(In)formal living: India's dual urban housing supply elasticities*

Arnab Dutta[1], Sahil Gandhi[2], and Richard K. Green †[3]

[1]*The University of Hong Kong*
[2]*University of Manchester*
[3]*University of Southern California & Urban Institute*

July 16, 2024

Abstract

We study housing supply in the world's largest country. We show that droughts and highway infrastructure investments in one region of India affect housing demand in its other regions through migration. Using these migration-inducing shocks as demand-shifters, we estimate urban India's formal and informal housing (slum) supply elasticities. Our informal elasticity estimate provides empirical evidence of gentrification, that is, informal-to-formal conversions with rising rents. Our formal supply elasticity estimates indicate that Indian cities are supply inelastic.

JEL Classification: R21, R23, R31
Keywords: *Housing supply, migration, India*

*This project is funded by the USC Lusk Center for Real Estate. We thank Richard Arnott, Nate Baum-Snow, Marlon Boarnet, Leah Boustan, Jan Brueckner, S Chandrasekhar, Dragana Cvijanovic, Vernon Henderson, Matt Kahn, Rajat Kochhar, Anup Malani, Rakesh Mohan, Somik Lall, Jorge De La Roca, Ruozi Song, Vaidehi Tandel, William Wheaton, Yanos Zylberberg, and several others for their suggestions. We benefited from comments provided by participants at the annual ACSP, APPAM, ASSA-AREUEA, and international AREUEA conferences in 2021 and 2022, and seminars organized by the Cornell University, MIT, National University of Singapore, The University of Hong Kong, University of Cincinnati, and University of Southern California. An earlier version of this paper was circulated with the title "Distant Shocks, Migration, and Housing Supply in India." All errors are our own.

†Contact: richarkg@usc.edu

1 Introduction

The United Nations projects urban population growth of 1.9 billion between 2025 and 2050, of which more than 1.8 billion will be in developing countries. Half of this urban population increase in developing countries will be in slums.[1] The impact of this increased demand on the price of housing will be determined by the elasticity of housing supply. Ours is the first paper to estimate *both* informal and formal housing supply elasticities for a developing country. Our particular focus is India, the most populous country in the world.

We estimate the decadal supply elasticity of informal housing and formal housing in urban India to be -0.49 and 1.36 respectively. The negative supply elasticity of informal housing may seem counterintuitive, but it is consistent with redevelopment processes studied in Henderson et al. (2021, 2016). More specifically, the negative elasticity reflects gentrification, by which we mean informal-to-formal housing conversion. A simultaneous increase in rents paid and land values around slums attract real estate developers (Gechter and Tsivanidis, 2023; Harari and Wong, 2024). Developers clear slums for the construction of formal residential and commercial real estate space (Bhan, 2009). Alternatively, government and non-government programs convert informal units to formal ones through *in situ* redevelopments (Dasgupta and Lall, 2009; Michaels et al., 2021). The formal housing estimate is within the range of national average metropolitan-level elasticities provided in the literature for the United States (US). Saiz (2010) estimates an average US elasticity of 1.75 for the period between 1970 and 2000. Baum-Snow and Han (2024) provides an average US elasticity of 1.25 for the period 2000-2010. Hsieh and Moretti (2019) estimates that zoning in the US reduces housing supply elasticity to the point where it costs 8.9 percent of GDP. The fact that India has a similar supply elasticity means that it is possibly costing itself GDP.

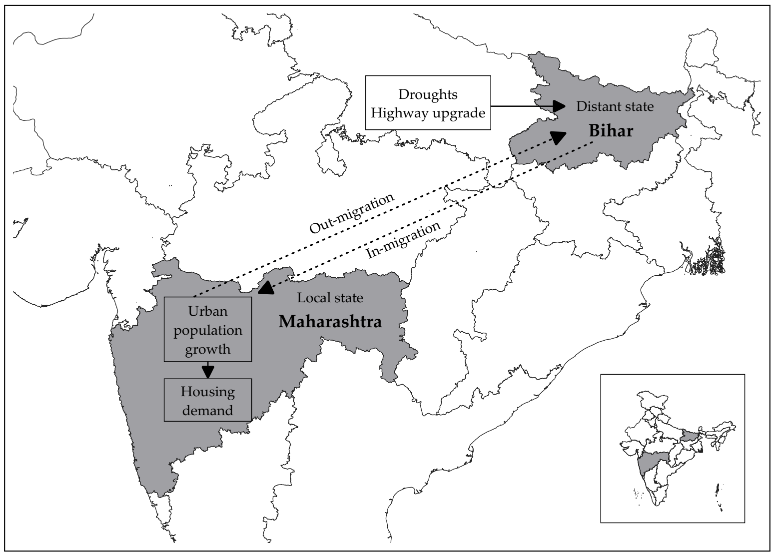
Estimating housing supply elasticities is hard. Supply elasticity estimation techniques require demand shifters, or instruments, that satisfy exclusion restrictions. Papers have used geographical constraints (Saiz, 2010), migration (Paciorek, 2013; Saiz, 2010), and labor demand shocks (Baum-Snow and Han, 2024) as instruments to estimate supply elasticities. The absence of data prevents us from following the techniques in Saiz and Baum-Snow and Han for estimating supply elasticity in India. Further, conventional demand shocks, such as Bartik shifters, are not reliable because a large informal labor sector in India prevents us from accurately characterising its employment structure.

We apply the Rosen-Roback spatial equilibrium framework (Roback, 1982; Rosen, 1979) to construct housing demand shifters. We argue that migration-inducing shocks (such as drought and highway upgrade) occurring in one region can affect the demand for housing in another without affecting the housing supply in the latter. To illustrate the spatial equilibrium mechanisms we take the example of two Indian states—Maharashtra and Bihar (see figure 1). The negative rainfall shocks and the highway upgrade program will affect wages and rents in Bihar

[1] The United Nations (2023, p.34) estimates slum population will increase from 1.1 billion to 2 billion by 2050.

(distant state), thereby, causing a state of spatial disequilibrium. The disequilibrium will spur migration between Bihar and Maharashtra, affecting urban population growth in Maharashtra.[2] Such an exogenous change in the population of the local region, in turn, causes a shift in its demand for housing. Thus, the negative rainfall shocks and the highway upgrade program in the *distant* state of Bihar will act as demand shifters for urban housing markets in the *local* state of Maharashtra.[3] We empirically test this hypothesis using data from India's Census and National Sample Surveys between 2001 and 2011.

Figure 1: Shock-induced migration's impact on housing demand



Source: Based on Census of India (2011c).
Note: Map presents a snapshot of India with the *local* state of Maharashtra and the *distant* state of Bihar highlighted. Negative rainfall shocks and highway upgradation occurs in Bihar. The resulting inter-state migration affects housing demand in Maharashtra through population growth.

In our empirical analysis, we first use state-level data to show that negative rainfall shocks and a highway upgrade program increased inter-state migration in India between 2001 and 2011. We find an additional rainfall deficit month during the decade increased the decadal out-migration from affected states by 1.1%.[4] We also find a distant state's inclusion in the Golden Quadrilateral (GQ) highway upgrade program increased migration both to and from such states.[5] This

[2] Imbert et al. (2022) use exogenous agriculture income shocks as instruments for rural-urban migration in China. Boustan (2010) use agricultural productivity shocks in the southern United States (US) as instruments for the great Black migration to the northern US during the post-war period. Gorback and Keys (2020) construct instruments from tax policies and international capital flows for estimating housing supply elasticities in the US.

[3] Through the rest of the paper, the terms *distant* and *local* are consistently used to refer to regions that experience shocks and regions where the housing markets are affected by the shocks, respectively.

[4] Rainfall shocks are defined as the number of rainfall deficit months during the decade. For a given month, a state is in rainfall deficit if the state's total rainfall in the month is less than 80% of the long-term normal. Consistent with past literature (Bhavnani and Lacina, 2017; Jayachandran, 2006) and the definitions used by the India Meteorological Department (IMD) to designate regions as rainfall deficient, we measure the negative rainfall shock variable as the number of months when absolute rainfall was less than 80% of the long-term normal.

[5] The Golden Quadrilateral (GQ) or the National Highways Development Project Phase I (NHDP I) was introduced as a highway upgrade program by the Central government of India in 2000, and it came into effect in 2001. Its primary

increased inter-state mobility led to urban population growth in the local state.

Next, using district-level data, we show that distant states' shock-induced urban population growth increased the demand for housing in local districts. Our findings indicate that both the negative rainfall shocks and the GQ highway upgrade program are strong instruments for the number of formal and informal houses in local urban markets. Using these demand shocks, we estimate urban India's national housing supply elasticities for informal and formal housing to be -0.49 and 1.36 respectively.

We conduct robustness checks to see if we satisfy the exclusion restriction. For example, negative rainfall shocks and GQ implementation in one state may *directly* affect population and housing growth in neighboring states. We estimate our elasticities with negative rainfall shocks and the GQ program implementation by limiting observations to states that are not neighbors. Another exclusion restriction concern is that if the rainfall shocks and the GQ program implementation in the distant state led to the migration of construction workers then there would have been housing supply effects in the local state. The majority of such construction migrant workers in India move on a short-term basis. We conduct a second set of robustness checks by redefining the migration variables to exclude short-run migrants who moved less than a year before the Census enumeration. Our results are robust to these checks.

We also estimate formal housing supply elasticities for the 12 largest states of India and compare them to US metropolitan-level elasticities estimated by Saiz (2010). The most supply elastic state of Maharashtra (home of Mumbai) has a supply elasticity comparable to very elastic metropolitan markets in Texas and the least supply elastic state of West Bengal (home of Kolkata) has substantially lower elasticity values than the least supply elastic metropolitan areas (Miami and Los Angeles) in the US.

Related literature. We make two contributions to the academic literature. First, we estimate the supply elasticity of informal housing. Informal housing has been studied in the literature because its existence is associated with poverty (Marx et al., 2013) and institutional frictions (Henderson et al., 2021), such as a lack of property rights (Brueckner and Selod, 2009) and formal housing regulations (Cavalcanti et al., 2019). Niu et al. (2021) underscored the important role played by informal housing markets in reducing urbanization costs in Chinese cities by providing low-income migrants with cheaper housing. In India, informal housing accounts for a formidable share of the overall stock. Hence, without an informal housing supply elasticity estimate, our understanding of housing markets in India will be incomplete. To the best of our knowledge, this paper provides the first informal housing supply elasticity estimate using direct observations on quantities and rents. The closest attempt at estimating an informal housing supply elasticity figure has been made by Niu et al. (2021) for Chinese cities. However, they calculate proxy informal housing elasticities using shares of village areas in the total urban built-up area on the edges of cities.

goal was to upgrade preexisting highways connecting the four largest metropolitan areas of India—Chennai, Delhi, Kolkata, and Mumbai—from two lanes to four lanes (Ghani et al., 2016).

We show that the informal housing supply elasticity in India is negative. This negative informal elasticity estimate implies the existence of gentrification in developing countries that happens through the conversion of informal houses into formal buildings (Bhan, 2009; Dasgupta and Lall, 2009; Gechter and Tsivanidis, 2023; Harari and Wong, 2024). The literature on informal housing has theorized that land parcels with informal housing units are redeveloped over time as rents in a city increase (Henderson et al., 2021). Gechter and Tsivanidis (2023) show that new formal housing developments in India cause gentrification in neighbourhoods with informal housing. To the best of our knowledge, empirical evidence on the rates at which informal units are converted into formal units in response to rising rents does not yet exist in the literature.

Second, we provide a policy-relevant housing supply elasticity estimate for the largest country in the world. Prior academic literature has predominantly focused on developed countries like the US (Baum-Snow and Han, 2024; Green et al., 2005; Saiz, 2010).[6] Papers have shown that regulations (Diamond, 2017; Glaeser et al., 2005; Quigley and Raphael, 2005) and natural land constraints, like undulating terrains (Saiz, 2010), reduce the supply elasticity of housing in metropolitan areas of the US. Similar supply constraints also exist in developing countries like India. Harari (2020) shows that many Indian cities are land-constrained. Land and housing markets in Indian cities are heavily regulated with floor-area-ratio (FAR)[7] limits (Bertaud and Brueckner, 2005; Brueckner and Sridhar, 2012; Nagpal and Gandhi, 2024), ceilings on vacant land ownership[8], and rent control laws (Gandhi et al., 2022). Studies have found that these regulations impose significant building costs on Indian developers (Bertaud and Brueckner, 2005; Brueckner and Sridhar, 2012; Harari, 2020). We add to the growing number of papers that investigate housing (Gandhi et al., 2021) and land markets (Duranton et al., 2015; Sood, 2019; Tandel et al., 2023) in India.

Outline. The rest of this paper is organized as follows. In section 2, we describe the data used for analysis and present some descriptives on housing and migration in India. Section 3 provides a discussion of the Rosen-Roback spatial equilibrium setting applied in this paper. Section 4 presents the empirical analysis along with the housing supply elasticity estimates. We present a discussion of the instruments, robustness checks, and state-level supply elasticities in section 5. Concluding remarks are given in section 6.

[6]Some studies have estimated housing supply elasticities in other countries, such as Australia (McLaughlin, 2012), China (Wang et al., 2012), Italy (Accetturo et al., 2021), and the United Kingdom (Malpezzi and Maclennan, 2001).

[7]The FAR of a building is equal to its total floor area divided by the area of the land parcel on which it is built (Bertaud and Brueckner, 2005). Lower FAR values indicate lower building height, and hence, stricter building regulations.

[8]Vacant land ownership was restricted in the largest urban areas of India under the urban land ceiling laws between the 1970s and the 2000s. These laws required firms and individuals to sell vacant land beyond a ceiling limit to the government at below-market prices (Dutta and Gandhi, 2023; Sridhar, 2010).

2 Data and descriptives

We gather data from the National Sample Survey Organization of India (NSS), the Census of India, and the India Meteorological Department (IMD). We construct datasets at the state and district levels. We use the state-level datasets to study inter-state migration and population growth. Using district-level data we study the impact of distant state shocks on local district-level outcomes. We estimate our elasticity figures using the district-level datasets. We constructa wide form panel for both datasets based on variable values from the Census years 2001 and 2011, which we then use to construct first-differenced variables for the analysis. In this section, we first discuss the definition of informality of housing before providing a brief description of the datasets used in the analysis. We then present some descriptives.

2.1 Informality of housing

We use an approach to identify informal and formal housing that is similar to Gechter and Tsivanidis (2023).[9] We classify houses as informal if their roofs *or* walls are made of non-durable substances such as grass, thatch, bamboo, plastic, polythene, mud, unburnt brick, and wood. On the other hand, formal houses' roofs *and* walls are made of durable substances like galvanized iron, metal, asbestos sheets, burnt bricks, stone, and concrete. This definition is closely linked to the Indian Census definitions of *katcha* and *pucca* (synonyms for non-durable and durable housing, respectively). Note that *katcha* or non-durable houses are not synchronous with slum housing in India. This is because there is a gap between the government's definition of slums and the actual living conditions of those residing in slums and non-slum neighborhoods (Rains and Krishna, 2020; Rains et al., 2019). The use of durability of housing as an indicator for informality provides consistency in the measurement of the informal housing stock.

