



## The Review of Regional Studies

*The Official Journal of the Southern Regional Science Association*



# Does Titling Matter? Evidence from Housing Markets in India\*

Antonio M. Bento<sup>a</sup>, Somik V. Lall<sup>b</sup>, and Joel R. Landry<sup>c</sup>

<sup>a</sup>*Sol Price School of Public Policy and Department of Economics, University of Southern California and National Bureau of Economic Research, USA*

<sup>b</sup>*The World Bank*

<sup>c</sup>*John and Willie Leone Family Department of Energy and Mineral Engineering, The Pennsylvania State University, USA*

---

**Abstract:** In this paper, we exploit the plausibly exogenous random relocation of slum dwellers in Pune, India to evaluate whether the provision of a formal title is capitalized into property values. Using propensity score matching methods, which address important sources of bias that are present in conventional Ordinary Least Squares (OLS) hedonic methods, we find no effect of titling. Our results suggest that slum improvements may be sufficient to enhance informal tenure rights and that formalizing those rights with a title has little additional impact. An analysis of heterogeneous treatment effects from titling confirms this conclusion, providing preliminary evidence when titling may complement or substitute for other slum interventions.

**Keywords:** titling, formal property rights, policy interventions in slums

**JEL Codes:** D23, H31, I32, O17, R21

---

## 1. INTRODUCTION

Nearly a third of the world's urban population live in slums (UN, 2003). Located 125 miles east of Mumbai in India's Maharashtra state, Pune is representative of this global trend with nearly one million slum dwellers residing in a city of 3.1 million (ORGCC, 2011). The conditions that characterize slum life—informal home ownership and squatting, concentrated poverty, and inadequate access to housing, basic public services, formal credit, and job markets—have disproportionate impacts on the welfare of slum dwellers that can exacerbate their social exclusion. Several policy options have been considered to improve the welfare of slum dwellers, including: upgrading the housing quality of slums onsite, extending or expanding municipal services and infrastructure to slum communities, and providing households with a legal title to their dwelling (UN, 2003). This paper evaluates titling as a strategy for improving the welfare of the urban poor by estimating whether titling has been capitalized into housing values.

Titling is simply a formal, legal provision of a property right. Individuals who lack property rights may underinvest in their property if they are unable to exclude others from seizing the gains from their investments (Demsetz, 1967; Alchian and Demsetz, 1973). Previous literature suggests

---

\* Acknowledgements. We thank Carole Graham, the editor, and two anonymous reviewers for helpful comments and suggestions. All remaining errors are reserved by the authors.

Bento is a professor of public policy and economics at the University of Southern California and a research associate of the National Bureau of Economic Research. Lall is with the World Bank. Landry is an assistant professor of environmental and energy economics at The Pennsylvania State University, University Park. *Corresponding Author:* J.R. Landry E-mail: joelrlandry@psu.edu.

three primary mechanisms through which a formal title is likely to be capitalized into property values.<sup>1</sup> First, by providing households with secure tenure, households no longer face a threat of eviction or demolition, assuring households that investments into their housing stock will not be expropriated (Jimenez, 1984; Brasselle, Gaspart, and Platteau, 2002). Second, a title may allow households to offer their property as collateral, expanding their access to credit markets (Feder et al., 1988; Besley, 1995; Feder and Nishio, 1999; De Soto, 2000; Brasselle, Gaspar, and Platteau, 2002; Field and Torero, 2008). Finally, a title may lower the transaction costs realized when transferring houses between individuals (Besley, 1995; Brasselle, Gapsart, and Platteau, 2002; Lanjouw and Levy, 2002).

In this paper, we evaluate whether titling has been capitalized into housing values in Pune, India. We also examine how observables correlated with tenure security affect heterogeneity in the capitalization of titling into property values. Several other papers have examined these issues using different methods and in other urban contexts. Jimenez (1984), using an OLS hedonic, finds that the possession of a title increases the self-appraised estimated sale values of owner-occupied housing units by 58 percent and observed rents by 18 percent for renter occupied housing units in Davao, Philippines. Jimenez also shows that the benefit of possessing a title diminishes with a housing unit's age, providing initial evidence of *de facto* tenure tights. Performing a similar but extended analysis of titled and untitled homes in Manila, Friedman, Jimenez, and Mayo (1988) find that titling increases rental values by 11 percent and sale prices by 23 percent. In addition, they find that certain housing characteristics such as possessing a high quality roof also enhance tenure security and are more highly valued by housing units without a title. Unlike Jimenez (1984) and Friedman, Jimenez, and Mayo (1988) which estimate a hedonic OLS based upon revealed preferences, Lanjouw and Levy (2002) ask a random sample of titled (untitled) homeowners in Guayaquil, Ecuador what the sales price of their housing unit would be were they to forego (receive) a title. They also find that some household and community characteristics, such as gender of the head of household, affect perceived threats of eviction.

