 3 shows the pure relocation channel. In panel B column 4 we further shut down how the health effects of pollution may affect relocation choices. We reduce the steady state amount of pollution in Beijing by 50% (row 1 in panels A and B). Next, we lower *hukou* restrictions for each skill level (rows 2 and 3 in panels A and B) by 50%. Finally, we lower the *hukou* regulations by 50% while reducing steady state pollution (rows 4 and 5 in panels A and B).

Table D2: Distributional Consequences of Reducing Pollution in One City

	Skilled Wage in Beijing			Unskilled Wage in Beijing		
	Overall	Health	Relocate+Agglom	Overall	Health	Relocate+Agglom
	(1)	(2)	(3)	(4)	(5)	(6)
Reduce steady state PM2.5	2.807	5.907	-2.928	18.505	5.907	11.895
Reduce exogenous part of PM2.5	2.679	5.421	-2.601	16.903	5.421	10.892
Relax skilled hukou	-4.329	0.000	-4.329	12.587	0.000	12.587
Relax unskilled hukou	7.688	0.000	7.688	-6.434	0.000	-6.434
Reduce PM2.5 & relax skilled hukou	-1.483	5.907	-6.978	33.521	5.907	26.074
Reduce PM2.5 & relax unskilled hukou	10.585	5.907	4.417	10.721	5.907	4.545

Notes: This table shows the corresponding wage effects of the policy exercises in our main Table 7. We reduce the steady state amount of pollution in Beijing by 50% (row 1). We then reduce only the exogenous component of pollution by 50% (row 2). Next, we lower the *hukou* restrictions for each skill level (rows 3 and 4) by 50%, keeping pollution fixed. Finally (rows 5 and 6) we relax *hukou* regulations by 50% while reducing steady state pollution. The first 3 columns show the effect on the wage of college educated workers, whereas the last 3 columns show the effects on the wage of the non-college educated.

Table D3: The Productivity Effect of Relocating Pollution (no externalities and pure relocation)

	Change in GDP per Worker (%)			
	Overall changes (no externalities)	Health (no externalities)	Relocation (no externalities)	Relocation (pure relocation)
	(1)	(2)	(3)	(4)
Relocate steady state PM2.5	7.712	2.963	3.155	2.360
Relax hukou	2.318	0.000	2.318	2.318
Relax overall mobility constraints	6.314	0.000	6.314	6.314
Relocate PM2.5 & relax hukou	10.005	2.963	5.276	4.514
Relocate PM2.5 & lower migration costs	14.308	2.963	9.158	8.443

Notes: In this table, we do an additional exercise, shutting down agglomeration and congestion effects. Column 1 shows the overall gain to GDP. Column 2 shows the increase in GDP as a consequence of the health effects only. Column 3 shows the gain due to the re-allocation of labor channel only. In column 4 we further shut down how the health effects of pollution may affect relocation choices. We relocate PM2.5 in all cities based on the amount of skill-biased capital in the city (row 1). In row 2, we relax the *hukou* restriction in the 24 top tier cities by 50% (row 2). In row 3 we relax overall migration costs in the 24 high tier cities by 50%.

Table D4: Distributional Effects of Relocating Pollution Across Cities

	Skilled Wage			Unskilled Wage		
	Overall (1)	Health (2)	Relocate+Agglom (3)	Overall (4)	Health (5)	Relocate+Agglom (6)
Relocate steady state PM2.5	17.948	4.586	10.581	1.169	1.635	-1.890
Relocate exogenous part of PM2.5	9.030	2.451	5.737	1.042	1.042	-0.446
Relax hukou	3.871	0.000	3.871	1.936	0.000	1.936
Relax overall mobility constraints	10.124	0.000	10.124	5.362	0.000	5.362
Relocate PM2.5 & relax hukou	21.394	4.586	13.871	3.111	1.685	-0.113
Relocate PM2.5 & lower mig costs	28.107	4.4586	20.034	6.787	1.685	3.164

Notes: This table shows the corresponding wage effects of the policy exercises in our main Table 8. We relocate PM2.5 in all cities based on the amount of skill-biased capital in the city (row 1). In row 2, we relocate only the exogenous component of pollution. In addition to such relocations of pollution, we also relax the *hukou* restriction in the 24 top tier cities by 50% (row 3). In row 4 we relax overall migration costs in the 24 high tier cities by 50%. Columns 1-3 show the effects on skilled workers, while columns 4-6 show the effects on unskilled workers.

Table D5: Welfare Effects of Relocating Pollution Across Cities

	Skilled Welfare (1)	Unskilled Welfare (2)	Average Welfare (3)
Relocate steady state PM2.5	29.922	2.234	4.630
Relocate exogenous part of PM2.5	13.245	1.317	2.307
Relax hukou	9.919	0.920	1.698
Relax overall mobility constraints	17.523	10.007	10.658
Relocate PM2.5 & relax hukou	44.280	3.213	6.767
Relocate PM2.5 & lower migration costs	55.826	13.231	16.917

Notes: This table shows the corresponding welfare effects of the policy exercises in our main Table 8. We relocate PM2.5 in all cities based on the amount of skill-biased capital in the city (row 1). In row 2, we relocate only the exogenous component of pollution. In addition to such relocations, we also relax *hukou* restrictions in the 24 top tier cities by 50% (row 3). In row 4 we relax overall migration costs in the 24 high tier cities by 50%. Rows 3 and 4 keep pollution fixed. Rows 5 and 6 relocate pollution while relaxing migration costs.

Table D6: Distributional Effects of Pollution Regulation

	Skilled Wage			Unskilled Wage		
	Overall (1)	Health (2)	Relocate+Agglom (3)	Overall (4)	Health (5)	Relocate+Agglom (6)
Control PM2.5	3.539	1.508	1.905	3.651	1.560	1.979
Control PM2.5 & relax hukou	7.630	1.508	5.860	5.699	1.560	3.947
Control PM2.5 & lower mig costs	14.064	1.508	12.185	9.171	1.560	7.368

Notes: This table shows the corresponding wage effects of the policy exercises in our main Table 9. We reduce pollution according to the targets set by the 2013 *Air Pollution Prevention and Control Plan* (row 1). In addition to pollution regulations, we also relax the hukou restriction (row 2) and migration costs (row 3) in higher tier cities by 50%. Columns 1-3 show effects on skilled wages. Columns 4-6 on unskilled wages.