2.2 State-level data

The Census of India provides decennial data on in-migration containing details on the numbers of migrants by their time of movement (i.e., less than a year ago, 1-4 years ago, and 5-9 years ago, etc.), their previous state of residence, their sector of origin (rural or urban), and their current place of residence (urban or rural). We construct decadal inter-state migration variables based on the number of individuals who moved into urban areas of a state from both rural and urban areas of another state in the decade leading up to the Census years, 2001 and 2011. The Census datasets also provide the urban population and the urban surface area for a given state.

We obtain data on mean monthly per capita consumption from the National Sample Survey Organization of India and calculate real values based on the Consumer Price Index data provided by the Labor Bureau of India. We get our data on the expansion of National Highways Development Project Phase I, also known as the Golden Quadrilateral (GQ) highway upgrade

[9]See Kuffer et al. (2016) for a detailed literature review on the various definitions of informal settlements.

Table 1: Summary statistics of state-level variables

	2001		2011	
	Mean	SD	Mean	SD
Variable	(1)	(2)	(3)	(4)
# Months absolute rainfall <80% last decade	58	12	64	11
Inter-state migrants into urban last decade ('000)	319	543	452	691
State population urban (millions)	8	11	11	14
Mean monthly per capita consumption (INR)	890	274	1001	374
Urban surface area (sq. miles)*	915	1,128	2,337	2,760
N	35	35	35	35

Source: Authors' calculations.
Note: Table presents summary statistics of state-level variables. Migration given in thousands, population in millions, and urban surface area in square miles. All values rounded off to the nearest integer. Real monthly per capita consumption in 2001 Indian National Rupees (INR) is calculated using the Consumer Price Index (CPI) data from the Labor Bureau of India. In PPP terms, $1 = 10 INR in 2001. For exchange rates see: https://data.oecd.org/conversion/purchasing-power-parities-ppp.htm.
*Calculation excludes Arunachal Pradesh and Sikkim where urban surface area data is missing.

program, from Ghani et al. (2016). The GQ program upgraded pre-existing highways connecting the metros of Chennai, Delhi, Kolkata, and Mumbai from two lanes to four lanes. The program extended to 14 states and union territories of India.[10] And finally, we gather rainfall shock data from the Open Government Data (OGD) portal of the Central Government of India. This dataset is sourced from the India Meteorological Department. It reports the percentage deviation of rainfall from the long-term average on a monthly basis between 1901 and 2015. We use this data to construct rainfall shock variables at the state level.[11] The summary statistics of all state-level variables are given in table 1.[12]

2.3 District-level data

We obtain data on the number of formal and informal residential housing units at the district level from the Census of India (2011b). In addition, we get data on urban population and urban surface area from the Census of India (2011d). We also gather data on district-level mean monthly per capita consumption and the mean monthly housing rents for the various types of housing units in our analysis from the National Sample Survey Organization of India (2012). These rent and consumption values are inflation-adjusted to 2001 values based on the Consumer Price Index data provided by the Labor Bureau of India (2012).

[10]In figure A.1, we map the GQ-recipient states and union territories.

[11]The original data provides rainfall departure percentages for each of the 36 meteorological subdivisions in India. Meteorological subdivisions are roughly analogous to the state boundaries of India, with a few exceptions. Larger states consist of more than one subdivision, while some smaller states are clustered into one subdivision. We map these meteorological subdivisions to state boundaries and recalculate the rainfall departure values at the state level.

[12]Table 1 shows that the average number of rainfall deficit months during a decade was around 60, or five years. Although they look large, the numbers are consistent with country-wide severe drought spells that occurred in 2002, 2004, and 2009 (Mishra and Liu, 2014).

Table 2: Summary statistics of district-level variables

	2001		2011	
	Mean	SD	Mean	SD
Variable	(1)	(2)	(3)	(4)
Urban population ('000)	1,181	1,777	1,515	2,184
Informal housing units ('000)	46	42	46	39
Formal housing units ('000)	180	311	284	446
Mean informal rent (INR)	302	223	311	233
Mean formal rent (INR)	628	276	751	332
Mean rent for all residential houses (INR)	570	246	684	307
Mean monthly per capita consumption (INR)	1,051	247	1,110	338
Urban surface area (sq. miles)	104	126	134	142
Median no. of rooms per house	1.95	0.46	1.98	0.36
N	144	144	144	144

Source: Authors' calculations.
Note: Table presents summary statistics of district-level variables. Population and housing units given in thousands and urban surface area in square miles. Mean and SD of median rooms rounded off to two decimal places. All other values rounded off to the nearest integer. Real monthly per capita consumption and rents in 2001 Indian National Rupees (INR) is calculated using the Consumer Price Index (CPI) data from the Labor Bureau of India. In PPP terms, $1 = 10 INR in 2001. For exchange rates see: https://data.oecd.org/conversion/purchasing-power-parities-ppp.htm.

The number of districts in India varies over time because of realignment of boundaries. Between 2001 and 2011, the number of districts increased from 593 to 640. Furthermore, the NSS datasets before 2002 contain district information based on 1991 boundaries when there were roughly 460 districts. To obtain a time-consistent sample of districts, we rearrange the actual administrative district boundaries and create hypothetical district boundaries that encompass all contiguous districts that underwent realignment between 2001 and 2011. This rearrangement leaves us with 479 time-consistent districts.[13] For our formal elasticity estimates we lose 140 districts because of missing data. For our informal elasticity estimates we further lose 195 districts because of missing informal rent data. Our final district-level dataset consists of 144 districts, with a total urban population of 218 million as of 2011, accounting for roughly 60% of India's overall urban population of 377 million at the time.[14] The summary statistics of district-level variables for these 144 districts are given in table 2.

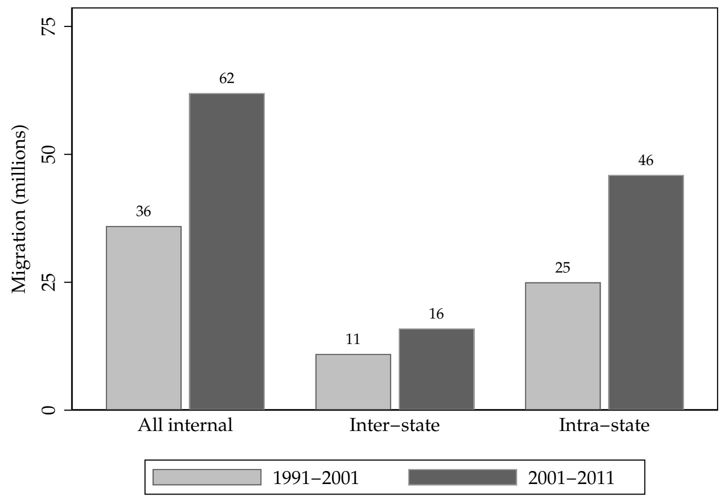
[13]A similar district matching exercise is performed by Dutta and Randolph (2022), Dutta et al. (2022), and Dutta (2024).

[14]Because of missing data on formal rent, consumption, housing units, and urban surface area, we lose 140 districts. We lose another 195 districts because of missing informal rent data, which is reported for fewer districts. We estimate the housing supply elasticity of formal units using the larger sample of 339 districts for which formal rent data are available and present the results in table 5. Note that this larger sample of 339 districts had a combined population of 349 million, or 93% of India's urban population, in 2011. Our formal elasticity estimate of 1.36, therefore, is based on almost the entirety of India's urban areas.

2.4 Descriptives

The growth in inter-state migration between the 1990s and the 2000s led to a shift in the demand for housing in India's urban housing markets. Over the last few decades, India's urbanization has been increasingly accompanied by migration. Figure 2 shows that the number of internal migrants in Indian cities increased by more than 70% from 36 million to 62 million. These numbers indicate a departure from previous trends because, historically, India is known for low levels of permanent internal mobility (Bhavnani and Lacina, 2017; Kone et al., 2018). Recent studies have shown that technological growth explains the increasing trends in India's internal migration since the 2000s (Dutta and Randolph, 2022; Ghose, 2021).

Figure 2: Internal migrants in urban India (1991-2011)



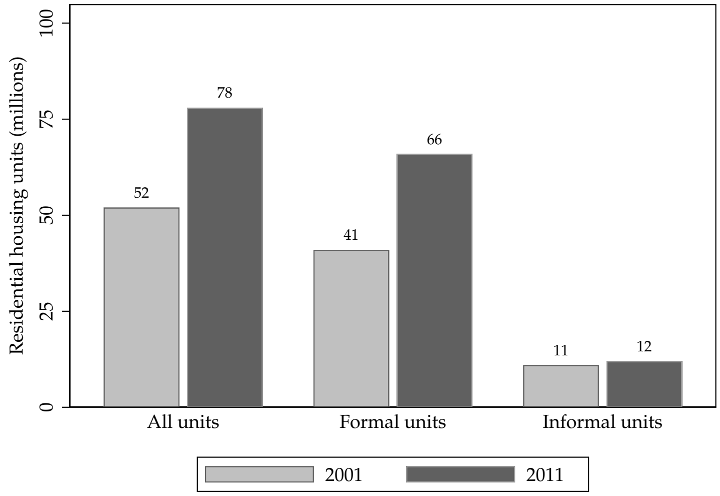
Source: Authors' calculations.
Note: Figure presents the total number of rural-urban and urban-urban migrants who moved to Indian cities between 1991-2001 and between 2001-2011. Bars labeled with their corresponding values.

At the same time, the housing stock grew rapidly with a growth rate between 2001 and 2011 of 50%. The overall housing stock in Indian cities, as of 2011, was 78 million (see figure 3). India's urban population at the time was 377 million, indicating a ratio of 4.8 persons per housing unit.[15] Indian housing markets have a variety of typologies. Out of the 78 million housing units in 2011, 66 million were formal houses made of concrete and metal. The remaining 12 million, or 15% of the housing stock, were informal houses, predominantly made from non-durable substances, such as thatch, mud, plastic, etc. Such a large presence of informal housing units is commonplace in developing countries, where they provide low-cost housing options to the urban poor (Bertaud, 2018; Glaeser, 2011; Niu et al., 2021).

Most of India's new housing demand comes from natural population growth. This is not

[15]For reference, the housing stock in the United States at the time was roughly 132 million, for a total population of 310 million, implying a ratio of 2.3. See https://fred.stlouisfed.org/series/ETOTALUSQ176N.

Figure 3: Housing stock in urban India (2001-2011)



Source: Authors' calculations.
Note: Figure presents the number of housing units in urban India. Formal housing units' roofs *and* walls are made of galvanized iron, metal, asbestos sheets, burnt bricks, stone, and concrete. Informal housing units' roofs *or* walls are made of grass, thatch, bamboo, plastic, polythene, mud, unburnt brick, and wood.

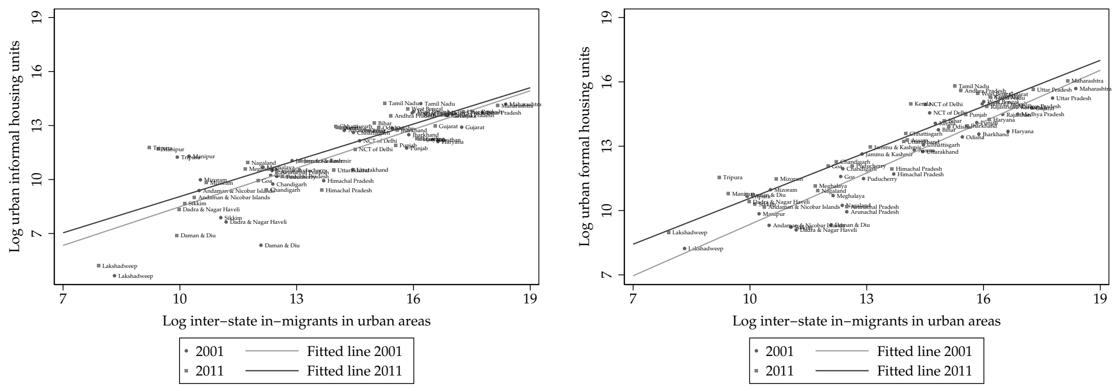
unusual worldwide. But a change in migration can lead to a change in housing demand. Inter-state movements in India constituted about 26% of overall decadal migration during the 2000s (16 million out of 62). Furthermore, table 1 indicates that state-level average inter-state migration during the 2000s was about 5% of the average state's population in 2011 (0.45 out of 11 million). Even though the volumes of inter-state migration seem low, they are not insignificant, especially when we look at these numbers from a growth perspective. Figure 2 indicates that inter-state migration grew from 11 million to 16 million between the 1990s and the 2000s, a non-trivial growth of over 45%.

We plot the log of formal and informal housing units against the log values of inter-state migration at the state-level. Figure 4 presents the scatter plots. A strong correlation exists between the size of the housing stock and the number of inter-state migrants received by states.

Increased inter-state migration and housing demand could reflect structural transformation. Figure A.2 shows that the urban population's share of migrants increased very little during this time. Inter-state migrants as a share of India's urban population remained flat at 4% between 2001 and 2011, meaning the increasing volumes of migration do not reflect structural transformation.

Another issue is that prior literature suggests that a major share of migrants to Indian cities are informal workers and they move into slums and not formal housing (Mitra, 2010; Srivastava, 2011). If this is true, migration shocks would more likely shift informal rather than formal housing demand. Other studies, however, have found that slums do not predominantly consist of migrants (Rains and Krishna, 2020). A large number of informal workers, mostly from poorer households, do migrate to Indian cities, but many affluent Indians also migrate and do so permanently rather than seasonally. For instance, the NSS data on employment and migration (National Sample Survey Organization of India, 2008) indicates that while about 12% of

Figure 4: Inter-state migration and housing in urban India



(a) Urban informal units and in-migrants

(b) Urban formal units and in-migrants

Source: Author's calculations.
Note: Figure in panel (a) presents a scatter plot of the log of state-level urban informal housing units and the log of inter-state migrants living in urban areas. Figure in panel (b) presents a scatter plot of the log of state-level urban formal housing units and the log of inter-state migrants living in urban areas

households had a seasonal migrant, about 27% of households had a former member that moved permanently for employment or education. Moreover, the NSS data also indicates that more affluent (in terms of consumption) and highly educated households were more likely to send a migrant out permanently and less likely to send migrants on a seasonal basis, which is consistent with past literature's findings (Morten, 2019).

3 Conceptual framework

We use the Rosen-Roback spatial equilibrium framework (Roback, 1982; Rosen, 1979) to analyze the effect of distant region shocks on inter-regional mobility and local housing demand. A shock that affects rents and incomes within a distant region induces spatial disequilibrium, spurring inter-regional mobility. Such mobility affects local housing demand if net inward mobility to the local region is non-zero. Distant region shocks that affect rents and incomes in the distant region act as demand shifters and can be used to estimate the local housing supply elasticity. In this section, we provide an analytical discussion of these effects.

3.1 Set-up

Consider an economy with a local region i where we are interested in estimating the housing supply elasticity and a distant region j that experiences exogenous shocks. Consistently throughout the paper, we refer to the distant region with the index j and local region with index i. The number of individuals occupying regions i and j are n_i and n_j, respectively. We assume that each individual is equivalent to a household in either region.[16] In locations i and j, individuals earn wages w_i and w_j, and derive utility from housing services $\{h_i, h_j\}$, a numeraire consumption good $\{c_i, c_j\}$, and location-specific exogenously-given amenities $\{a_i, a_j\}$. The preference structure is assumed to be strictly quasi-concave.