Estimates in the literature of the capitalization of titling into property values vary for several reasons. First, as noted by Lanjouw and Levy (2002) and many others, estimation of consumers' willingness to pay for a title is difficult because the assignment of titles across households is not typically exogenous. Households residing in housing units that possess a formal title will have selected into those properties because the expected net benefits from receiving a title may be higher. Titling interventions in which titling can be voluntarily obtained by households or where titling is purposively placed will cause similar selection issues. Because of the endogeneity of titling, a simple examination of the difference in property values between titled and untitled properties that are otherwise identical is likely to overstate the benefits of a titling program because of correlation of unobservables such as the threat of eviction with the receipt of a title. Lanjouw and Levy (2002), partially address the endogeneity issue by allowing households to self-report the value of the title for themselves, although whether a household possesses a title or not is voluntarily determined. Similar to Galiani and Schargrodsy (2010), we exploit a natural experiment where titling is plausibly exogenously assigned. In our sample, untitled households are randomly relocated in response to public infrastructure projects. Those households are then provided a new housing unit along with a title.

---

<sup>1</sup> In addition, a sizeable literature has examined the linkages between formal property rights and other dimensions of welfare that may not be capitalized into housing values; see Friedman, Jimenez, and Mayo (1988) and Galiani and Schargrodsy (2010) for reviews.

Second, there are considerable differences in the institutional, political, cultural, and economic contexts examined in the previous literature. Therefore, the baseline threat of eviction, access to credit markets, and social and political networks will differ across urban settings and, possibly, the estimated impacts of receiving a title. We provide the first estimate of the impact of titling for a major metropolitan city in India, Pune.

Third, and related to the second, are the ways in which housing unit and household characteristics are correlated with or 'produce' unobserved tenure security (i.e. the *ex post* threat of eviction facing households given their current choice of housing unit) in light of the baseline threat of eviction may differ across urban settings. This is important as the potential gains from titling interventions will be highly contingent on tenure security which cannot be directly measured or observed. Previous revealed preference studies assumed parametric functional forms to estimate the hedonic price function using OLS. Yet, the way in which titling interacts with unobserved tenure security, correlated housing unit, and household characteristics may be highly non-linear. Therefore, OLS estimates are likely to introduce specification bias through the imposition of a functional form (Rosen, 1974; Goodman, 1978; Cropper, Deck, and McConnell, 1988; Kuminoff, Parmeter, and Pope, 2010). Perhaps more importantly, the OLS estimator can suffer from common support failure bias when there is weak support between treatment and control groups on the basis of observed characteristics (Black and Smith, 2004; Ravallion, 2008).

Using survey data on slums from the World Bank's Pune Household Survey in which titling is plausibly exogenously assigned following random resettlement, when using an ordinary least squares (OLS) hedonic estimator we find a statistically significant effect from titling on owner occupied housing values of 20.2 percent of the estimated monthly rental value. This is near the lower bound estimate previously identified in the literature for owner occupied housing units. However, using a Propensity Score Matching (PSM) estimator, we find a much smaller effect of titling of 13.2 percent which is no longer statistically significant. Common support bias is the main driver of the difference between the two estimates. This is because many treatment households that receive a title also receive upgrades to their housing unit and/or improved access to municipal services, which tend to accompany resettlement; whereas only some control households also receive these improvements. Unobserved tenure security is correlated with some observable housing unit and household characteristics. For example, an untitled dwelling unit with brick walls is less likely to be seized than one without brick walls. Thus, our results suggest that slum improvements may already enhance informal rights to such an extent that formalizing those rights with a title provides little additional benefit as measured through the capitalization into property values. A careful review of treatment effect heterogeneity on the treated confirms this insight. We find that treatment effects for households that receive a title may be higher for households that have less tenure security as measured across several dimensions, such as possessing a dwelling unit without brick walls or residing in dense neighborhoods. Previous studies may overstate the benefits of titling if slum improvements jointly accompany titling if weak support over observables is correlated with tenure security, or otherwise impacts property values as well when titling is not exogenously assigned.

Our results broadly coincide with the mixed evidence regarding whether the provision of formal property rights improves the welfare of the poor. Slum dwellers in Pune have limited access to credit markets and possess strong informal tenure rights and so the provision of a formal title has little additional impact on welfare as capitalized into property values. Consequently, from the perspective of the capitalization of titling into housing values, interventions such as slum

upgrading and improvements in access to municipal services which improve *de facto* tenure security may be preferable to titling schemes.<sup>2</sup> Our analysis of treatment effect heterogeneity provides suggestive evidence as to when titling programs are likely to be complementary with or substitutable for other slum intervention strategies; and thus, how future titling programs may be better designed to achieve welfare gains.