Table D7: Welfare Effects of Pollution Regulation

	Skilled Welfare (1)	Unskilled Welfare (2)	Average Welfare (3)
Control PM2.5	7.179	1.095	1.622
Control PM2.5 & relax hukou	18.485	1.971	3.400
Control PM2.5 & lower migration costs	26.775	11.189	12.538

Notes: This table shows the corresponding welfare effects of the policy exercises in our main Table 9. We reduce pollution according to the 2013 *Air Pollution Prevention and Control Plan* (row 1). In addition to pollution regulations, we also relax the hukou restriction (row 2) and migration costs (row 3) in higher tier cities by 50%.

Table D8: Reducing Pollution in Beijing: Changes in GDP per worker (%)

	Panel A: Use Tombe and Zhu (2019) elasticity for all workers			
	Elasticity=1.19	Elasticity=1.4	Elasticity=1.5	Elasticity=1.61
Reduce steady state PM2.5	14.118	13.784	13.667	13.562
Reduce exogenous part of PM2.5	12.495	12.114	11.971	11.835
Relax skilled hukou	7.460	7.989	8.212	8.438
Relax unskilled hukou	-4.356	-4.357	-4.330	-4.283
Reduce PM2.5 & relax skilled hukou	21.983	22.234	22.369	22.525
Reduce PM2.5 & relax unskilled hukou	9.670	9.338	9.246	9.186
	Panel B: Use Tombe and Zhu (2019) elasticity for skilled workers			
	Elasticity=1.19	Elasticity=1.4	Elasticity=1.5	Elasticity=1.61
Reduce steady state PM2.5	14.507	14.517	14.528	14.541
Reduce exogenous part of PM2.5	12.896	12.884	12.883	12.886
Relax skilled hukou	7.737	8.574	8.931	9.294
Relax unskilled hukou	-3.948	-3.601	-3.450	-3.293
Reduce PM2.5 & relax skilled hukou	22.646	23.535	23.924	24.324
Reduce PM2.5 & relax unskilled hukou	10.487	10.865	11.034	11.212
	Panel C: Use Tombe and Zhu (2019) elasticity for unskilled workers			
	Elasticity=1.19	Elasticity=1.4	Elasticity=1.5	Elasticity=1.61
Reduce steady state PM2.5	14.144	13.738	13.554	13.359
Reduce exogenous part of PM2.5	12.505	12.088	11.904	11.711
Relax skilled hukou	7.919	7.603	7.456	7.298
Relax unskilled hukou	-4.159	-4.527	-4.664	-4.787
Reduce PM2.5 & relax skilled hukou	22.494	21.774	21.444	21.092
Reduce PM2.5 & relax unskilled hukou	9.906	9.112	8.783	8.459

Notes: In this counterfactual we reduce the steady state amount of pollution in Beijing by 50% (row 1). We then reduce only the exogenous component of pollution by 50% (row 2). Next, we lower the hukou restrictions for each skill level (rows 3 and 4) by 50%, keeping pollution fixed. Finally (rows 5 and 6) we relax hukou regulations by 50% while reducing steady state pollution.

Table D9: Relocating Pollution for China: Changes in GDP per worker (%)

	Panel A: Use Tombe and Zhu (2019) elasticity for all workers			
	Elasticity=1.19	Elasticity=1.4	Elasticity=1.5	Elasticity=1.61
Relocate steady state PM2.5	6.534	5.177	4.485	3.700
Relocate exogenous part of PM2.5	3.672	3.049	2.727	2.357
Relax hukou	2.796	3.614	4.059	4.593
Relax overall mobility constraints	7.564	9.253	10.115	11.111
Relocate PM2.5 & relax hukou	9.217	8.584	8.297	8.003
Relocate PM2.5 & lower migration costs	14.325	14.456	14.531	14.641
	Panel B: Use Tombe and Zhu (2019) elasticity for skilled workers			
	Elasticity=1.19	Elasticity=1.4	Elasticity=1.5	Elasticity=1.61
Relocate steady state PM2.5	7.103	6.636	6.434	6.228
Relocate exogenous part of PM2.5	3.889	3.629	3.517	3.402
Relax hukou	2.455	2.705	2.811	2.921
Relax overall mobility constraints	6.713	7.187	7.384	7.584
Relocate PM2.5 & relax hukou	9.456	9.169	9.045	8.920
Relocate PM2.5 & lower migration costs	14.068	13.946	13.888	13.828
	Panel C: Use Tombe and Zhu (2019) elasticity for unskilled workers			
	Elasticity=1.19	Elasticity=1.4	Elasticity=1.5	Elasticity=1.61
Relocate steady state PM2.5	6.253	5.435	5.012	4.527
Relocate exogenous part of PM2.5	3.515	3.197	3.034	2.848
Relax hukou	2.971	3.422	3.638	3.875
Relax overall mobility constraints	7.890	8.912	9.376	9.871
Relocate PM2.5 & relax hukou	9.071	8.689	8.474	8.219
Relocate PM2.5 & lower migration costs	14.291	14.447	14.466	14.454

Notes: In this counterfactual exercise we relocate PM2.5 in all cities based on the amount of skill-biased capital in the city (row 1). In row 2, we relocate only the exogenous component of pollution. In row 3, we relax hukou restrictions in the 24 top tier cities by 50%, keeping pollution fixed. In row 4 we relax overall migration costs to the 24 top tier cities by 50%, keeping pollution fixed.