The market-cleaning rents for housing services are r_i and r_j. The user-cost model relates each r to its corresponding market-clearing house price p through the equation: $r = p(K + T + D + E)$. Here, K is the cost of capital, T is the local property tax rate, D is the rate of depreciation, and E is the local rate of expected appreciation (Poterba, 1984). The fact that the market-clearing rent for housing services is an appropriate measure of the market-clearing price for housing as a composite commodity is well established in the literature (Brueckner, 1987; Mills, 1967).

Solving for the consumer's utility maximization problem, we get the demand for housing services, $h_i^d(r_i, w_i)$ and $h_j^d(r_j, w_j)$. The aggregate demand for housing, H_i^D and H_j^D, are products of total populations and individuals' housing demand:

$$H_\iota^D = n_\iota h_\iota^d(r_\iota, w_\iota) \quad \text{where} \quad \iota = \{i, j\} \tag{1}$$

The utility-maximizing demand functions also provide the indirect utilities, V_ι which, in turn, gives us the spatial equilibrium condition:

$$V_i(r_i, w_i, a_i) = V_j(r_j, w_j, a_j) = \bar{V} \tag{2}$$

The values of r and w adjust, such that, conditional on amenities the indirect utility is equal across both regions i and j. At this equilibrium, there are no gains to mobility between i and j.

3.2 Spatial disequilibrium

Our spatial disequilibrium model follows Boustan (2010).[17] Consider a shock z_j at the distant region j that does not affect amenities but changes the rent or income, or both, thus changing the utility of individuals at j. The shock z could be a negative shock like a drought or a positive shock like a highway upgrade program. Because z_j affects rent r_j and income w_j, it follows that $V_j(r_j(z_j), w_j(z_j), a_j)$ is an implicit function of z_j. Hence, in response to z_j, we have a state of spatial disequilibrium:

$$\tilde{V}_j = V_j(z_j) = V_j(r_j(z_j), w_j(z_j), a_j) \neq \bar{V} = V_i \tag{3}$$

Since there are gains to mobility because of the difference in V_i and $\tilde{V}_j$, the shock z_j will induce mobility between j and i until r and w adjust in both i and j, so that a new spatial equilibrium is realized, where $\tilde{V}_j = \tilde{V}_i = \tilde{V}$.

[16] As long as the number of households and the total population at i are monotonically related, relaxing this assumption will not alter the model mechanisms.

[17] The empirical literature has also examined how labor and housing market shocks affect inter-regional mobility in the United States (Molloy et al., 2011; Saks and Wozniak, 2011).

3.3 Mobility

We characterize mobility m between regions i and j as the vector $\{m_{ji}, m_{ij}\}$. The variable m_{ji} represents the number of individuals moving from j to i and m_{ij} denotes the number of individuals moving from i to j. Note that the populations at i and j are functions of the migration vector $\{m_{ji}, m_{ij}\}$. In-migration from j to i increases population at i and decreases population at j, and vice versa.

We allow for migration to be an implicit function of the shock z_j. To see this, suppose that $m_{ij}(.)$ and $m_{ji}(.)$ are two distinct functions of the indirect utilities V_i and V_j. Spatial equilibrium implies that the net movement between two regions in equilibrium should be equal to zero. Therefore, we will have $m_{ji}(\bar{V}, \bar{V}) = m_{ij}(\bar{V}, \bar{V})$ in spatial equilibrium which implies net zero movement between i and j. Now, in response to the shock z_j the indirect utility at the distant state j changes from $\bar{V}$ to $V_j(z_j) = \tilde{V}$. The resulting migration functions can be written as follows:

$$m_{ij}(\bar{V}, V_j(z_j)) = m_{ij}(z_j); \quad m_{ji}(\bar{V}, V_j(z_j)) = m_{ji}(z_j) \tag{4}$$

Equation (4) implies that both in- and out-migration are implicit functions of the shock z_j.

We further make two assumptions about the shock z_j. First, we restrict the universe of shocks to only those that have a net non-zero effect on migration in at least one direction. Second, if the distant shock z_j affects rents and wages at i, then such effects are only through migration. These two assumptions are analogous to the relevance condition and exclusion restriction underlying our instrumental variable specification.

3.4 Housing demand

If the shock z_j leads to net non-zero mobility towards the local region, *i.e.* net in-migration, then housing demand in the local region will increase. The opposite will happen if the shock leads to net non-zero mobility away from the local region, *i.e.*, net out-migration.

Consequently, a distant shock-inducing net non-zero inter-regional mobility acts as a demand shifter in local housing markets. The distant shock affects rents and incomes in the distant region. This, in turn, changes the indirect utility in the distant region, thereby inducing a state of spatial disequilibrium in the economy. The resulting difference in utilities across the two regions implies gains to mobility. Individuals move across regions. This movement causes a change in the local population, thus affecting local housing demand.

It is important to note here, that, distant state shocks such as highway upgrades can lead to increased trading of commodities across states (Donaldson, 2018) which can, in turn, affect commodity prices, thereby, causing a second channel of impact on housing demand through substitution effects. This is particularly true if both the distant and the local states were recipients of the highway upgrade program. However, it is unlikely that the highway upgrade would have had a long-term impact on consumers' budget constraints since long-term non-housing commodity prices are affected by other factors such as international prices.

We empirically test whether the Golden Quadrilateral (GQ) highway upgrade in India had any impact on recipient states' commodity prices by regressing price changes of baskets of non-housing commodities consumed by urban households in a state on the state's GQ recipient status. We find that a state's inclusion in the GQ program had no impact on its commodity prices in urban areas. The results from this regression are given in table A.1. We, therefore, ignore the trade channel of impact of distant states' inclusion in the GQ program on local housing demand.

3.5 Housing supply

Suppose that the total housing stock H_i^S in the region i is supplied through a competitive market. In market equilibrium, we have $H_i^S = H_i^D(r_i, w_i, n_i)$. Assuming that the supply function is log-linear, the inverse supply function at i is:

$$\log(r_i) = \frac{1}{\eta_i}\log(H_i^S) \tag{5}$$

where, the housing supply elasticity at i is η_i. Estimating η_i in equation (5) presents a classic endogeneity problem because we only observe market equilibrium values of r_i and H_i^S. We hence need exogenous demand shifters to trace the slope $1/\eta_i$ of the inverse supply curve. We discussed in section 3.4 how exogenous shocks z_j in a distant region can act as a demand shifter at i if the shock z_j induces net non-zero mobility between i and j. We can write the reduced form effect of z_j on the aggregate demand for housing services as follows:

$$\log(H_i^D) = \beta z_j \tag{6}$$

Our discussion in section 3.4 implies that β could be either negative or positive, depending on the relative magnitude of the in- and out-migration effects of the shock z_j. The predicted $\log(\hat{H}_i^D)$ obtained after estimating the parameter β can be substituted in equation (5) to estimate η_i. The predicted $\log(\hat{H}_i^D)$ may not be exogenous if the shock-induced migration leads to labor market changes in the local region and, in turn, affects construction wages. In such a case, β might include supply-side factors as well, a concern we address using robustness checks discussed in section 5.1.2.

We hypothesize that z_j is a demand shifter for formal and informal houses. Since these housing categories are different markets, their slopes will be different.

4 Empirical implementation

In this section, we use empirical analysis to test the impact of distant shocks on migration and population growth and to estimate housing supply elasticities. Specifically, we estimate the impact of two distant shocks, droughts and highway infrastructure investments, on inter-state mobility, and the subsequent impact of such mobility on local urban population growth. Next, we use the same shocks as instruments for local urban population growth and estimate the urban population growth impact on local housing demand. Finally, we use the shocks as demand shifters and estimate the supply elasticites for formal and informal housing.

4.1 Distant shocks, inter-state migration, and urban population growth

We empirically estimate the effect of distant state shocks on inter-state mobility and the resulting impact of mobility on a local state's urban population growth. For this estimation, we use state-level panel data with two periods—Census years 2001 and 2011. An observation is a state-state pair.

4.1.1 Empirical strategy

Consider the local state i and the distant state j. Migration flows between i and j are given as m_k where $k = \{ji, ij\}$. m_{ji} represents the number of individuals moving from j to i and m_{ij} denotes the number of individuals moving from i to j. Our goal is to isolate the impact of migration flows, m_{ji} and m_{ij}, on the local state's urban population growth $\Delta \log(n_i)$ between 2001 and 2011.

We estimate the following first difference equation using data on $i - j$ pairs of Indian states between 2001 and 2011:

$$\Delta \log(n_i) = \lambda_1 \Delta \log(m_{ji}) + \lambda_2 \Delta \log(m_{ij}) + \lambda_3 \Delta x_i + v_{ij} \tag{7}$$

where, x_i is a vector of controls that include the log of per capita consumption (proxy for income), urban surface area, and urban surface area squared at i.[18] The error term is given by v_{ij} and Δ represents changes in the variable values between 2001 and 2011. All time invariant unobservables will be eliminated in the first-difference set-up.

Inter-state migration flows are clearly endogenous to local urban population growth. For the two endogenous variables m_{ji} and m_{ij}, we therefore use two shocks occurring at the distant state j as instruments. We call the distant state shock vector, $z_j = \{s_j, g_j\}$. Here, s_j represents the change in the number of rainfall deficit months in the previous decade; and g_j is a dummy variable equal to one if state j was a recipient of the National Highways Development Project Phase I or the Golden Quadrilateral (GQ) highway upgrade program and zero otherwise. A state is in rainfall deficit in a given month if the state's rainfall levels in the month are less than 80% of the long-term normal (Bhavnani and Lacina, 2017; Jayachandran, 2006). We estimate the impact

[18] We include the urban surface area of i as a control to account for changes in the total urban surface area across states. Total urban surface areas of states and districts change over time because smaller urban-like settlements in India are reclassified and declassified as Census towns each Census year. Census towns are governed by rural administrative bodies but have urban-like features, such as a population of at least 5,000, a population density of at least 400 persons per square kilometer, and at least 75% of the male workforce employed in the non-agricultural sector (Pradhan, 2017; Tandel et al., 2019). Since we have aggregated data for all urban areas, controlling for the urban surface area allows us to mitigate any effect on our dependent variables that can be attributed to the change in the urban surface area itself. We also control for urban surface area squared to account for the possibility of a non-linear relationship between urban growth and surface area arising due to non-linear transport cost increase with distance away from centers of cities (Henderson et al., 2021).

of s_j and g_j on migration flows m_k between i and j using the following first-stage equation:

$$\Delta \log(m_k) = \alpha_1 s_j + \alpha_2 g_j + \alpha_3 \Delta x_i + \varphi_k \tag{8}$$

where, x_i is as defined for equation (7). The error term is given by φ_k. The identification of the coefficients α_1 and α_2 comes from the variation in the rainfall shocks and highway upgrade status between different states.

4.1.2 Effect of inter-state migration on local urban population growth

In the first stage, we estimate the impact of distant state shocks on inter-state migration. We then use the distant state shocks as instruments for migration and estimate migration's impact on local urban population growth. The two instruments are: (i) decadal changes in the number of months when absolute rainfall was less than 80% of the long-term normal at j and (ii) a dummy equal to one if state j was a recipient of the Golden Quadrilateral (GQ) highway upgrade program, and zero otherwise. The results are given in table 3.

We observe a number of things in table 3. First, while the rainfall shock at j had a positive impact on migration from j to i, it did not affect migration from i to j. One additional month of absolute rainfall less than 80% of the long-term normal at j increased migration from j to i by approximately 1.1%.[19] This is consistent with the literature that negative rainfall shocks spur outward mobility from affected regions in India (Bhavnani and Lacina, 2017; Rosenzweig and Udry, 2014).

Second, the highway upgrade at j had a positive and significant impact on migration from both j to i and i to j. A distant state's inclusion into the GQ program increased migration from j to i by 27% and migration from i to j by 21%. On the one hand, the labor demand shock from firm relocation along the highway (Ghani et al., 2016) would have increased movement toward states included in the GQ program (Bartik, 1993). On the other hand, higher incomes, possibly resulting from the labor demand shock, would have spurred movement outward from those states because higher earners are better insured against uncertain migration outcomes (Bryan et al., 2014; Morten, 2019; Munshi and Rosenzweig, 2016).

In the second-stage, we see that an increase in migration from j to i caused the urban population at i to increase at the same rate as migration. The parameter estimate of the impact of migration from i to j on urban population growth at i is negative but not significant. Our evidence implies that the net impact of the distant shock-induced migration was to cause an increase in urban population at i.

We report the F-stat on excluded instruments which pass the threshold of 10 indicating that the instruments are not weak.

[19] Coefficients in log linear regressions are approximations of how changes in X variables lead to percentage point changes in Y variables. From now on the word approximately is implied.

Table 3: Distant shock-induced migration and local urbanization

	First-stage		Second-stage
	Dependent variable = ΔLog indicators		
	Migration 0-9 yrs.		Urban population
	$i \leftarrow j$	$i \rightarrow j$	at i
$\Delta = 2011 - 2001$	(1)	(2)	(3)
ΔLog migration 0-9 yrs. $i \leftarrow j$			1.018***
			(0.244)
ΔLog migration 0-9 yrs. $i \rightarrow j$			-0.379
			(0.385)
Δ#Months rainfall $<$ 80% at j	0.011***	-0.002	
	(0.003)	(0.003)	
GQ highway at j dummy	0.268***	0.213***	
	(0.043)	(0.046)	
Controls	Y	Y	Y
F-stat on excluded instruments	29.7	10.6	
N	1,028	1,028	1,028
Adj. R^2	0.179	0.121	

Source: Authors' calculations.
Note: Table presents two-stage least squares first-difference regression estimates using a panel of state-state pairs in India between 2001 and 2011. State i is the local region and state j the distant. Excluding missing pairs, there are 1,028 $i - j$ state pairs consisting of 35 states and union territories across India. Migration $i \leftarrow j$ is in-migration to i and $i \rightarrow j$ is out-migration from i. Migrants are those who moved in the previous decade. Differences are calculated as first differences of values between the years 2001 and 2011. Log in- and out-migration are instrumented by the decadal change in the number of months when rainfall was less than 80% of the long-term normal in state j and the GQ at j dummy, which equals one if state j was recipient of the Golden Quadrilateral highway upgrade program, and zero otherwise. Consumption, urban surface area, and urban surface area squared at i are the controls. For estimates of controls, see online appendix table C.1. Robust standard errors presented in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

4.2 Urban population growth and housing demand

We estimate the impact of distant shock-induced urban population growth on housing demand in the local region. For this exercise, we use district-level variables for the local region and state-level variables (shocks) for the distant region. Each observation is a district-state pair. Here as well, we have a panel data with two periods—2001 and 2011.

4.2.1 Identification

Following our use of notations from the earlier sections, growth in urban housing units is given as $\Delta \log(H_i)$ and urban population growth as $\Delta \log(n_i)$ in the local district i. We estimate the impact of $\Delta \log(n_i)$ on $\Delta \log(H_i)$, with the following equation:

$$\Delta \log(H_i) = \sigma_1 \Delta \log(n_i) + \sigma_2 \Delta x_i + \vartheta_i \tag{9}$$

where, x_i is a vector of covariates which consists of the log of district-level mean per capita consumption, the urban surface area, and the urban surface area squared.[20] The error term is given by ϑ_i. As before, Δdenotes changes in variables between 2001 and 2011.