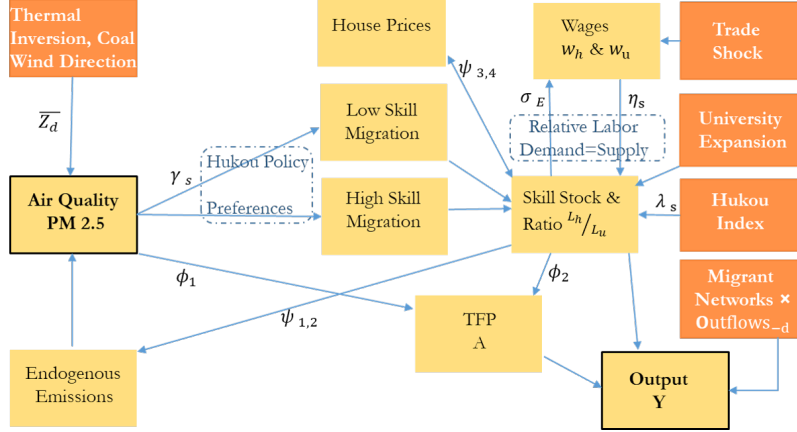
Table D10: 2013 City-level Pollution Caps: Changes in GDP per Worker (%)

	Panel A: Use Tombe and Zhu (2019) elasticity for all workers			
	Elasticity=1.19	Elasticity=1.4	Elasticity=1.5	Elasticity=1.61
Reduce steady state PM2.5	3.940	4.337	4.528	4.737
Reduce exogenous part of PM2.5	6.911	8.177	8.837	9.5605
Reduce PM2.5 & relax unskilled hukou	11.778	13.936	15.024	16.265
	Panel B: Use Tombe and Zhu (2019) elasticity for skilled workers			
	Elasticity=1.19	Elasticity=1.4	Elasticity=1.5	Elasticity=1.61
Reduce steady state PM2.5	3.614	3.611	3.609	3.606
Reduce exogenous part of PM2.5	6.214	6.467	6.575	6.684
Reduce PM2.5 & relax unskilled hukou	10.561	11.043	11.243	11.445
	Panel C: Use Tombe and Zhu (2019) elasticity for unskilled workers			
	Elasticity=1.19	Elasticity=1.4	Elasticity=1.5	Elasticity=1.61
Reduce steady state PM2.5	3.944	4.331	4.512	4.708
Reduce PM2.5 & relax skilled hukou	7.096	7.971	8.384	8.833
Reduce PM2.5 & relax unskilled hukou	12.118	13.575	14.240	14.951

Notes: In this counterfactual we reduce pollution according to the targets set by the 2013 Air Pollution Prevention and Control Plan (row 1). In addition to pollution regulations, we also relax the hukou restriction in high-tier cities (row 2), and overall migration costs (row 3) by 50

E Additional Model Derivations

Figure E1: The Flowchart of Our Spatial GE Model



Notes: We summarize our model's primary relationships in this flowchart for convenience.

E.1 Deriving Labor Supply and Welfare

In this appendix we derive the labor supply curve from the worker utility function in our main text.

$$V_{jsod} = \mu_{jsd} w_{sd} Z_d^{-\gamma_s} h p_d^{-v_s} a_d \xi_{sod} \exp^{-M_{sod}} \quad (7)$$

Workers will pick the destination with the highest value of $V_{jsod} = \widetilde{w_{sd}} \mu_{jsd}$, where we define $\widetilde{w_{sd}} \equiv w_{sd} Z_d^{-\gamma_s} h p_d^{-v_s} a_d \xi_{sod} \exp^{-M_{sod}}$ to be a composite of wages, costs and amenities. The probability that someone from origin o picks destination 1 is given by:

$$\begin{aligned} \pi_{so1} &= Pr[\widetilde{w_{so1}} \mu_{s1} > \widetilde{w_{sod'}} \mu_{sd'}] \quad \forall d' \neq 1 \\ &= Pr\left[\mu_{sd'} < \frac{\widetilde{w_{so1}} \mu_{s1}}{\widetilde{w_{sod'}}}\right] \quad \forall d' \neq 1 \\ &= \int \frac{dF}{d\mu_{s1}} (\mu_{s1}, \omega_{so1} \mu_{s1}, \dots, \omega_{soD} \mu_{sD}) d\mu_{s1} \end{aligned} \quad (E.1)$$

where we define $\omega_{sod} \equiv \frac{\widetilde{w_{sp1}}}{w_{sod'}}$. We assume that the preferences are distributed with the following Frechet distribution:

$$F(\mu_{s1}, \dots, \mu_{sD}) = \exp \left\{ - \left[\sum_{d=1}^D \mu_{sd}^{-\eta_s} \right] \right\} \quad (\text{E.2})$$

So the derivative of the CDF is given by:

$$\frac{dF}{d\mu_{s1}} = \eta_s \mu_{s1}^{-\eta_s-1} \exp \left\{ - \left[\sum_{d=1}^D \mu_{sd}^{-\eta_s} \right] \right\} \quad (\text{E.3})$$

This derivative evaluated at $(\mu_{s1}, \omega_{so1}\mu_{s1}, \dots, \omega_{soD}\mu_{sD})$, allows us to determine the probability of choosing destination 1, given by π_{1os} :

$$\begin{aligned} \pi_{so1} &= \int \eta_s \mu_{s1}^{-\eta_s-1} \exp \left\{ - \left[\sum_{d=1}^D (\omega_{sod}\mu_{sd})^{-\eta_s} \right] \right\} d\mu_{s1} \\ &= \frac{1}{\sum_{d=1}^D \omega_{sod}^{-\eta_s}} \int \left(\sum_{d=1}^D \omega_{sod}^{-\eta_s} \right) \mu_{s1}^{-\eta_s-1} \exp \left\{ - \left[\mu_{s1}^{-\eta_s-1} \left(\sum_{d=1}^D \omega_{sod}^{-\eta_s} \right) \right] \right\} d\mu_{s1} \\ &= \frac{1}{\sum_{d=1}^D \omega_{sod}^{-\eta_s}} \int dF(\mu) \\ &= \frac{1}{\sum_{d=1}^D \omega_{sod}^{-\eta_s}} \cdot 1 = \frac{(\widetilde{w_{so1}})^{\eta_s}}{\sum_{d=1}^D (\widetilde{w_{sod}})^{\eta_s}} \end{aligned} \quad (\text{E.4})$$

The third line comes from the properties of the Frechet distribution, where we know that the term in the integral of the second line is simply the PDF with a shape parameter η , and a scale parameter $\sum_{d=1}^D \omega_{sod}^{-\eta_s}$. Expanding on the definitions for $\widetilde{w_{sod}}$, and scaling up the probability by the size of the skilled workforce P_{os} by origin, we derive labor supply by skill and destination (in our main text):

$$\pi_{sod} = \frac{\left[w_{sd} Z_d^{-\gamma_s} h p_d^{-v_s} a_{sd} \xi_{sod} \exp^{-M_{sod}} \right]^{\eta_s}}{\sum_{d'} \left(w_{sd'} Z_{d'}^{-\gamma_s} h p_{d'}^{-v_s} a_{sd'} \xi_{sod'} \exp^{-M_{sod'}} \right)^{\eta_s}} \text{ and } L_{sd} = \sum_o P_{os} \pi_{sod} \quad (9)$$

The Frechet assumptions also allow us to measure aggregate welfare. Using equation 7, we can integrate over the the location preference μ_{jsd} , conditional on choosing a destination.

$$\begin{aligned}
E[V_{jsod}|d] &= (\widetilde{w_{sod}}) E[\mu_{jsd}|d] \\
&= (\widetilde{w_{sod}}) \pi_{sod}^{-\frac{1}{\eta_s}} \Gamma\left(1 - \frac{1}{\eta_s}\right) \\
&= \left(\sum_{d'} \left(w_{sd'} Z_{d'}^{-\gamma_s} h p_{d'}^{-v_s} a_{sd'} \xi_{sod'} \exp^{-M_{sod'}}\right)^{\eta_s}\right)^{\frac{1}{\eta_s}} \Gamma\left(1 - \frac{1}{\eta_s}\right), \quad (E.5)
\end{aligned}$$

where Γ is the gamma function, and is constant across cities.

Average city utility may depend on *hukou* costs. For instance, if a high-amenity city has a very restrictive *hukou* policy it may have a high average utility as those who originate from this city already have access to the amenities without paying *hukou* costs. We define average utility for those from city o to be:

$$\overline{V_{so}} \equiv \left(\sum_{d'} \left(w_{sd'} Z_{d'}^{-\gamma_s} h p_{d'}^{-v_s} a_{sd'} \xi_{sod'} \exp^{-M_{sod'}}\right)^{\eta_s}\right)^{\frac{1}{\eta_s}} \quad (E.6)$$

The equation shows that the average utility depends on the average option value migrating to any other city, and the ‘utility’ earned there. This average is scaled by the Frechet shape parameter η_s as it captures the dispersion in tastes across locations. The utility of those in city o is a decreasing function of migration costs to all other cities, as the option value of moving to those cities fall. We can therefore, rewrite the average utility as a function of *hukou* restrictions, and the labor supply as a function of utility in the manner described in the main text, by using the above set of equations:

$$\log \pi_{sod} = -\eta_s \log \overline{V_{so}} + \eta_s (\log w_{sd} - v_s \log h p_d) + \eta_s \log a_{sd} - \eta_s \gamma_s \log Z_d - \eta_s M_{sod} + \widetilde{\xi_{sod}}, \quad (10)$$

E.2 Elasticity of Capital, and Skill-biased Capital

So far the model assumes that capital is perfectly supplied at the rate R^* . If however, capital was fixed at a value $\bar{K}_d$ in a city, it would not change the skill-premia. The average

earnings for a worker with skill s in district d would be:

$$\begin{aligned} \log w_{sd} = \log \left(\frac{\partial Y_d}{\partial \ell_{sd}} \right) &= \log \theta_{sd} + \log \rho + \left[1 - \left(1 - \frac{1}{\sigma_E} \right) \left(\frac{1}{\rho} \right) \right] \log Y_d \\ &+ \left(1 - \frac{1}{\sigma_E} \right) \left(\frac{1}{\rho} \right) (\log A_d + (1 - \rho) \log \bar{K}_d) - \frac{1}{\sigma_E} \log L_{sd} \quad (\text{E.7}) \end{aligned}$$

Here the modified term $\left(\frac{1}{\rho} \right) (\log A_d + (1 - \rho) \log \bar{K}_d)$ is common across skill levels, and not affect skill premia. It varies across cities, just as TFP in the main model. We can similarly (re)define a modified TFP term that includes the immobile capital.

We can explicitly model skill biased capital as affecting the productivity parameter θ_{sd} . Below, we explicitly model skill biased capital to show how flexible forms of introducing it do not influence the estimation. In the following set up, the noticeable changes are where equation 3 has been modified into equation E.10, which includes an elasticity of substitution between labor ℓ_{sd} and skill biased capital k_{sd} represented by σ_s :

$$Y_d = A_d L_d^\rho K_d^{(1-\rho)} \quad (\text{E.8})$$

$$L_d = \left(\sum_s \theta_{sd} L_{sd}^{\frac{\sigma_E-1}{\sigma_E}} \right)^{\frac{\sigma_E}{\sigma_E-1}} \quad (\text{E.9})$$

$$L_{sd} = \left(\Lambda_s k_{sd}^{\frac{\sigma_s-1}{\sigma_s}} + (1 - \Lambda_s) \ell_{sd}^{\frac{\sigma_s-1}{\sigma_s}} \right)^{\frac{\sigma_s}{\sigma_s-1}}, \quad (\text{E.10})$$

where ℓ_{sd} is the supply of workers of skill s , and L_{sd} is now a labor aggregate over workers and capital. Given this new set up, earnings can be represented by equation E.11, instead of equation 4 in our main text:

$$\log w_{sd} = \log \left(\frac{\partial Y_d}{\partial \ell_{sd}} \right) = \frac{1}{\rho} \log A_d + \log \tilde{\rho} + \log \theta_{sd} (1 - \Lambda_s) + \frac{1}{\sigma_E} \log L_d + \left(\frac{1}{\sigma_s} - \frac{1}{\sigma_E} \right) \log L_{sd} - \frac{1}{\sigma_s} \log \ell_{sd}, \quad (\text{E.11})$$

E.3 On Existence and Uniqueness of the Equilibrium

Section 5 presents the model and equilibrium. Here we describe the determinants behind the existence and uniqueness of the equilibrium defined in Section 5.3.