A least square estimate of σ_1 in equation (9) will produce biased estimates because of omitted variable bias. There are unobservables that simultaneously affect urban population growth and housing demand. We therefore instrument $\Delta \log(n_i)$ with rainfall shocks s_j and the GQ highway upgrade dummy g_j, defined in section 4.1.1.

With data on $i - j$ district-state pairs we estimate the following first-stage equation:

$$\Delta \log(n_i) = \mu_1 s_j + \mu_2 \delta g_j + \mu_3 \Delta x_i + \nu_{ij} \tag{10}$$

where, all variables are as defined before and the error term is given by ν_{ij}. The predicted exogenous population growth from estimating equation (10) is used to estimate equation (9).

4.2.2 Effect of local urban population growth on housing demand

The results from estimating equations (9) and (10) with first difference models using data on $i - j$ district-state pairs in India between 2001 and 2011 are presented in table 4.[21]

In the first-stage regressions, we find that both rainfall shocks and the highway upgrade program at the distant state j led to urban population growth at i. An additional month of rainfall level less than 80% of the long-term normal at state j led to an increase in urban population by 1.1% at i. The inclusion of state j in the GQ program increased the urban population by 12% at i. This is consistent with the findings in table 3 that the shocks had a positive effect on migration from j to i and not on migration from i to j and that such in-migration from j to i led to urban population growth at i.

In the second-stage regression results, we see that the distant state shock-induced urban population growth had a positive impact on the number of informal and formal housing units. A 1% increase in urban population led to a 0.21% increase in demand for informal houses and 1.9% increase in demand for formal houses. Urban population growth's stronger impact on formal housing units compared to informal units is consistent with lower floor space usage in informal housing.[22]

We report three diagnostic tests: (i) the F-statistic on excluded instruments, (ii) the Montiel-Olea-Pflueger effective F-statistic (Olea and Pflueger, 2013), and (iii) the Sargan Hansen J-statistic's

[20] The average number of rooms in housing units at i may be correlated with the right-hand side variables as well as the number of housing units. This would imply omitted variable bias. We run regressions with the median number of rooms per housing unit in a district i as an additional control and find that the results are largely similar (see table A.2).

[21] The estimates in table 4 are based on the sample of 4,896 $i - j$ district-state pairs for which we have both formal and informal rent data. In table A.3, we present the estimates for the impact of instrumented urban population growth on formal housing demand using the entire sample of 11,526 $i - j$ district-state pairs for which we have formal rent data. The estimates in table A.3 are largely similar to those given in columns (1) and (3) of table 4.

[22] Data from the National Sample Survey Organization of India (2012) housing conditions survey conducted in 2012 across India indicates that the average per capita floor space consumed in formal housing was 77 square feet, and in informal housing was 52 square feet.

Table 4: Distant shock-induced local urbanization and housing demand

	First-stage	Second-stage	
	Dependent variable = ΔLog indicators		
	Urban population at i	Housing units at i	
		Informal	Formal
$\Delta = 2011 - 2001$	(1)	(2)	(3)
ΔLog urban population at i		0.205***	1.855***
		(0.025)	(0.026)
Δ#Months rainfall $< 80\%$ at j	0.011***		
	(0.000)		
GQ highway at j dummy	0.124***		
	(0.005)		
Controls	Y	Y	Y
F-stat on excluded instruments	1,018		
Montiel-Olea-Pflueger Eff. F-stat.	929		
Sargan-Hansen J-stat p-value		0.701	0.420
N	4,896	4,896	4,896
Adj. R^2	0.603		

Source: Authors' calculations.
Note: Table presents two-stage least squares first-difference regression estimates using a panel of district-state pairs in India between 2001 and 2011. District i is the local region and state j the distant. Excluding missing pairs, there are 4,896 $i - j$ district-state pairs consisting of the 35 states and union territories and 144 districts of India in our sample. Differences are calculated as first differences of values between the years 2001 and 2011. Log urban population at i is instrumented by the decadal change in the number of months when rainfall was less than 80% of the long-term normal in state j and the GQ at j dummy, which equals one if state j was recipient of the Golden Quadrilateral highway upgrade program, and zero otherwise. Consumption, urban surface area, and urban surface area squared at i are the controls. For estimates of controls, see online appendix table C.2. Robust standard errors presented in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

p-value. The F-statistics are above 10 indicating that the instruments are not weak. The J-statistic's p-values are reported because we have two instruments for one endogenous variable. The insignificance of the J-statistics suggests that the overidentification restrictions hold.

In an additional set of analyses, we implement estimating equations (9) and (10) using one instrument at a time. We report these results in table A.4. The estimates remain largely similar when we use either just rainfall shocks or only the GQ dummy as an instrument for urban population growth at i, compared to the main estimates presented in table 4.

4.3 Demand shifters and housing supply elasticity estimation

We estimate the inverse supply elasticity of housing for urban India. We use district-level panel data as in section 4.2 to construct first-differenced variables. Each observation is a district-state pair.

4.3.1 Estimating equations

Suppose that the supply of housing in response to housing market rent changes $\Delta \log(r_i)$ at district i is given by $\Delta \log(H_i^S)$. Ideally, we would like to estimate the following inverse supply equation:

$$\Delta \log(r_i) = \eta \Delta \log(H_i^S) + \varepsilon_i \tag{11}$$

where η is the inverse supply elasticity and ε_i is the error term.

However, we do not observe $\Delta \log(H_i^S)$. Instead, we know the market equilibrium quantities of housing units $\Delta \log(H_i)$. This creates an endogeinity problem. Addressing this problem requires demand shifters. In section 3, we discuss the conceptual framework for using shocks occurring in a distant state j, that affect rents and incomes at j, as demand shifters for housing in district i. The results in sections 4.1 and 4.2 show that distant state shocks shift housing demand through population changes resulting from migration.

We use rainfall shocks s_j and the GQ highway upgrade dummy g_j defined in section 4.1 as instruments for $\Delta \log(H_i)$. We estimate the following first stage equation using data on $i - j$ district-state pairs in India between 2001 and 2011:

$$\Delta \log(H_i) = \beta_1 s_j + \beta_2 g_j + \beta_3 \Delta x_i + \epsilon_{ij} \tag{12}$$

where, all variables are as defined before and ϵ_{ij} is the error term.[23] We estimate two sets of equations, one each for informal and formal housing units.

4.3.2 Housing supply elasticity estimates

We estimate supply elasticities for the two different types of urban housing units in our analysis—informal and formal. We estimate the coefficients in equation (12) with first difference models using data on $i - j$ district-state pairs in India between 2001 and 2011. The results are presented in table 5.[24]

In the first-stage regressions, we see that both rainfall shocks and the GQ highway upgrade program at a distant state j had a strong positive effect on both types of housing units in district i. An additional month with rainfall levels less than 80% of the long-term average at j increased the demand for informal houses by 0.2% and formal houses by 2.1% at i (columns (1) and (2)). The inclusion of state j in the GQ program led to a 2.7% higher demand for informal houses and 23% higher demand for formal houses in district i (columns (1) and (2)). These effects are consistent with the results from tables 3 and 4 which show that the distant shocks induced local urban population growth through increased inter-state migration and, therefore, led to higher

[23] A relevant point here is that contrary to the existing literature on housing supply elasticity estimation, we do not control for construction cost in equation (12). This is because we do not have any data on construction costs at the district level in India.

[24] The Ordinary Least Squares (OLS) estimates of equation (11) are presented in table A.6.

Table 5: Housing demand shifters and inverse supply elasticity estimation

	First-stage			Second-stage		
	Dependent variable = ΔLog indicators					
	Housing units at i			Rents at i		
	Informal	Formal	Formal (full data)	Informal	Formal	Formal (full data)
$\Delta = 2011 - 2001$	(1)	(2)	(3)	(4)	(5)	(6)
ΔLog informal units at i				-2.060***		
				(0.564)		
ΔLog formal units at i					0.617***	0.738***
					(0.037)	(0.029)
Δ#Months rainfall $<$ 80% at j	0.002***	0.021***	0.020***			
	(0.000)	(0.001)	(0.000)			
GQ highway at j dummy	0.027***	0.226***	0.214***			
	(0.007)	(0.007)	(0.004)			
Controls	Y	Y	Y	Y	Y	Y
Implied housing supply elasticities				-0.49	1.62	1.36
F-stat on excluded instruments	27.4	1,700	3,697			
Montiel-Olea-Pflueger Eff. F-stat.	27.6	1,449	3,201			
Sargan-Hansen J-stat p-value				0.938	0.798	0.799
N	4,896	4,896	11,526	4,896	4,896	11,526
Adj. R^2	0.075	0.659	0.674			

Source: Authors' calculations.
Note: Table presents two-stage least squares first-difference regression estimates using a panel of district-state pairs in India between 2001 and 2011. District i is the local region and state j the distant. Excluding missing pairs, there are 4,896 $i - j$ district-state pairs in columns (1), (2), (4) & (5) consisting of 144 districts and 35 states and union territories of India in our sample. Columns (2) and (5) is restricted to the 144 districts for which we have rent data for both informal and formal housing. Excluding missing pairs, there are 11,526 $i - j$ district-state pairs in columns (3) & (6) consisting of the 35 states and union territories and 339 districts of India in our complete sample (full data) on formal rents. Differences are calculated as first differences of values between the years 2001 and 2011. Log housing units at i are instrumented by the decadal change in the number of months when rainfall was less than 80% of the long-term normal in state j and the GQ at j dummy, which equals one if state j was recipient of the Golden Quadrilateral highway upgrade program, and zero otherwise. Consumption, urban surface area, and urban surface area squared at i are the controls. For estimates of controls, see online appendix table C.3 and table C.4. Robust standard errors presented in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

housing demand in the local district.

The second-stage results provide the inverse supply elasticity estimates for urban India. First, we find that the inverse elasticity for informal houses is -2.06 (column (4)) implying a decadal informal housing supply elasticity of -0.49. A negative value for housing supply elasticity is counter intuitive. However, a negative elasticity of supply for *informal* houses indicates that a process of urban gentrification is underway in Indian cities. We define gentrification as informal-to-formal housing conversion, and it occurs in two ways. A simultaneous increase in rents paid by slum dwellers and land values around slums attract real estate developers (Gechter and Tsivanidis, 2023; Harari and Wong, 2024). Slums are cleared for the construction of formal residential and commercial real estate space (Bhan, 2009). Slums are also upgraded through various government and non-government programs implementing *in situ* redevelopments (Dasgupta and Lall, 2009; Michaels et al., 2021; Rains and Krishna, 2020; Rains et al., 2019). The empirical evidence on gentrification is also consistent with existing economic theories about urbanization in developing countries. In a monocentric-city model, Henderson et al. (2021) shows how informal houses are converted to formal houses when the bid-rents are driven up in a developing country's city through productivity shocks.[25]

And second, the inverse elasticity for formal houses is 0.62 (column (5)), implying a decadal formal supply elasticity of 1.62. The regression sample used in column (5) of table 5 is restricted to the 144 districts for which we have rent data for both informal and formal housing. Since informal rent data is missing for a large number of districts, we estimate another formal elasticity value using the full sample of 339 districts for which we have formal rents. The latter formal supply elasticity estimate is 1.36 (column (6) in table 5).

To get a sense of what these elasticity estimates mean, we turn to the literature on housing supply elasticity estimation in the United States (US). Saiz (2010) estimates a national average metropolitan-level supply elasticity of 1.75 for the period 1970-2000. Baum-Snow and Han (2024), on the other hand, provides a more recent estimate of 1.25 during the decade of 2000-2010. A lower elasticity estimate for the 2000s (as opposed to 1970-2000) indicates that US housing markets have become more inelastic over time because of increasingly regulated land use. Other papers have proposed that national housing supply elasticities in the US are between 1 and 3 (Gyourko et al., 2008). Both our formal housing supply elasticity estimates (1.36 and 1.62) lie within the range of estimated US supply elasticities. Studies have argued that the US is supply inelastic (Baum-Snow and Han, 2024; Glaeser et al., 2008; Green et al., 2005; Saiz, 2010) implying that India's formal housing supply is also inelastic. This indicates the existence of regulatory and institutional frictions in Indian markets that impede the new construction of houses.

As in section 4.2.2, we report the F-statistics on excluded instruments, the Montiel-Olea-Pflueger effective F-statistics, and the Sargan Hansen J-statistic's p-values in table 5. The F-statistics are higher than 10 for all first-stage regressions implying that the instruments are not weak. We report the J-statistics' p-values because we have two instruments for one endogenous variable in all regressions. The insignificance of the J-statistic means that the overidentification restrictions are satisfied.

We also re-estimate equations (11) and (12) using rainfall shocks and the GQ dummy individually as instruments for housing units. These results, presented in table A.5, show estimates that are very similar to the ones given in table 5.

5 Discussion

In this section, we first discuss the validity of our instruments and implement robustness checks to address endogeneity concerns. Next, we show state-level elasticities for some of the largest states in India.

[25]The negative supply elasticity for informal housing does not necessarily imply the absence of new informal housing. It simply means that more old informal housing is redeveloped into formal housing than new informal housing is created. The dynamic reflects how capital to land ratios increase as land values increase (Brueckner, 1987).

5.1 Validity of instruments

Instruments need to satisfy the relevance condition and the exclusion restriction. We discuss how rainfall shocks and the GQ dummy satisfy these conditions.

5.1.1 Relevance

For the relevance condition to hold the instruments will have to be highly correlated with the endogenous variables. The IV literature is extensive on how weak instruments can result in inconsistent estimation of IV regressions.[26] We report the F-statistic on excluded instruments and the Montiel-Olea-Pflueger effective F-statistic.[27] All the reported diagnostics indicate that the instruments are not weak. Furthermore, prior literature's findings on the economic impacts of negative rainfall shocks and highway infrastructure are consistent with our results.

Negative rainfall shocks act as negative income shocks in most parts of India due to largely rainfall-dependent agricultural practices. Hence, rainfall levels less than 80% of the long-term normal induce drought-like conditions and are unfavorable for agricultural output. A body of literature has examined the relationship between rainfall shocks and agricultural output (Jayachandran, 2006), and the subsequent impact of rainfall shocks on migration (Morten, 2019; Rosenzweig and Udry, 2014). Rainfall shocks have also been used as an instrument to study civil conflict (Sarsons, 2015) and dowry deaths in India (Sekhri and Storeygard, 2014). Bhavnani and Lacina (2017) construct an instrument from negative rainfall shocks to estimate the effect of inter-state migration flows on fiscal federalism in India. Our empirical findings that negative rainfall shocks affect migration is therefore not new to the literature. Our use of rainfall shocks as an instrument for housing demand is, however, is novel.