When bilateral migration costs are present we make a few other standard assumptions

that help meet sufficient conditions for the existence of a spatial equilibrium (Allen et al., 2020): M_{sod} are finite, the graph of the matrix of costs is strongly connected, and they are quasi-symmetric. The connectivity assumption simply implies that there is a sequential path of finite bilateral migration costs that can link any two cities o and d . The quasi-symmetry assumption, which is not entirely necessary for existence (but does aid the solution), simply says that one portion of the costs is symmetric. That is, $M_{sod} = M_{so}M_{sd}\widetilde{M_{sod}}$, and $\widetilde{M_{sod}} = \widetilde{M_{sdo}}$. So in moving from Shanghai to Beijing, there may be a component of the cost that is Beijing specific (say, related to Beijing *hukou* policy), a component related to leaving Shanghai (say, its large airport), and a component that is Beijing-Shanghai specific (say, the distance between the two, or number of train connections). This last bilateral component is assumed to be symmetric for ease of proving the existence of an equilibrium.

Additionally, our model contains congestion forces (such as pollution and house prices) and agglomeration (effects on TFP). The existence of multiple equilibria often depends on the relative strength of agglomeration and congestion forces (Allen et al., 2020). An equilibrium is unique if congestion forces are at least as large as the agglomeration forces. That is, the parameters ψ_1, ψ_2, ψ_3 and ψ_4 that determine congestion are meaningful in magnitude, relative to ϕ_1 and ϕ_2 that drive agglomeration. More skilled workers raise TFP (via ϕ_2), yet may lead to more congestion, via higher house prices (via ψ_3 and ψ_4) and more pollution (via ψ_1 and ψ_2), which in turn may lower TFP (via ϕ_1). Given the meaningful congestion forces we may expect a unique equilibrium.

When solving for equilibrium, our model converges to the same unique equilibrium across different starting values. This does not necessarily imply the equilibrium is globally unique. Yet, like other work (Ahlfeldt et al., 2015) we envision that if there were to be multiple equilibria for a different set of parameter values, we would select the counterfactual equilibrium closest to the observed real-world outcome.

E.4 Alternative Specification of Housing Market

We assume a fixed amount of land for construction is available in each city. Developers are price-takers in the competitive market, and they sell houses at local housing prices $P_{d,sale}$ which are set through equilibrium. The inputs to housing productions include developable land (e_d) and construction materials (m_d) (as in Saiz (2010), Diamond (2016) and Piyapromdee (2021)). The aggregate housing production in city d is:

$$Q_d = b_d e^{\varphi} m_d^{(1-\varphi)} \quad (\text{E.12})$$

where Q_d is the quantity of houses in city d , b_d is the productivity in the housing production specific for the city, and φ denotes the share of housing production accruing to land. The price of construction materials P_m is exogenous. The land price is a function of the quantity of houses, $P_{d,e} = Q_{d,h}^{\kappa}$; κ is the elasticity of land price with respect to housing production. Developers maximize their profits. In the steady state equilibrium, housing sale prices $P_{d,sale}$ equal the discounted values of the price of housing services per year hp_d . Thus, the housing supply equation can be expressed as:

$$\ln hp_d = \ln(cc_d) + \varphi \kappa \ln Q_d \quad (\text{E.13})$$

where the construction costs $\ln(cc_d) = \ln i + \ln[\frac{1}{b_d} \varphi^{-\varphi} (1-\varphi)^{-(1-\varphi)} P_m^{(1-\varphi)}]$, and i is the interest rate in housing market.

We implicitly assume all housing is owned by absentee landlords (as in [Hsieh and Moretti \(2018\)](#)). Following [Anagol et al. \(2021\)](#), we aggregate individual housing demand in each location. Based on the Cobb-Douglas utility function of equation (7), the aggregate demand for housing services in city d is:

$$Q_d = \sum_s \frac{v_s L_{sd} w_{sd}}{hp_d}, s = low, high \quad (\text{E.14})$$

where v_s represents the average share of housing expenditure over wage for cohorts with skill level s . Substituting equation [E.14](#) into equation [E.13](#), the price of housing services per year is:

$$\ln hp_d = \ln \widetilde{cc}_d + \chi \sum_s v_s L_{sd} w_{sd} \quad (\text{E.15})$$

where $\ln \widetilde{cc}_d = \frac{\ln cc_d}{1+\varphi\kappa}$ is the equilibrium construction cost. $\chi = \frac{\varphi\kappa}{1+\varphi\kappa}$ represents the elasticity of housing price with respect to housing demand.

We combine our IVs for population (baseline migration network interacted with out-migration flows from each province) and for skill ratio (college expansion IV) to address the endogeneity of housing demand ($\sum_s v_s L_{sd} w_{sd}$) in equation [E.15](#). Table [E1](#) reports our estimates of equation [E.15](#). Based the estimated housing price elasticity with respect to

housing demand (0.839), we predict housing prices using equation E.15 in our counterfactual exercises. The counterfactual results are presented in Tables E2 and E3.

Table E1: The Elasticity of Housing Price with Respect to Housing Demand

	Dependent variable: Log (Housing Price)					
	(1)	(2)	(3)	(4)	(5)	(6)
Log(Housing Demand)	0.882*** (0.192)	0.839*** (0.165)	0.685*** (0.107)	0.683*** (0.109)	0.695*** (0.0953)	0.719*** (0.112)
Observations	121	121	121	121	121	121
City Control	Yes	Yes	No	No	Yes	Yes
Weather Controls	No	Yes	No	Yes	No	Yes
Regions FE	No	No	Yes	Yes	Yes	Yes
F-Test of ivs	9.066	10.59	13.86	12.80	14.39	11.45

Notes: Robust standard errors are reported in parentheses. We report Kleibergen-Paap rk Wald F statistic.

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

Table E2: The Productivity Effect of Relocating Pollution in One City

	Change in GDP per Worker in China (%)			
	Overall changes (1)	Health (2)	Relocation (3)	Relocation +Agglom (4)
Relocate steady state PM2.5	14.002	5.907	4.273	7.644
Reduce exogenous part of PM2.5	12.599	5.446	3.670	6.763
Relax skilled hukou	7.901	0.000	4.905	7.901
Relax unskilled hukou	-3.960	0.000	-4.608	-3.960
Relocate PM2.5 & Relax skilled hukou	22.245	5.907	8.712	15.427
Relocate PM2.5 & Relax unskilled hukou	9.985	5.9079	-0.063	3.851

Notes: In this counterfactual exercise we reduce the steady state amount of pollution in Beijing by 50% (row 1). We then reduce only the exogenous component of pollution by 50% (row 2). Next, we lower hukou restrictions for each skill level (rows 3 and 4) by 50%, keeping pollution fixed. Rows 5 and 6 lower the hukou regulations by 50% while reducing steady state pollution. Column 1 shows the gain to overall GDP per worker. Column 2 shows the component purely explained by the health-productivity channel. Column 3 shows the pure relocation channel, and Column 4 also incorporates agglomeration as a consequence of relocation.

Table E3: The Productivity Effect of Relocating Pollution Across Cities

	Change in GDP per Worker in China (%)			
	Overall changes (1)	Health (2)	Relocation (3)	Relocation +Agglom (4)
Relocate steady state PM2.5	7.930	2.749	3.170	3.331
Reduce exogenous part of PM2.5	4.181	1.509	1.992	2.086
Relax hukou	2.342	0.000	2.123	2.342
Relax overall mobility constraints	6.281	0.000	5.804	6.281
Relocate PM2.5 & relax hukou	10.207	2.749	5.052	5.422
Relocate PM2.5 & lower migration costs	14.554	2.749	8.636	9.335

Notes: In this counterfactual exercise we relocate PM2.5 in all cities based on the amount of skill-biased capital in the city (row 1). In row 2, we relocate only the exogenous component of pollution. In row 3, we relax hukou restrictions in the 24 top tier cities by 50%, keeping pollution fixed. In row 4 we relax overall migration costs to the 24 top tier cities by 50%, keeping pollution fixed. Column 1 shows overall gains to GDP. Column 2 shows the GDP increase as a consequence of health effects only. Column 3 shows gains due to the re-allocation of labor only. Column 4 shows the gain to GDP accounting for changes in TFP due to changes in re-location and the agglomeration of skilled workers.

F Data Appendix

Pollution Data: In our baseline analysis, city-level annual PM2.5 concentrations are measured using the Global Annual PM2.5 Grids derived from satellite data by [Van Donkelaar et al. \(2016\)](#). We employ PM2.5 satellite data, since it has the following advantages. First, official PM2.5 data were not available for most Chinese cities prior to 2012, and only a subset of cities had access to official PM10 data prior to 2012. However, our satellite PM2.5 data have been available since 1998 for all the cities in China. Second, fine particles (diameter < 2.5 μm) are more hazardous than larger particles (2.5 μm < diameter < 10 μm) in terms of mortality, cardiovascular and respiratory endpoints, and PM2.5 is considered to be the best indicator of the level of health risks from air pollution.¹ Finally, a potential concern of the official air quality data is that it may be manipulated by the local government ([Ghanem and Zhang, 2014](#); [Greenstone et al., 2020](#)), however our satellite data are immune to any underlying data manipulation.

[Van Donkelaar et al. \(2016\)](#) estimate ground-level PM2.5 by combining Aerosol Optical Depth (AOD) retrievals from the NASA MODIS, MISR, and SeaWiFS, which are subsequently calibrated to global ground-based observations of PM2.5 using Geographically Weighted Regression (GWR). The raster grids of this ground calibrated PM2.5 data have a high grid cell resolution of 0.01 degree. This yields a comprehensive and reliable measurement of air quality for a wide range of cities in China, covering all the prefecture, sub-provincial and provincial cities. The correlation between satellite PM2.5 data and monitor-based PM2.5 data in China is up to 0.8 ([Freeman et al., 2019](#)). Moreover, our satellite-derived PM2.5 data cover periods before and after China's disclosure of official PM2.5 data, allowing us to analyze the migration response to the disclosure of pollution information.

We use the Air Quality Index (AQI) released by Ministry of Environmental Protection (MEP) for robustness checks ([AQI, 2014](#)). AQI is an overall indicator for air pollution concentration calculated using multiple atmospheric pollutants including SO₂, NO₂, PM₁₀, PM_{2.5}, O₃ and CO.

Migration Data: In our baseline analysis, we obtain data on individual migration decisions from the 2015 Population Census of China ([Census, 2015](#)). China conducts its

¹see WHO report: <http://www.who.int/mediacentre/news/releases/2014/air-quality/en/>.

population census every five years; the 2015 census is the latest census with restricted public access. The census records a wide range of demographic and economic characteristics of individuals, including age, gender, education, employment, hukou location, current residential city, and marriage status. We limit the sample to the working age population. In the Census, migrants are defined as those who are away from their hukou city for more than six months. Hukou status determines citizens' access to state-provided goods (such as schools for children) and services (like marriage registries or passport renewals). Given the strong (forced) attachment to one's hukou city, when a person's location of residence differs, it can be reliably characterized as migration. We sometimes use baseline city-level population in 2000 as the regression weights, and data on population in 2000 come from the 2000 Population Census of China (Census, 2000).