The GQ highway upgrade project has been documented in the literature as a positive economic shock since it affected firm relocation along the highway in states through which it passed (Abeberese and Chen, 2021; Ghani et al., 2016). Based on economic theory and evidence, we should expect two countervailing effects of the highway upgrade program in a state. First, due to firm relocation along the highways, we would expect to see a growth in employment in the recipient states, leading to migration towards such states (Molloy et al., 2011; Saks and Wozniak, 2011). And second, the firm and employment growth may also lead to a positive income shock in recipient states.[28] This would lead to more movement out of recipient states because higher incomes insure individuals against risky migration outcomes (Bryan et al., 2014; Morten, 2019;

[26] The most recent study is by Mikusheva and Sun (2022) who propose a jackknifed version of the Anderson-Rubin weak IV test for the multiple-IV case.

[27] Note that the Montiel-Olea-Pflueger effective F-statistic can be calculated only in regressions with one endogenous variable (Olea and Pflueger, 2013). We, therefore, present it in tables 4 and 5 which have one endogenous variable and not table 3 which contains two endogenous variables.

[28] Whether the market clearing wage would change in GQ recipient states will depend on net-migration resulting from the GQ program-induced labor demand shock. If net-migration increased the labor supply more than the labor demand induced by the GQ program, then there might as well have been a decrease in the market clearing wage. However, given that our estimates show out-migration resulting from the GQ program implementation, the likelihood of a wage increase is higher.

Munshi and Rosenzweig, 2016). Our empirical findings are consistent with these two countervailing effects of the GQ program on in- and out-migration.

Both instruments used in the analysis are well grounded in theory and the existing literature. Given that the instruments pass the diagnostics as well, we argue that the relevance condition holds.

5.1.2 Endogeneity concerns and robustness checks

The instruments we chose are strong predictors of our endogenous variables of interest, but we have yet to discuss whether they satisfy the exclusion restriction. For example, negative rainfall shocks and GQ implementation in one state may *directly* affect population and housing growth in neighboring states.

Rainfall shocks may be spatially correlated. For instance, monsoonal floods in the neighboring states of Bihar, West Bengal, and Orissa often occur simultaneously. Bhavnani and Lacina (2017) recognize this issue and control for rainfall shocks in both migrants' origin and destination states. We take another approach to addressing this issue, which we shall discuss later in this section.

The inclusion of one state in the GQ program will not affect the inclusion of another because these highways were constructed on trade routes built during ancient and colonial times. For instance, one part of the GQ program, the National Highway II (NH2), was constructed on portions of the Grand Trunk Road that was first built by the emperor Chandragupta Maurya during the 3rd century BCE and later redeveloped under the rule of emperor Sher Shah Suri, the Mughals, and the British Raj (Elisseeff, 2000; Thapar, 2015). Contiguous states may still have a higher likelihood of being on the same trade route. Such contiguity effects will, however, be absorbed by the first-difference framework because there were no changes in state boundaries between 2001 and 2011. But, the GQ upgrade program implementation across contiguous states might have potentially affected housing supply if better contiguous-state road networks led to higher trading, and thus, reduced prices of construction material.[29]

We address these exclusion restriction concerns through robustness checks. In these robustness checks, we estimate equations (7) and (8) with negative rainfall shocks and the GQ program implementation occurring in only non-contiguous states as the instruments. The results from these regressions are in table B.1 and table B.2. The coefficient estimates in table B.1 are very similar in magnitude to those seen in table 3. In table B.2, we present the housing supply elasticity estimation results with non-contiguous state shocks and do not find any material changes from the main estimates given in table 5.

Another exclusion restriction concern is that if the rainfall shocks and the GQ program implementation in the distant state led to migration of construction workers then there would have been housing supply effects in the local state. This is because a migration-induced labor supply shock in the construction sector would have led to changes in construction wages, affecting

[29]Even though we expect construction material to be traded across neighboring states, it is unlikely that such trading occurs across non-contiguous states since substances such as cement are heavy and difficult to transport.

construction costs. This is particularly relevant in the Indian context since a large number of Indians migrate for construction work. However, such migrant workers are more likely to move seasonally for one to six months before moving back to their homes. The NSS survey on migration and employment (National Sample Survey Organization of India, 2008) conducted in 2007-08 indicates that 36% of seasonal inter-state migrants move for construction work compared to only 1.5% of long-term inter-state migrants. Hence, if we eliminate short-run migrants from our analysis, we alleviate the endogeneity concern arising from the housing supply effects of distant state shocks. We conduct a second set of robustness checks by redefining the migration variables in equation (7) to exclude short-run migrants who moved less than a year before the Census enumeration. The results from this analysis are presented in appendix B. We continue to see that rainfall shocks and the GQ program are valid instruments.Table B.3 shows that a 1% increase in long-term migration from j to i increased urban population growth at i by 1.48%, and out-migration from i had no impact.

The final concern with identification is if rainfall shocks and the GQ program implementation led to the migration of firms as well as individuals across states. If firms moved, there may have been a restructuring of industries in regions which did not receive the shocks in ways that can affect population growth and housing supply. This, however, is unlikely to be a concern in our identification. First, negative rainfall shocks predominantly affect agriculture in India (Jayachandran, 2006). Any sectoral reallocation of labor resulting from such shocks will likely not lead to the movement of firms (Emerick, 2018).

Second, past literature has shown that the GQ program led to firm concentration *along* the highways (Ghani et al., 2016). If only the distant state was a recipient of the GQ program, firm location in such states (as implied by Ghani et al. (2016)) would only impact local housing markets through migration. In this case, there would be no endogeneity concern with firm relocation.

In cases where both the distant and local states were recipients of the GQ program, only firm movement across non-contiguous states will be an identification concern. This is because the robustness check with non-contiguous state shocks eliminates contiguous distant and local state pairs which were GQ recipients.

Our robustness check using non-contiguous states alleviates the firm relocation problem. Because of differences in property rights, laws, policies, regulations, culture, and languages, firms rarely move across district boundaries, let alone from one non-contiguous state to another. Abeberese and Chen (2021) found that only 12% of over 10,000 firms surveyed by the Annual Survey of Industries (ASI) between 1999 and 2007 relocated across districts of India, with the average firm moving only 42 miles (68 kms).

5.2 State-level Elasticities

To get a sense of spatial heterogeneity in the overall national-level housing supply elasticity estimates of India, we provide some state-level elasticity figures for formal housing units. We do this by exploiting district-level variation within states for 12 of the largest states in India.[30] We estimate only the formal supply elasticities because not enough districts reported informal rents during the study period.[31] The formal housing supply elasticities are given in table 6. For comparison, table 6 also presents the Metropolitan Statistical Area (MSA) in the US that is closest to an Indian state in terms of its supply elasticity estimate based on Saiz (2010).[32]

Table 6: State-level formal supply elasticities

	Urban population in 2011 (millions)	Formal elasticity	Similar US MSAs (Saiz, 2010)
State	(1)	(2)	(3)
Maharashtra	51	3.06	Austin, TX
Odisha	7	2.05	Mobile, AL
Tamil Nadu	35	1.92	Fresno, CA
Andhra Pradesh	28	1.63	Phoenix, AZ
Gujarat	26	1.31	Las Vegas, NV
Madhya Pradesh	20	1.25	Detroit, MI
Uttar Pradesh	44	1.17	Newark, NJ
Rajasthan	17	1.06	Jacksonville, FL
Karnataka	24	0.75	New York, NY
Haryana	9	0.54	Miami, FL
Bihar	12	0.49	
West Bengal	29	0.38	

Source: Authors' calculations and Saiz (2010).
Note: All reported states have observations on formal housing rents and quantities for at least 14 districts. States arranged in decreasing order of elasticity values. All elasticity values rounded off to two decimal places. Last column shows MSAs in the United States (US) that have comparable housing supply elasticities. There are no metropolitan areas in the US with elasticity figures comparable to Bihar and West Bengal. Among Saiz's estimates, Miami is the least supply elastic MSA in the US with an elasticity value of 0.6.

Table 6 shows that there is considerable variation in formal supply elasticities across states of India. On the higher end, Maharashtra has a supply elasticity of 3.06, comparable to Austin, Texas, in the US. Considering that cities in Texas are some of the most supply elastic in the US, Maharashtra's elasticity reflects a very responsive market. Given that cities in Maharashtra have low FARs ranging from 1 to 2 (Sridhar, 2010) the high estimated supply elasticity seems surprising. However, many cities in Maharashtra utilizes Transferable Development Rights that relax these strict FARs, making formal housing markets more responsive (Phatak, 2013).

It is also not surprising that states like Orissa, Andhra Pradesh, and Gujarat have relatively higher elasticities. A report from The World Bank (2009) on doing business in India ranked Bhubaneshwar (capital of Orissa), Hyderabad (then capital of Andhra Pradesh) and Ahmedabad,

[30] We report the 12 states in which at least 14 districts reported a formal rent figure for both 2001 and 2011.
[31] For instance, Uttar Pradesh, the largest state in India, reported informal rents for only 5 districts in both 2001 and 2011.
[32] Baum-Snow and Han (2024) have also estimated housing supply elasticities for US metropolitan areas using Census-tract level elasticities. However, our estimation exercise more closely resembles the Saiz (2010) estimation of metropolitan-level elasticities and so we compare our results to his.

as having among the swiftest building construction permit processes in India.[33] The states of Bihar and West Bengal have the lowest formal housing supply elasticities in India, with values of 0.49 and 0.38, respectively. These elasticity values are lower than the least supply elastic MSAs in the US: Miami and Los Angeles-Long Beach. The same World Bank report estimated that the average time required to obtain a construction permit in Kolkata (capital of West Bengal) was 258 days, higher than any other city in India.

6 Conclusion

Indian cities are adding more people than any others in the world. Moreover, many urban dwellers in Indian and other developing country cities live in informal housing. Despite this, the economic literature on housing in urban India and informal housing more generally is sparse. We fill these gaps by estimating both the informal and formal supply elasticities of housing in urban India. Our estimated elasticities also allows us to observe gentrification in Indian cities.

We apply the Rosen-Roback spatial equilibrium framework to construct demand shifters. We show that shocks that induce inter-regional migration affect housing demand through population growth in regions that did not receive shocks. We use negative rainfall shocks and a highway upgrade program implementation in one state as demand shifters for another. We find that both negative rainfall shocks and the highway upgrade implementation induced inter-state migration in India during the 2000s. This increased migration led to urban population growth, and hence, increased demand for housing in regions which did not receive the shocks. Using these demand shifters, we estimate national-level housing supply elasticity figures for urban India's informal and formal housing units.

We show that the informal housing supply elasticity in India is negative. This negative informal elasticity estimate implies the existence of gentrification in India that happens through the conversion of informal houses into formal buildings. Gentrification generally leads to displacement, and the shock of eviction is highly costly to those who are displaced. Consequently, gentrification creates a Pareto problem: while it may lead to an overall increase in wealth, it diminishes the well-being of those who are displaced. Gechter and Tsivanidis (2023) find that within the Mumbai context, the surplus created by gentrification there is more than sufficient to compensate those displaced by it.

Our formal housing supply elasticity estimates for India are similar to estimates for the USA, which are inelastic (Baum-Snow and Han, 2024; Saiz, 2010). Our results on formal housing point to the role of regulations in restricting the responsiveness of housing markets in India (see Bertaud and Brueckner, 2005). Unfortunately, no one has to this point developed a land use regulatory index for subgeorgaphies of India similar to the Wharton Residential Land Use Regulatory Index for metropolitan areas in the US.[34] Developing such an index for India and

[33] The number of days taken to obtain a construction permit in Ahmedabad, Bhubaneshwar, and Hyderabad are, respectively, 144, 149, and 80 (The World Bank, 2009).

[34] Several studies have used the Wharton Residential Land Use Regulatory Index to establish a relationship between

investigating its relationship to housing supply elasticity is a ripe topic for future research.

National and state-level housing supply elasticity estimates do not paint an accurate picture of metropolitan-level elasticities. Even withing a state, differences in city sizes and regulations can produce a range of metropolitan level elascticities. Nevertheless, aggregate measures of supply elasticities at the state and national levels provide useful benchmark for policymakers. Further research with metropolitan-level price and new construction data would be required to provide granular elasticity estimates.

regulatory stringency and housing supply elasticities in the US (see Green et al., 2005; Gyourko and Molloy, 2015; Paciorek, 2013).

References

Abeberese, A. B. and Chen, M. (2021). Intranational trade costs, product scope and productivity: Evidence from India's Golden Quadrilateral project. *Journal of Development Economics*, page 102791. (Cited on 23, 25)

Accetturo, A., Lamorgese, A., Mocetti, S., and Pellegrino, D. (2021). Housing supply elasticity and growth: Evidence from Italian cities. *Journal of Economic Geography*, 21(3):367–396. (Cited on 5)

Bartik, T. J. (1993). Who benefits from local job growth: Migrants or the original residents? *Regional studies*, 27(4):297–311. (Cited on 16)

Baum-Snow, N. and Han, L. (2024). The microgeography of housing supply. *Journal of Political Economy*, 132(6):1897–1946. (Cited on 2, 5, 22, 26, 27)

Bertaud, A. (2018). Order without design: How markets shape cities. (Cited on 9)

Bertaud, A. and Brueckner, J. K. (2005). Analyzing building-height restrictions: Predicted impacts and welfare costs. *Regional Science and Urban Economics*, 35(2):109–125. (Cited on 5, 27)

Bhan, G. (2009). "This is no longer the city I once knew". Evictions, the urban poor and the right to the city in millennial Delhi. *Environment and Urbanization*, 21(1):127–142. (Cited on 2, 5, 21)

Bhavnani, R. R. and Lacina, B. (2017). Fiscal federalism at work? Central responses to internal migration in India. *World Development*, 93:236–248. (Cited on 3, 9, 15, 16, 23, 24)

Boustan, L. P. (2010). Was postwar suburbanization "white flight"? Evidence from the black migration. *The Quarterly Journal of Economics*, 125(1):417–443. (Cited on 3, 12)

Brueckner, J. K. (1987). The structure of urban equilibria: A unified treatment of the Muth-Mills model. *Handbook of Regional and Urban Economics*, 2(20):821–845. (Cited on 12, 22)

Brueckner, J. K. and Selod, H. (2009). A theory of urban squatting and land-tenure formalization in developing countries. *American Economic Journal: Economic Policy*, 1(1):28–51. (Cited on 4)

Brueckner, J. K. and Sridhar, K. S. (2012). Measuring welfare gains from relaxation of land-use restrictions: The case of India's building-height limits. *Regional Science and Urban Economics*, 42(6):1061–1067. (Cited on 5)

Bryan, G., Chowdhury, S., and Mobarak, A. M. (2014). Underinvestment in a profitable technology: The case of seasonal migration in Bangladesh. *Econometrica*, 82(5):1671–1748. (Cited on 16, 23)

Cavalcanti, T., Da Mata, D., and Santos, M. (2019). On the determinants of slum formation. *The Economic Journal*, 129(621):1971–1991. (Cited on 4)

Census of India (2001-2011a). *D-series Migration Tables (2001-2011)*. Office of the Registrar General and Census Commissioner, New Delhi. (Cited on 6)

Census of India (2001-2011b). *H-series Housing Tables (2001-2011)*. Office of the Registrar General and Census Commissioner, New Delhi. (Cited on 7)

Census of India (2001-2011d). *Primary Census Abstract Tables (2001-2011)*. Office of the Registrar General and Census Commissioner New Delhi. (Cited on 6, 7)