In addition, we also construct an individual-level longitudinal panel using the China Labor-force Dynamics Survey (CLDS) and use an alternative definition of internal migration to examine the robustness of our empirical pattern. The CLDS data record individual histories of location changes for a sample of 21,086 individuals in 14,226 households across 29 provinces of China. A probability-proportional-to-size sampling (PPS) procedure based on population size and administrative units is adopted to ensure that the survey is nationally representative. As a result, the distribution of sample size across cities in the CLDS is consistent with the geographic distribution of population in China. The CLDS is a national longitudinal social survey, with detailed information on education, work and migration experience. Since the survey asks retrospective migration histories of each individual, we are able to construct a longitudinal panel of location histories between 2008 and 2016. The CLDS allows us to account for individual-specific unobservables, track those who have moved multiple times and those who have moved and returned home. When we analyze migration choices using the longitudinal panel data, we define migration to be an indicator for whether an individual changed city locations between years, regardless of whether they change their hukou status.

In this paper, we refrain from looking at within-city sorting for a few reasons: First, the wage and local labor market are likely to consist of the larger city as a whole. And our spatial structural model analysis aims to quantify the productivity consequences of moving workers from high productivity to low productivity cities in response to relocating pollution. Second, within-city variation can only explain about 19% of the spatial dispersion of

air quality in China.

Trade Data: Information on city exports and imports is derived from the China Customs Database, which covers the universe of Chinese exports and imports, and was harmonized and generously provided by the University of California, Davis, Center for International Data (Feenstra et al., 2018). We utilize information on the quantity and value of exports classified by the Harmonized System for all international transactions from China. The data reports the annual trade information on values, quantities, and partner countries at the HS 8-digit level for all Chinese cities in the period under investigation (i.e., 1997 to 2014). As the industry classifications used in tariffs and the China Customs Database (i.e., HS 6-digit) are different from the one in the Annual Survey of Industrial Production (i.e., Chinese Standard Industrial Classification 4-digit), we correspond them to the International Standard Industrial Classification (ISIC) Revision three at the 4-digit level to construct various trade shock measures in practice. We construct a city-level measure that captures the differential impact of PNTR across Chinese prefecture cities based on their pre-2001 industrial activity. Existing industrial activity is measured by each industry's share of total city exports, *prior* to the conferral of PNTR, calculated from the China Customs Database.

Exports are categorized by the destination country and city of origin. The 4-digit city codes provided in the customs data identify a level of geography more disaggregated than the standard prefecture cities in China. Hence, we aggregate city codes in the customs data up to the prefecture level, based on the reported city name. In the end, the original 479 city codes in the customs data are aggregated to 313 prefecture cities, including four municipalities. We do not include exports categorized as process and assembly or process with imported materials, given the heterogeneous response to such processing trade (Yu, 2015).

The tariff data comes from the Trade Analysis and Information System (TRAINS) database, which is maintained by the United Nations Conference on Trade and Development (UNCTAD). The raw tariff data is withdrawn with the simple average at the level of country-HS 6-digit. For additional balance tests, we use data on contract intensity and export licenses. Using data from Nunn (2007), we measure industry contract intensity in 1997 as the fraction of intermediate inputs employed by firms requiring relationship-specific investments. We use data from Bai et al. (2017) on the fraction of export revenues

within an industry covered under export licenses in 2000.

Data on IVs for Pollution: We obtain information on large-scale power plants, their coal consumption, and plant-level electricity generation from China Electric Power Yearbook and China Energy Statistical Yearbook. As in [Freeman et al. \(2019\)](#), the large-scale thermal power plants are defined as the thermal power plants whose installed-capacities are larger than 1 million KW. We manually collect information on the establishment year of plants, the angle between their locations and annual prevailing wind direction of each city, and the distance to each city.

We collect data on thermal inversions from the Modern-Era Retrospective Analysis for Research and Applications, Version 2 (MERRA-2), which records the 6-hour air temperature at different atmospheric layers ([GMAO, 2015](#)). For each 6-hour period, we calculate the temperature change from the first to the second above ground atmospheric layer. If the temperature change is positive, a thermal inversion occurs and the difference in temperatures measures the strength of thermal inversions. We calculate the annual occurrence and the annual sum of thermal inversion strength from the 6-hour data.

Other Data: We collect data on baseline city characteristics from the City Statistical Yearbooks([CCSY, 2016](#)). The city characteristic variables include population, GDP, total electricity consumption, industrial electricity consumption, industrial structure (the product value at service sector / manufacture sector), industrial SO₂ emission, industrial waste water emission and industrial dust emission, teacher student ratio, the number of hospitals per capita, and the number of doctors per capita.

We obtain data on weather amenities from China Meteorological Data Service Center. Our weather condition variables include temperature, wind speed, sunshine duration and humidity. We manually calculate the distance from each city to Shanghai seaport, to Tianjin seaport, and to Shenzhen seaport. Since the 2015 Census does not record individual-level wages, we use the CLDS to calculate city-and-education specific average wage.

We measure city-level housing prices using yearly average data on housing rents from the Xitai Real Estate Big Data depository ([XiTai, 2015](#)). The data has been collected since 2005, and covered 337 cities in collaboration with the China National Bureau of Statistics, and the China National Development and Reform Commission. We compare these data

to the purchase price of residential properties from the statistical yearbooks, and find a correlation of 0.93.

We derive information from a large-scale university expansion in China at the turn of the century that suddenly expanded college enrollment by 20% in certain cities, to identify skilled-worker agglomeration effects. Data on the number of college students and graduates at city level are from the China City Statistical Yearbooks. We create measures of the number of college graduates by city and year, and use the change in graduates from 2001-05 (cohorts just before and after the university expansion policy) as our instrument for the agglomeration of skilled workers.