Census of India (2011c). *Indian Administrative Atlases*. Office of the Registrar General and Census Commissioner New Delhi. (Cited on 3, 36)

Dasgupta, B. and Lall, S. V. (2009). Assessing benefits of slum upgrading programs in second-best settings. In *urban Land markets: Improving Land management for successful urbanization*, pages 225–251. Springer. (Cited on 2, 5, 21)

Diamond, R. (2017). Housing supply elasticity and rent extraction by state and local governments. *American Economic Journal: Economic Policy*, 9(1):74–111. (Cited on 5)

Donaldson, D. (2018). Railroads of the Raj: Estimating the impact of transportation infrastructure. *American Economic Review*, 108(4-5):899–934. (Cited on 13)

Duranton, G., Ghani, S. E., Goswami, A. G., Kerr, W., and Kerr, W. R. (2015). The misallocation of land and other factors of production in India. *World Bank Policy Research Working Paper*, (7221). (Cited on 5)

Dutta, A. (2024). Rising temperatures, informal housing, and inequality. *Available at SSRN 4785001*. (Cited on 8)

Dutta, A. and Gandhi, S. (2023). Land use deregulation and housing supply under institutional frictions. *Available at SSRN 4411151*. (Cited on 5)

Dutta, A., Gandhi, S., and Green, R. K. (2022). Do urban regulations exacerbate rural-urban inequality? Evidence from rent control in India. *Evidence from rent control in India (April 25, 2022)*. (Cited on 8)

Dutta, A. and Randolph, G. (2022). Telecom expansion and internal migrants in indian cities. *Available at SSRN 4094458*. (Cited on 8, 9)

Elisseeff, V. (2000). *The silk roads: Highways of culture and commerce*. Berghahn Books. (Cited on 24)

Emerick, K. (2018). Agricultural productivity and the sectoral reallocation of labor in rural India. *Journal of Development Economics*, 135:488–503. (Cited on 25)

Gandhi, S., Green, R. K., and Patranabis, S. (2022). Insecure property rights and the housing market: Explaining India's housing vacancy paradox. *Journal of Urban Economics*, page 103490. (Cited on 5)

Gandhi, S., Tandel, V., Tabarrok, A., and Ravi, S. (2021). Too slow for the urban march: Litigations and the real estate market in Mumbai, India. *Journal of Urban Economics*, 123:103330. (Cited on 5)

Gechter, M. and Tsivanidis, N. (2023). Spatial Spillovers from High-Rise Developments: Evidence from the Mumbai Mills. *Working Paper*. (Cited on 2, 5, 6, 21, 27)

Ghani, E., Goswami, A. G., and Kerr, W. R. (2016). Highway to success: The impact of the Golden Quadrilateral project for the location and performance of Indian manufacturing. *The Economic Journal*, 126(591):317–357. (Cited on 4, 7, 16, 23, 25, 36)

Ghose, D. (2021). Trade, Internal Migration, and Human Capital. (Cited on 9)

Glaeser, E. (2011). *Triumph of the city: How urban spaces make us human*. Pan Macmillan. (Cited on 9)

Glaeser, E. L., Gyourko, J., and Saiz, A. (2008). Housing supply and housing bubbles. *Journal of urban Economics*, 64(2):198–217. (Cited on 22)

Glaeser, E. L., Gyourko, J., and Saks, R. (2005). Why is Manhattan so expensive? Regulation and the rise in housing prices. *The Journal of Law and Economics*, 48(2):331–369. (Cited on 5)

Gorback, C. S. and Keys, B. J. (2020). Global capital and local assets: House prices, quantities, and elasticities. Technical report, National Bureau of Economic Research. (Cited on 3)

Green, R. K., Malpezzi, S., and Mayo, S. K. (2005). Metropolitan-specific estimates of the price elasticity of supply of housing, and their sources. *American Economic Review*, 95(2):334–339. (Cited on 5, 22, 28)

Gyourko, J. and Molloy, R. (2015). Regulation and housing supply. In *Handbook of regional and urban economics*, volume 5, pages 1289–1337. Elsevier. (Cited on 28)

Gyourko, J., Saiz, A., and Summers, A. (2008). A new measure of the local regulatory environment for housing markets: The Wharton Residential Land Use Regulatory Index. *Urban Studies*, 45(3):693–729. (Cited on 22)

Harari, M. (2020). Cities in bad shape: Urban geometry in India. *American Economic Review*, 110(8):2377–2421. (Cited on 5)

Harari, M. and Wong, M. (2024). Slum upgrading and long-run urban development: Evidence from indonesia. (Cited on 2, 5, 21)

Henderson, J. V., Regan, T., and Venables, A. J. (2021). Building the city: From slums to a modern metropolis. *The Review of Economic Studies*, 88(3):1157–1192. (Cited on 2, 4, 5, 15, 22)

Henderson, J. V., Venables, A. J., Regan, T., and Samsonov, I. (2016). Building functional cities. *Science*, 352(6288):946–947. (Cited on 2)

Hsieh, C.-T. and Moretti, E. (2019). Housing constraints and spatial misallocation. *American Economic Journal: Macroeconomics*, 11(2):1–39. (Cited on 2)

Imbert, C., Seror, M., Zhang, Y., and Zylberberg, Y. (2022). Migrants and firms: Evidence from china. *American Economic Review*, 112(6):1885–1914. (Cited on 3)

India Meteorological Department, P. (1901-2017). *India Meteorological Department Datasets*. http://www.imd.gov.in. (Cited on 6, 7)

Jayachandran, S. (2006). Selling labor low: Wage responses to productivity shocks in developing countries. *Journal of Political Economy*, 114(3):538–575. (Cited on 3, 15, 23, 25)

Kone, Z. L., Liu, M. Y., Mattoo, A., Ozden, C., and Sharma, S. (2018). Internal borders and migration in India. *Journal of Economic Geography*, 18(4):729–759. (Cited on 9)

Kuffer, M., Pfeffer, K., and Sliuzas, R. (2016). Slums from space—15 years of slum mapping using remote sensing. *Remote Sensing*, 8(6):455. (Cited on 6)

Labor Bureau of India, D. (2001-2012). *Consumer Price Index Series*. http://www.labourbureau.gov.in/. (Cited on 6, 7, 8)

Malpezzi, S. and Maclennan, D. (2001). The long-run price elasticity of supply of new residential construction in the United States and the United Kingdom. *Journal of Housing Economics*, 10(3):278–306. (Cited on 5)

Marx, B., Stoker, T., and Suri, T. (2013). The economics of slums in the developing world. *Journal of Economic perspectives*, 27(4):187–210. (Cited on 4)

McLaughlin, R. B. (2012). New housing supply elasticity in Australia: a comparison of dwelling types. *The Annals of Regional Science*, 48(2):595–618. (Cited on 5)

Michaels, G., Nigmatulina, D., Rauch, F., Regan, T., Baruah, N., and Dahlstrand, A. (2021). Planning ahead for better neighborhoods: Long-run evidence from tanzania. *Journal of Political Economy*, 129(7):2112–2156. (Cited on 2, 21)

Mikusheva, A. and Sun, L. (2022). Inference with many weak instruments. *The Review of Economic Studies*, 89(5):2663–2686. (Cited on 23)

Mills, E. S. (1967). An aggregative model of resource allocation in a metropolitan area. *The American Economic Review*, 57(2):197–210. (Cited on 12)

Mishra, A. and Liu, S. C. (2014). Changes in precipitation pattern and risk of drought over india in the context of global warming. *Journal of Geophysical Research: Atmospheres*, 119(13):7833–7841. (Cited on 7)

Mitra, A. (2010). Migration, livelihood and well-being: Evidence from indian city slums. *Urban Studies*, 47(7):1371–1390. (Cited on 10)

Molloy, R., Smith, C. L., and Wozniak, A. (2011). Internal migration in the United States. *Journal of Economic Perspectives*, 25(3):173–96. (Cited on 12, 23)

Morten, M. (2019). Temporary migration and endogenous risk sharing in village India. *Journal of Political Economy*, 127(1):1–46. (Cited on 11, 16, 23)

Munshi, K. and Rosenzweig, M. (2016). Networks and misallocation: Insurance, migration, and the rural-urban wage gap. *American Economic Review*, 106(1):46–98. (Cited on 16, 24)

Nagpal, G. and Gandhi, S. (2024). Scaling heights: Affordability implications of zoning deregulation in india. *Working paper*. (Cited on 5)

National Sample Survey Organization of India (1999-2012). *Indian National Sample Surveys, Rounds 55, 58, 66, and 69*. http://mospi.nic.in. (Cited on 6, 7, 18)

National Sample Survey Organization of India (2008). *Indian National Sample Surveys, Round 64*. http://mospi.nic.in. (Cited on 10, 25)

Niu, D., Sun, W., and Zheng, S. (2021). The role of informal housing in lowering China's urbanization costs. *Regional Science and Urban Economics*, 91:103638. (Cited on 4, 9)

Olea, J. L. M. and Pflueger, C. (2013). A robust test for weak instruments. *Journal of Business & Economic Statistics*, 31(3):358–369. (Cited on 18, 23)

Paciorek, A. (2013). Supply constraints and housing market dynamics. *Journal of Urban Economics*, 77:11–26. (Cited on 2, 28)

Phatak, V. (2013). *Land Based Fiscal Tools and Practices for Generating Additional Financial Resources*. Ministry of Urban Development, Government of India. (Cited on 26)

Poterba, J. M. (1984). Tax subsidies to owner-occupied housing: an asset-market approach. *The Quarterly Journal of Economics*, 99(4):729–752. (Cited on 12)

Pradhan, K. C. (2017). *Unacknowledged urbanisation: The new Census Towns in India*. Springer. (Cited on 15)

Quigley, J. M. and Raphael, S. (2005). Regulation and the high cost of housing in California. *American Economic Review*, 95(2):323–328. (Cited on 5)

Rains, E. and Krishna, A. (2020). Precarious gains: Social mobility and volatility in urban slums. *World Development*, 132:105001. (Cited on 6, 10, 21)

Rains, E., Krishna, A., and Wibbels, E. (2019). Combining satellite and survey data to study Indian slums: Evidence on the range of conditions and implications for urban policy. *Environment and Urbanization*, 31(1):267–292. (Cited on 6, 21)

Roback, J. (1982). Wages, rents, and the quality of life. *Journal of Political Economy*, 90(6):1257–1278. (Cited on 2, 11)

Rosen, S. (1979). Wage-based indexes of urban quality of life. *Current Issues in Urban Economics*, pages 74–104. (Cited on 2, 11)

Rosenzweig, M. R. and Udry, C. (2014). Rainfall forecasts, weather, and wages over the agricultural production cycle. *American Economic Review*, 104(5):278–83. (Cited on 16, 23)

Saiz, A. (2010). The geographic determinants of housing supply. *The Quarterly Journal of Economics*, 125(3):1253–1296. (Cited on 2, 4, 5, 22, 26, 27)

Saks, R. E. and Wozniak, A. (2011). Labor reallocation over the business cycle: New evidence from internal migration. *Journal of Labor Economics*, 29(4):697–739. (Cited on 12, 23)

Sarsons, H. (2015). Rainfall and conflict: A cautionary tale. *Journal of Development Economics*, 115:62–72. (Cited on 23)

Sekhri, S. and Storeygard, A. (2014). Dowry deaths: Response to weather variability in India. *Journal of Development economics*, 111:212–223. (Cited on 23)

Sood, A. (2019). Land Market Frictions and Manufacturing in India. *Available at SSRN 3468519*. (Cited on 5)

Sridhar, K. S. (2010). Impact of land use regulations: Evidence from India's cities. *Urban Studies*, 47(7):1541–1569. (Cited on 5, 26)

Srivastava, R. (2011). Internal migration in india. *Human Development in India*. (Cited on 10)

Tandel, V., Gandhi, S., and Tabarrok, A. (2023). Building networks: Investigating the quid pro quo between local politicians & developers. *Journal of Development Economics*, 164:103138. (Cited on 5)

Tandel, V., Hiranandani, K., and Kapoor, M. (2019). What's in a definition? a study on the suitability of the current urban definition in india through its employment guarantee programme. *Journal of Asian Economics*, 60:69–84. (Cited on 15)

Thapar, R. (2015). *The Penguin history of early India: from the origins to AD 1300*. Penguin UK. (Cited on 24)

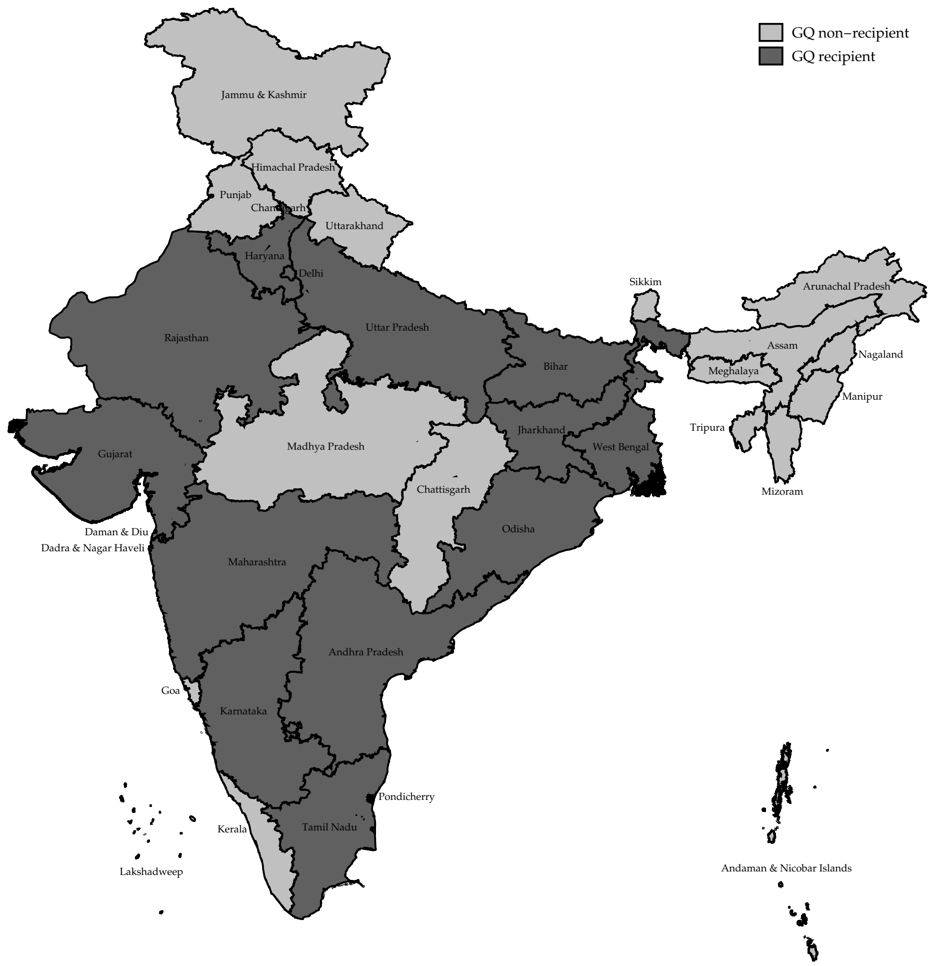
The World Bank (2009). *Doing business in India 2009*. The International Bank for Reconstruction and Development / The World Bank. (Cited on 26, 27)

United Nations (2023). *The sustainable development goals report 2023*. UN. (Cited on 2)

Wang, S., Chan, S. H., and Xu, B. (2012). The estimation and determinants of the price elasticity of housing supply: Evidence from China. *Journal of Real Estate Research*, 34(3):311–344. (Cited on 5)

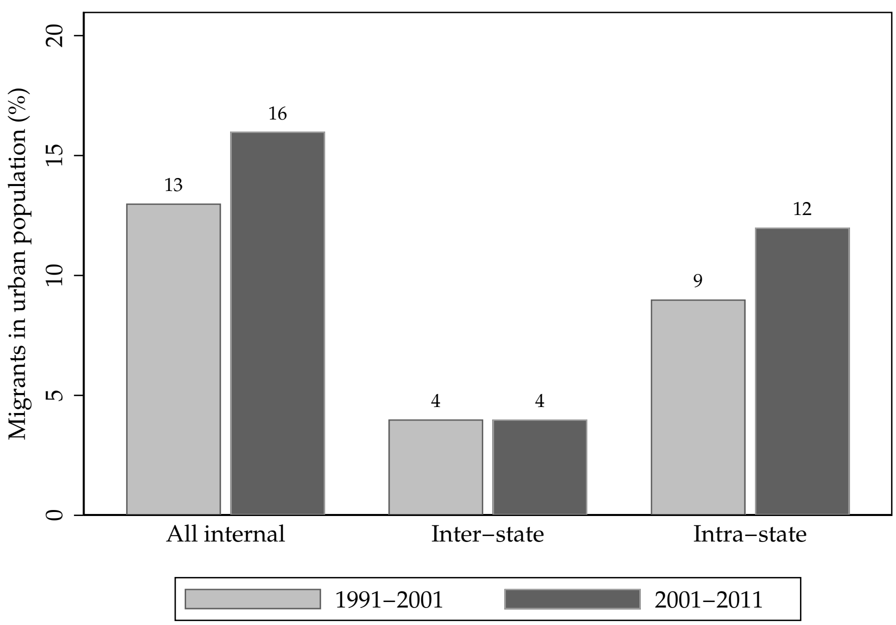
Appendix A Figures and tables

Figure A.1: Map of Golden Quadrilateral recipient states in India



Data Source: Based on Census of India (2011c); Ghani et al. (2016).
Note: Figure presents a map of India with the 35 states and union territories demarcated. Light-colored states were not recipients of the National Highways Development Project Phase I or the Golden Quadrilateral (GQ) highway upgrade project. Dark-colored states were part of the GQ program.

Figure A.2: Share of decadal migrants in urban population of India



Source: Author's calculations.
Note: Figure presents the urban population's share of rural-urban and urban-urban migrants that moved during 1991-2001 and 2001-2010 by migrants' last residence (same or different state). Bars labeled with their corresponding values.

Table A.1: Golden Quadrilateral and urban commodity prices

	Dependent variable = ΔLog price
$\Delta = 2011 - 2001$	(1)
GQ highway dummy	0.037
	(0.042)
ΔLog weekly wage	-0.017
	(0.115)
ΔLog urban population	-0.271***
	(0.063)
ΔUrban surface area	-0.022
	(0.027)
ΔUrban surface area squared	-0.000
	(0.002)
N	33
R^2	0.435

Source: Authors' calculations.
Note: Table presents first-difference regression estimates using a sample of 33 states in India between 2001 and 2011. The states of Sikkim and Arunachal Pradesh are excluded because of missing urban area data. Differences are calculated as first differences of values between the years 2001 and 2011. Commodity prices for a given state are calculated by first obtaining the prices of non-housing commodity baskets in the state's urban areas, then multiplying the computed prices with the national average of each basket consumed by urban households between 2001 and 2011, and finally adding the products for all baskets. The GQ highway dummy equals one if the state was a recipient of the GQ program, and zero otherwise. Urban surface area is in 1000 sq. miles. Robust standard errors presented in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.2: Distant shock-induced local urbanization and housing demand with median rooms

	First-stage	Second-stage	
	Dependent variable = ΔLog indicators		
	Urban population at i	Housing units at i	
		Informal	Formal
$\Delta = 2011 - 2001$	(1)	(2)	(3)
Δ Log urban population at i		0.237***	1.847***
		(0.024)	(0.026)
Δ #Months rainfall $< 80\%$ at j	0.011***		
	(0.000)		
GQ highway at j dummy	0.124***		
	(0.005)		
Δ Log consumption at i	0.009	-0.210***	0.040**
	(0.013)	(0.013)	(0.018)
Δ Urban surface area at i	5.115***	0.502**	-1.462***
	(0.202)	(0.205)	(0.225)
Δ Urban surface area at i squared	-5.513***	-1.380***	1.130***
	(0.417)	(0.323)	(0.309)
Δ Median no. rooms per unit at i	-0.007	-0.142***	0.035***
	(0.008)	(0.007)	(0.007)
F-stat on excluded instruments	983		
Montiel-Olea-Pflueger Eff. F-stat.	928		
Sargan-Hansen J-stat p-value		0.728	0.423
N	4,896	4,896	4,896
Adj. R^2	0.603		

Source: Authors' calculations.
Note: Table presents two-stage least squares first-difference regression estimates using a panel of district-state pairs in India between 2001 and 2011. District i is the local region and state j the distant. Excluding missing pairs, there are 4,896 $i - j$ district-state pairs consisting of the 35 states and union territories and 144 districts of India in our sample. Differences are calculated as first differences of values between the years 2001 and 2011. Log urban population at i is instrumented by the decadal change in the number of months when rainfall was less than 80% of the long-term normal in state j and the GQ at j dummy, which equals one if state j was recipient of the Golden Quadrilateral highway upgrade program, and zero otherwise. Consumption, urban surface area, urban surface area squared, and median number of rooms per housing unit at i are the controls. Urban surface area unit in 1000 sq. miles. Robust standard errors presented in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.3: Shock-induced local urbanization and formal housing demand with full sample

	First-stage	Second-stage
	Dependent variable = ΔLog indicators	
	Urban population at i	Formal housing units at i
$\Delta = 2011 - 2001$	(1)	(2)
Δ Log urban population at i		1.830***
		(0.016)
Δ#Months rainfall $< 80\%$ at j	0.011***	
	(0.000)	
GQ highway at j dummy	0.118***	
	(0.003)	
ΔLog consumption at i	0.023***	0.002
	(0.006)	(0.007)
ΔUrban surface area at i	6.581***	-2.617***
	(0.153)	(0.184)
ΔUrban surface area at i squared	-7.556***	2.904***
	(0.334)	(0.258)
F-stat on excluded instruments	2,194	
Montiel-Olea-Pflueger Eff. F-stat.	2,088	
Sargan-Hansen J-stat p-value		0.395
N	11,526	11,526
Adj. R^2	0.638	

Source: Authors' calculations.

Note: Table presents two-stage least squares first-difference regression estimates using a panel of district-state pairs in India between 2001 and 2011. District i is the local region and state j the distant. Excluding missing pairs, there are 11,526 $i - j$ district-state pairs consisting of the 35 states and union territories and 339 districts of India in our complete sample on formal rents. Differences are calculated as first differences of values between the years 2001 and 2011. Log urban population at i is instrumented by the decadal change in the number of months when rainfall was less than 80% of the long-term normal in state j and the GQ at j dummy, which equals one if state j was recipient of the Golden Quadrilateral highway upgrade program, and zero otherwise. Consumption, urban surface area, and urban surface area squared at i are the controls. Urban surface area unit in 1000 sq. miles. Robust standard errors presented in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.4: Distant shock-induced local urbanization and housing demand with one IV

Panel (a): Only rainfall instrument			
	First-stage	Second-stage	
	Dependent variable = ΔLog indicators		
	Urban population at i	Housing units at i	
		Informal	Formal
$\Delta = 2011 - 2001$	(1)	(2)	(3)
ΔLog urban population at i		0.200***	1.866***
		(0.030)	(0.030)
Δ#Months rainfall $<$ 80% at j	0.014***		
	(0.000)		
Controls	Y	Y	Y
F-stat on excluded instruments	1,321		
Montiel-Olea-Pflueger Eff. F-stat.	1,240		
N	4,896	4,896	4,896
Adj. R^2	0.563		
Panel (b): Only GQ instrument			
	First-stage	Second-stage	
	Dependent variable = ΔLog indicators		
	Urban population at i	Housing units at i	
		Informal	Formal
$\Delta = 2011 - 2001$	(1)	(2)	(3)
ΔLog urban population at i		0.213***	1.839***
		(0.033)	(0.033)
GQ highway at j dummy	0.176***		
	(0.005)		
Controls	Y	Y	Y
F-stat on excluded instruments	1,433		
Montiel-Olea-Pflueger Eff. F-stat.	980		
N	4,896	4,896	4,896
Adj. R^2	0.563		

Source: Authors' calculations.

Note: Table presents two-stage least squares first-difference regression estimates using a panel of district-state pairs in India between 2001 and 2011. District i is the local region and state j the distant. Excluding missing pairs, there are 4,896 $i - j$ district-state pairs consisting of the 35 states and union territories and 144 districts of India in our sample. Differences are calculated as first differences of values between the years 2001 and 2011. In panel (a), log urban population at i is instrumented by the decadal change in the number of months when rainfall was less than 80% of the long-term normal in state j. In panel (b), log urban population at i is instrumented by the GQ at j dummy, which equals one if state j was recipient of the Golden Quadrilateral highway upgrade program, and zero otherwise. Consumption, urban surface area, and urban surface area squared at i are the controls. Robust standard errors presented in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.5: Housing demand shifters and inverse supply elasticity estimation with one IV

Panel (a): Rainfall instrument alone				
	First-stage		Second-stage	
	Dependent variable = ΔLog indicators			
	Housing units at i		Rents at i	
	Informal	Formal	Informal	Formal
$\Delta = 2011 - 2001$	(1)	(2)	(3)	(4)
ΔLog informal units at i			-2.089***	
			(0.679)	
ΔLog formal units at i				0.622***
				(0.044)
Δ#Months rainfall $<$ 80% at j	0.003***	0.027***		
	(0.000)	(0.001)		
Controls	Y	Y	Y	Y
F-stat on excluded instruments	38.2	2,059		
Montiel-Olea-Pflueger Eff. F-stat.	38.1	1,875		
N	4,896	4,896	4,896	4,896
Adj. R^2	0.072	0.607		
Panel (b): GQ instrument alone				
	First-stage		Second-stage	
	Dependent variable = ΔLog indicators			
	Housing units at i		Rents at i	
	Informal	Formal	Informal	Formal
$\Delta = 2011 - 2001$	(1)	(2)	(3)	(4)
ΔLog informal units at i			-2.028***	
			(0.695)	
ΔLog formal units at i				0.609***
				(0.049)
GQ highway at j dummy	0.037***	0.323***		
	(0.006)	(0.006)		
Controls	Y	Y	Y	Y
F-stat on excluded instruments	35.7	2,596		
Montiel-Olea-Pflueger Eff. F-stat.	35.8	1,391		
N	4,896	4,896	4,896	4,896
Adj. R^2	0.071	0.578		

Source: Authors' calculations.
Note: Table presents two-stage least squares first-difference regression estimates using a panel of district-state pairs in India between 2001 and 2011. District i is the local region and state j the distant. Excluding missing pairs, there are 4,896 $i - j$ district-state pairs consisting of 144 districts and 35 states and union territories of India in our sample. Differences are calculated as first differences of values between the years 2001 and 2011. In panel (a), log housing units at i are instrumented by the decadal change in the number of months when rainfall was less than 80% of the long-term normal in state j. In panel (b), log housing units at i are instrumented by the GQ at j dummy, which equals one if state j was recipient of the Golden Quadrilateral highway upgrade program, and zero otherwise. Consumption, urban surface area, and urban surface area squared at i are the controls. Robust standard errors presented in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.6: Ordinary Least Squares estimation of housing supply equation

	Dependent variable = ΔLog rent	
	Informal	Formal
$\Delta = 2011 - 2001$	(1)	(2)
ΔLog informal units at i	-0.065	
	(0.043)	
ΔLog formal units at i		0.313***
		(0.023)
ΔLog consumption at i	-0.142***	0.492***
	(0.048)	(0.041)
ΔUrban surface area at i	1.999***	1.510***
	(0.471)	(0.377)
ΔUrban surface area at i	1.206	-1.423**
	(0.781)	(0.722)
N	4,896	4,896
Adj. R^2	0.024	0.139

Source: Authors' calculations.
Note: Table presents Ordinary Least Squares first-difference regression estimates using a panel of district-state pairs in India between 2001 and 2011. District i is the local region and state j the distant. Excluding missing pairs, there are 4,896 $i - j$ district-state pairs consisting of 144 districts and 35 states and union territories of India in our sample. Differences are calculated as first differences of values between the years 2001 and 2011. Consumption, urban surface area, and urban surface area squared at i are the controls. Urban surface area unit in 1000 sq. miles. Robust standard errors presented in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Appendix B IV Robustness checks

We conduct two robustness checks to address endogeneity concerns with the use of rainfall shocks and the GQ program implementation in the distant state as instruments for migration, urban population growth, and housing demand in the local state. In the first robustness check, we estimate the impact of non-contiguous distant state shocks on inter-state migration and the resulting impact of such migration on urban population growth in the local state. In the second, we estimate the impact of distant shocks on long-term migration.

B.1 Non-contiguous state shocks

To address concerns that rainfall shocks and the GQ program implementation may have housing supply effects across neighboring states we conduct robustness checks by eliminating non-contiguous state pairs in our analysis. In the regressions using equations (7) and (8), we restrict the universe of shocks z_j to include only those that occurred in non-contiguous distant states. In other words, an observation in these regressions is equivalent to an $i - j$ non-contiguous state pair.

The coefficient estimates in table B.1 are very similar in magnitude to those seen in table 3. Every additional month of rainfall levels below 80% of long-term normal in distant non-contiguous states increased migration from such states by 1%. A distant non-contiguous state's inclusion in the GQ program increased migration from such states by 29% and to such states by 22%. Just as in table 3, in-migration led to a positive significant impact on urban population growth, whereas out-migration had no impact. In table B.2, we present the housing supply elasticity estimation results with non-contiguous state shocks and do not find any changes from the main estimates given in table 5.

One contentious issue in the results given in table B.1 is that the instruments for migration from i to j (column (2) of table B.1) seem to be weak—they do not pass the F-test criterion. This is not an unexpected finding, given that non-contiguous states' shocks will produce lower volumes of migration compared to shocks that occur in all states. In other words, the strength of the instruments reduce by design. However, the F-stat in column (2) of table B.1 misses the strong IV criterion marginally. Moreover, our results from previous sections and this one suggest that the local housing demand shock from distant events (contiguous or not) is primarily through in-migration ($i \leftarrow j$) and not out-migration ($i \rightarrow j$). This is because out-migration has no impact on local urban population growth, whereas in-migration increases local urban population. As long as the latter holds true, we need not worry about the weak instrument problem in our supply estimation exercise. This is confirmed by the fact that both the negative rainfall shocks and the highway upgrade program at distant non-contiguous states are strong instruments for the number of housing units in the local region (see table B.2).