References

- Ahlfeldt, G., S. Redding, D. Sturm, and N. Wolf (2015). The Economics of Density: Evidence from the Berlin Wall. *Econometrica* 83(6), 2127–2189.
- Allen, T., C. Arkolakis, and Y. Takahashi (2020). Universal Gravity. *Journal of Political Economy* 128(2), 393–433.
- Anagol, S., F. V. Ferreira, and J. M. Rexer (2021). Estimating the economic value of zoning reform.
- AQI, . (2014). *Air Quality Index*. Ministry of Environmental Protection in China[publisher]. <https://www.mee.gov.cn/>.
- Arceo, E., R. Hanna, and P. Oliva (2016). Does the Effect of Pollution on Infant Mortality Differ between Developing and Developed Countries? Evidence from Mexico City. *The Economic Journal* 126(591), 257–280.
- Bai, X., K. Krishna, and H. Ma (2017). How you export matters: Export mode, learning and productivity in china. *Journal of International Economics* 104, 122–137.
- CCSY, . (1994–2016). *China City Statistical Yearbooks*. China Statistics Press[publisher].
- CCSY, . (2018). *China City Statistical Yearbooks*. China Statistics Press[publisher].
- Census, . (2000). *Population Census Data*. National Bureau of Statistics of China[publisher]. <https://www.stats.gov.cn/sj/pcsj/>.
- Census, . (2015). *Population Census Data*. National Bureau of Statistics of China[publisher]. <https://www.stats.gov.cn/sj/pcsj/>.
- CFPS, . (2016). *China Family Panel Studies*. The Institute of Social Science Survey (ISSS) of Peking University[publisher]. <http://www.issp.pku.edu.cn/cfps/>.

- CGSS, . (2013). *Chinese General Social Survey*. National Survey Research Center, Renmin University of China[publisher]. <http://www.cnsda.org/index.php?r=projects/view&id=93281139>.
- Chen, S., P. Oliva, and P. Zhang (2022). The Effect of Air Pollution on Migration: Evidence from China. *Journal of Development Economics* 156, 102833.
- Chen, Y., A. Ebenstein, M. Greenstone, and H. Li (2013). Evidence on the Impact of Sustained Exposure to Air Pollution on Life Expectancy from China's Huai River Policy. *Proceedings of the National Academy of Sciences of the United States of America* 110(32), 12936–12941.
- Diamond, R. (2016). The Determinants and Welfare Implications of US Workers' Diverging Location Choices by Skill: 1980-2000. *American Economic Review* 106(3), 479–524.
- Ebenstein, A., M. Fan, M. Greenstone, G. He, and M. Zhou (2017). New Evidence on the Impact of Sustained Exposure to Air Pollution on Life Expectancy from China's Huai River Policy. *Proceedings of the National Academy of Sciences* 114(39), 10384–10389.
- Feenstra, R., H. Deng, C. Hong, P. Luck, A. Ma, H. Ma, S. Yao, G. Wright, and M. Xu (2018). Chinese and hong kong international trade data. *Center for International Data, University of California, Davis*.
- Freeman, R., W. Liang, R. Song, and C. Timmins (2019). Willingness to Pay for Clean Air in China. *Journal of Environmental Economics and Management*, 188–216.
- Ghanem, D. and J. Zhang (2014). Effortless Perfection: Do Chinese Cities Manipulate Air Pollution Data? *Journal of Environmental Economics and Management* 68(2), 203–225.
- GMAO, . (2015). *MERRA-2 inst6-3d-ana-Np: 3d,6-Hourly, Instantaneous, Pressure-Level, Analysis, Analyzed Meteorological Fields V5.12.4*, Greenbelt, MD,. USA, Goddard Earth Sciences Data and Information Services Center (GES DISC).
- Greenstone, M., G. He, R. Jia, and T. Liu (2020). Can Technology Solve the Principal-Agent Problem? Evidence from China's War on Air Pollution. *American Economic Review: Insights*.
- Hicks, D., P. March, and P. Oliva (2015). Air Pollution and Procyclical Mortality: Causal Evidence from Thermal Inversions. *Working Paper*.
- Hsieh, C. T. and E. Moretti (2018). Housing Constraints and Spatial Misallocation. *American Economic Journal: Macroeconomics* 11(2), 1–39.

- Jans, J., P. Johansson, and J. P. Nilsson (2018). Economic Status, Air Quality, and Child Health: Evidence from Inversion Episodes. *Journal of Health Economics* 61, 220–232.
- Khanna, G., W. Liang, M. Mobarak, and R. Song (2024). *Replication Data for: The Productivity Consequences of Pollution-Induced Migration in China*. American Economic Association.
- Nunn, N. (2007). Relationship-specificity, incomplete contracts, and the pattern of trade. *The Quarterly Journal of Economics* 122(2), 569–600.
- Piyapromdee, S. (2021). The Impact of Immigration on Wages, Internal Migration, and Welfare. *The Review of Economic Studies* 88(1), 406–453.
- Saiz, A. (2010). The Geographic Determinants of Housing Supply. *The Quarterly Journal of Economics* 125(3), 1253–1296.
- Tombe, T. and X. Zhu (2019). Trade, Migration, and Productivity: A Quantitative Analysis of China. *American Economic Review* 109(5), 1843–72.
- USGS, . (2010). *Global 30 Arc-Second Elevation (GTOPO30)*. US Department of Interior[publisher]. <https://earthexplorer.usgs.gov/>.
- Van Donkelaar, A., R. Martin, M. Brauer, N. C. Hsu, R. Kahn, R. Levy, A. Lyapustin, A. Sayer, and D. Winker (2016). Global Estimates of Fine Particulate Matter Using a Combined Geophysical-statistical Method with Information from Satellites, Models, and Monitors. *Environmental Science & Technology* 50(7), 3762–3772.
- XiTai, . (2015). *Housing Price Data*. XiTai Company[publisher]. www.creprice.cn.
- Xu, X., J. Liu, S. Zhang, R. Li, C. Yan, and S. Wu (2018). *China's Multi-Period Land Use Land Cover Remote Sensing Monitoring Data Set (CNLUCC) (In Chinese)*. Resource and Environment Data Cloud Platform[publisher]. <https://www.resdc.cn/doi/doi.aspx?DOIId=54>.
- Yu, M. (2015). Processing Trade, Tariff Reductions and Firm Productivity: Evidence from Chinese Firms. *The Economic Journal* 125(585), 943–988.