Table B.1: Distant non-contiguous states' shock-induced migration and local urbanization

	First-stage		Second-stage
	Dependent variable = ΔLog indicators		
	Migration 0-9 yrs.		Urban population
	$i \leftarrow j'$	$i \rightarrow j'$	at i
$\Delta = 2011 - 2001$	(1)	(2)	(3)
ΔLog migration 0-9 yrs. $i \leftarrow j'$			1.052***
			(0.277)
ΔLog migration 0-9 yrs. $i \rightarrow j'$			-0.447
			(0.442)
Δ#Months rainfall $< 80\%$ at j'	0.010***	-0.003	
	(0.004)	(0.003)	
GQ highway at j' dummy	0.287***	0.218***	
	(0.047)	(0.051)	
ΔLog consumption at i	-0.300*	-0.041	0.503**
	(0.168)	(0.246)	(0.196)
ΔUrban surface area at i	0.035	0.229***	0.169*
	(0.031)	(0.039)	(0.097)
ΔUrban surface area at i squared	0.007**	-0.016***	-0.021**
	(0.003)	(0.003)	(0.009)
F-stat on excluded instruments	26.0	9.13	
N	908	908	908
Adj. R^2	0.163	0.107	

Source: Authors' calculations.
Note: Table presents two-stage least squares first-difference regression estimates using a panel of non-contiguous state-state pairs in India between 2001 and 2011. State i is the local region and the non-contiguous state j' the distant. Excluding missing pairs, there are 908 $i - j'$ non-contiguous state pairs consisting of the 35 states and union territories in India. Migration $i \leftarrow j$ is in-migration to i and $i \rightarrow j$ is out-migration from i. Migrants are those who moved in the previous decade. Differences are calculated as first differences of values between the years 2001 and 2011. Log in- and out-migration are instrumented by the decadal change in the number of months when rainfall was less than 80% of the long-term normal in state j' and the GQ at j' dummy, which equals one if state j' was recipient of the Golden Quadrilateral highway upgrade program, and zero otherwise. Consumption, urban surface area, and urban surface area squared at i are the controls. Urban surface area unit in 1000 sq. miles. Robust standard errors presented in parentheses. * p < 0.10, ** p < 0.05, *** p < 0.01.

B.2 Long-term migration

We redefine the migration variables m_{ji} and m_{ij} in equations (7) and (8) to represent the number of individuals who moved during the 1-9 years before Census enumeration. In other words, we exclude migrants who moved less than a year before Census enumeration. We rerun the regressions using equations (7) and (8). The results from these regressions are given in table B.3.

As in the case of overall migration (table 3), the distant state's inclusion in the GQ highway upgrade program had a positive significant effect on long-term migration from j to i and from i to j, albeit with a smaller magnitude of effect. A distant state's inclusion in the GQ program increased long-term migration from j to i by 22% and from i to j by 19%. However, contrary to ta-

Table B.2: Supply elasticity estimation with non-contiguous state shocks

	First-stage		Second-stage	
	Dependent variable = ΔLog indicators			
	Housing units at i		Rents at i	
	Informal	Formal	Informal	Formal
$\Delta = 2011 - 2001$	(1)	(2)	(3)	(4)
ΔLog informal units at i			-2.026***	
			(0.609)	
ΔLog formal units at i				0.601***
				(0.039)
Δ#Months rainfall $<$ 80% at j'	0.001***	0.022***		
	(0.001)	(0.001)		
GQ highway at j' dummy	0.035***	0.219***		
	(0.008)	(0.008)		
ΔLog consumption at i	-0.192***	0.055***	-0.475***	0.439***
	(0.016)	(0.019)	(0.136)	(0.045)
ΔUrban surface area at i	1.943***	7.982***	6.671***	-2.141***
	(0.181)	(0.347)	(1.532)	(0.532)
ΔUrban surface area at i	-3.089***	-9.173***	-5.908**	2.819***
	(0.397)	(0.720)	(2.328)	(0.807)
Implied housing supply elasticities			-0.49	1.66
F-stat on excluded instruments	21.7	1,492		
Montiel-Olea-Pflueger Eff. F-stat.	22.1	1,261		
Sargan-Hansen J-stat p-value			0.514	0.400
N	4,224	4,224	4,224	4,224
Adj. R^2	0.072	0.661		

Source: Authors' calculations.
Note: Table presents two-stage least squares first-difference regression estimates using a panel of non-contiguous district-state pairs in India between 2001 and 2011. District i is the local region and a state j' that is non-contiguous with district i's state the distant. Excluding missing pairs, there are 4,224 $i - j'$ non-contiguous district-state pairs consisting of the 35 states and union territories and 144 districts of India in our sample. Differences are calculated as first differences of values between the years 2001 and 2011. Log housing units at i are instrumented by the decadal change in the number of months when rainfall was less than 80% of the long-term normal in state j' and the GQ at j' dummy, which equals one if state j' was recipient of the Golden Quadrilateral highway upgrade program, and zero otherwise. Consumption, urban surface area, and urban surface area squared at i are the controls. Urban surface area unit in 1000 sq. miles. Robust standard errors presented in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

ble 3, here we see that the negative rainfall shock at j does not affect long-term outward mobility. Instead, negative rainfall shocks in the distant state reduce long-term movement towards such regions. Even though the coefficient p-values are different, the overall results still indicate that the distant state shock-induced long-term migration led to urban population growth in the local state. Similar to table 3, here we see that 1% increase in long-term migration from j to i increased urban population growth at i by 1.48%, and out-migration from i had no impact. The higher impact of long-term (compared to short-term) migration on urban population growth is consistent

Table B.3: Distant shock-induced long-term migration and local urbanization

	First-stage		Second-stage
	Dependent variable = ΔLog indicators		
	Migration 1-9 yrs.		Urban population
	$i \leftarrow j$	$i \rightarrow j$	at i
$\Delta = 2011 - 2001$	(1)	(2)	(3)
ΔLog migration 1-9 yrs. $i \leftarrow j$			1.477***
			(0.413)
ΔLog migration 1-9 yrs. $i \rightarrow j$			-0.655
			(0.493)
Δ#Months rainfall $<$ 80% at j	0.005	-0.007**	
	(0.003)	(0.003)	
GQ highway at j dummy	0.215***	0.192***	
	(0.041)	(0.045)	
ΔLog consumption at i	-0.199	-0.283	0.295
	(0.149)	(0.221)	(0.277)
ΔUrban surface area at i	0.003	0.206***	0.230**
	(0.028)	(0.035)	(0.107)
ΔUrban surface area at i squared	0.008***	-0.014***	-0.028***
	(0.002)	(0.003)	(0.010)
F-stat on excluded instruments	16.3	10.7	
N	1,013	1,013	1,013
Adj. R^2	0.126	0.082	

Source: Authors' calculations.
Note: Table presents two-stage least squares first-difference regression estimates using a panel of state-state pairs in India between 2001 and 2011. State i is the local region and state j the distant. Excluding missing pairs, there are 1,028 $i-j$ state pairs consisting of 35 states and union territories across India. Migration $i \leftarrow j$ is in-migration to i and $i \rightarrow j$ is out-migration from i. Migrants are those who moved in the previous decade and were at their destinations for at least a year. Differences are calculated as first differences of values between the years 2001 and 2011. Log in- and out-migration are instrumented by the decadal change in the number of months when rainfall was less than 80% of the long-term normal in state j and the GQ at j dummy, which equals one if state j was recipient of the Golden Quadrilateral highway upgrade program, and zero otherwise. Consumption, urban surface area, and urban surface area squared at i are the controls. Urban surface area unit in 1000 sq. miles. Robust standard errors presented in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

with long-term movers settling down and starting families of their own at their destinations.

Appendix C Online tables

Table C.1: Distant shock-induced migration and local urbanization with controls

	First-stage		Second-stage
	Dependent variable = ΔLog indicators		
	Migration 0-9 yrs.		Urban population
	$i \leftarrow j$	$i \rightarrow j$	at i
$\Delta = 2011 - 2001$	(1)	(2)	(3)
ΔLog migration 0-9 yrs. $i \leftarrow j$			1.018***
			(0.244)
ΔLog migration 0-9 yrs. $i \rightarrow j$			-0.379
			(0.385)
Δ#Months rainfall < 80% at j	0.011***	-0.002	
	(0.003)	(0.003)	
GQ highway at j dummy	0.268***	0.213***	
	(0.043)	(0.046)	
ΔLog consumption at i	-0.297*	-0.062	0.477***
	(0.159)	(0.233)	(0.171)
ΔUrban surface area at i	0.039	0.220***	0.142*
	(0.028)	(0.036)	(0.082)
ΔUrban surface area at i squared	0.006**	-0.015***	-0.018**
	(0.002)	(0.003)	(0.007)
F-stat on excluded instruments	29.7	10.6	
N	1,028	1,028	1,028
Adj. R^2	0.179	0.121	

Source: Authors' calculations.
Note: Table presents two-stage least squares first-difference regression estimates using a panel of state-state pairs in India between 2001 and 2011. State i is the local region and state j the distant. Excluding missing pairs, there are 1,028 $i - j$ state pairs consisting of 35 states and union territories across India. Migration $i \leftarrow j$ is in-migration to i and $i \rightarrow j$ is out-migration from i. Migrants are those who moved in the previous decade. Differences are calculated as first differences of values between the years 2001 and 2011. Log in- and out-migration are instrumented by the decadal change in the number of months when rainfall was less than 80% of the long-term normal in state j and the GQ at j dummy, which equals one if state j was recipient of the Golden Quadrilateral highway upgrade program, and zero otherwise. Consumption, urban surface area, and urban surface area squared at i are the controls. Urban surface area unit is in 1000 sq. miles. Robust standard errors presented in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table C.2: Distant shock-induced local urbanization and housing demand with controls

	First-stage	Second-stage	
	Dependent variable = ΔLog indicators		
	Urban population at i	Housing units at i	
		Informal	Formal
$\Delta = 2011 - 2001$	(1)	(2)	(3)
ΔLog urban population at i		0.205***	1.855***
		(0.025)	(0.026)
Δ#Months rainfall $< 80\%$ at j	0.011***		
	(0.000)		
GQ highway at j dummy	0.124***		
	(0.005)		
ΔLog consumption at i	0.009	-0.198***	0.037**
	(0.013)	(0.014)	(0.018)
ΔUrban surface area at i	5.124***	0.824***	-1.540***
	(0.202)	(0.220)	(0.226)
ΔUrban surface area at i squared	-5.527***	-1.840***	1.242***
	(0.418)	(0.348)	(0.310)
F-stat on excluded instruments	1,018		
Montiel-Olea-Pflueger Eff. F-stat.	929		
Sargan-Hansen J-stat p-value		0.701	0.420
N	4,896	4,896	4,896
Adj. R^2	0.603		

Source: Authors' calculations.

Note: Table presents two-stage least squares first-difference regression estimates using a panel of district-state pairs in India between 2001 and 2011. District i is the local region and state j the distant. Excluding missing pairs, there are 4,896 $i - j$ district-state pairs consisting of the 35 states and union territories and 144 districts of India in our sample. Differences are calculated as first differences of values between the years 2001 and 2011. Log urban population at i is instrumented by the decadal change in the number of months when rainfall was less than 80% of the long-term normal in state j and the GQ at j dummy, which equals one if state j was recipient of the Golden Quadrilateral highway upgrade program, and zero otherwise. Consumption, urban surface area, and urban surface area squared at i are the controls. Urban surface area unit in 1000 sq. miles. Robust standard errors presented in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table C.3: Housing demand shifters and inverse supply elasticity estimation with controls

	First-stage		Second-stage	
	Dependent variable = ΔLog indicators			
	Housing units at i		Rents at i	
	Informal	Formal	Informal	Formal
$\Delta = 2011 - 2001$	(1)	(2)	(3)	(4)
ΔLog informal units at i			-2.060***	
			(0.564)	
ΔLog formal units at i				0.617***
				(0.037)
Δ#Months rainfall $<$ 80% at j	0.002***	0.021***		
	(0.000)	(0.001)		
GQ highway at j dummy	0.027***	0.226***		
	(0.007)	(0.007)		
ΔLog consumption at i	-0.196***	0.054***	-0.515***	0.453***
	(0.015)	(0.018)	(0.127)	(0.041)
ΔUrban surface area at i	1.877***	7.961***	6.665***	-2.176***
	(0.167)	(0.322)	(1.408)	(0.504)
ΔUrban surface area at i	-2.976***	-9.008***	-5.824***	2.814***
	(0.368)	(0.673)	(2.126)	(0.759)
F-stat on excluded instruments	27.4	1,700		
Montiel-Olea-Pflueger Eff. F-stat.	27.6	1,449		
Sargan-Hansen J-stat p-value			0.938	0.798
N	4,896	4,896	4,896	4,896
Adj. R^2	0.075	0.659		

Source: Authors' calculations.
Note: Table presents two-stage least squares first-difference regression estimates using a panel of district-state pairs in India between 2001 and 2011. District i is the local region and state j the distant. Excluding missing pairs, there are 4,896 $i - j$ district-state pairs consisting of 144 districts and 35 states and union territories of India in our sample. Differences are calculated as first differences of values between the years 2001 and 2011. Log housing units at i are instrumented by the decadal change in the number of months when rainfall was less than 80% of the long-term normal in state j and the GQ at j dummy, which equals one if state j was recipient of the Golden Quadrilateral highway upgrade program, and zero otherwise. Consumption, urban surface area, and urban surface area squared at i are the controls. Urban surface area unit in 1000 sq. miles. Robust standard errors presented in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table C.4: Formal elasticity estimation with full sample

	First-stage	Second-stage
	Dependent variable = ΔLog indicators	
	Formal housing units at i	Formal rent at i
$\Delta = 2011 - 2001$	(1)	(2)
ΔLog formal units at i		0.738***
		(0.029)
Δ#Months rainfall $<$ 80% at j	0.020***	
	(0.000)	
GQ highway at j dummy	0.214***	
	(0.004)	
ΔLog consumption at i	0.044***	0.175***
	(0.010)	(0.024)
ΔUrban surface area at i	9.423***	-6.879***
	(0.217)	(0.485)
ΔUrban surface area at i squared	-10.919***	8.023***
	(0.500)	(0.767)
Implied housing supply elasticity		1.36
F-stat on excluded instruments	3,697	
Montiel-Olea-Pflueger Eff. F-stat.	3,201	
Sargan-Hansen J-stat p-value		0.799
N	11,526	11,526
Adj. R^2	0.674	

Source: Authors' calculations.

Note: Table presents two-stage least squares first-difference regression estimates using a panel of district-state pairs in India between 2001 and 2011. District i is the local region and state j the distant. Excluding missing pairs, there are 11,526 $i - j$ district-state pairs consisting of the 35 states and union territories and 339 districts of India in our complete sample on formal rents. Differences are calculated as first differences of values between the years 2001 and 2011. Log housing units at i are instrumented by the decadal change in the number of months when rainfall was less than 80% of the long-term normal in state j and the GQ at j dummy, which equals one if state j was recipient of the Golden Quadrilateral highway upgrade program, and zero otherwise. Consumption, urban surface area, and urban surface area squared at i are the controls. Urban surface area unit in 1000 sq. miles. Robust standard errors presented in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.