The rest of the paper is organized as follows. Section 2 introduces a simple conceptual model that illustrates the relationship between the baseline threat of eviction, the *ex post* threat of eviction given housing unit and household characteristics, the marginal willingness to pay for a title, and the implications for the failure of the common support when evaluating titling. Section 3 discusses the data and reviews summary statistics. Section 4 presents the conventional hedonic model and estimates. Section 5 details our propensity score matching methods and estimates and reports the results from our heterogeneous treatment effects analysis. Finally, Section 6 concludes.

## 2. CONCEPTUAL MODEL

In this section we construct a simple conceptual model that illustrates the relationship between unobserved tenure security, titling, and the distributions of household and housing characteristics. This model is an extension of the theoretical model originally developed by Jimenez (1984). Consider a household  $i$  that receives exogenous income,  $m_i$ . A housing unit  $j$  can be represented as a bundle of housing and neighborhood characteristics,  $\mathbf{q}_j = [\mathbf{x}_j, \mathbf{z}_j, t_j]$ , where  $t_j = 1$  when the housing unit has a formal title and  $t_j = 0$ , otherwise,  $\mathbf{x}_j$  are a vector of housing and neighborhood characteristics that are orthogonal to the likelihood of being evicted<sup>3</sup> (e.g. the color of the flooring), and  $\mathbf{z}_j$  are a vector of housing and neighborhood characteristics that may alter the likelihood of eviction (e.g. durability of the walls, safety of the neighborhood, having a toilet indoors, or access to piped water).

The household has utility given by  $u(n_i, \mathbf{x}_i, \mathbf{z}_i, \mathbf{w}_i)$ , where  $n_i$  is the amount of a composite good consumed by the household whose price is normalized to one and  $\mathbf{w}_i$  is a vector of household characteristics. For simplicity, we do not consider a direct benefit to receiving a title other than through the elimination of the threat of eviction. We assume that the rental price function is separable with respect to a title such that  $r_j = r(\mathbf{x}_j, \mathbf{z}_j) + p(\mathbf{x}_j, \mathbf{z}_j)t_j$ .  $p_j = p(\mathbf{x}_j, \mathbf{z}_j)$  reflects the marginal change in rental value from possessing a title for a housing unit with characteristics  $\mathbf{x}_j$  and  $\mathbf{z}_j$ . In the context of the housing market equilibrium, this also equals the household's marginal willingness to pay for a title. To keep the exposition simple, we drop household and housing unit subscripts except when necessary in the analysis that follows.

A household who buys a titled housing unit *ex ante* maximizes  $u(n^S, \mathbf{x}^S, \mathbf{z}^S, \mathbf{w})$  subject to budget constraint  $m = n^S + r(\mathbf{x}^S, \mathbf{z}^S) + p(\mathbf{x}^S, \mathbf{z}^S)$ . The solution provides the *ex ante* indirect utility function when the household buys a titled house,  $V^S(r^S + p^S, \mathbf{w})$ ,<sup>4</sup> where  $r^S =$

<sup>2</sup> Given data limitations, we are not able to directly evaluate which of these interventions would be preferred. Moreover, there are other welfare benefits from titling that are not fully capitalized into property values as well as differences in efficiency costs that may make titling a preferred intervention strategy.

<sup>3</sup> Alternatively, the housing unit may be demolished or expropriated. Moreover, in some contexts housing units may face less implicit or explicit access to municipal or other governmental services. For example, an untitled property may be less likely to file a police report when their untitled home is robbed. Further, a household whose home is destroyed due to a flood event may not be able to obtain relocation or relief assistance unless they possess a title to their property. The subsequent analysis would cover threats such as these albeit with slight modifications to the model below.

<sup>4</sup> We suppress income in what follows since it can be subsumed in  $\mathbf{w}$ .

$$r\left(\mathbf{x}^S(r^S + p^S, \mathbf{w}), \mathbf{z}^S(r^S + p^S, \mathbf{w})\right), \text{ and } p^S = p\left(\mathbf{x}^S(r^S + p^S, \mathbf{w}), \mathbf{z}^S(r^S + p^S, \mathbf{w})\right).$$

A household who buys an untitled housing unit *ex ante* may be evicted with probability  $\pi(Z, W, \pi_0)$ , where  $\pi_0$  is the initial baseline threat of eviction.<sup>5</sup> Neither  $\pi(\cdot)$  nor  $\pi_0$  can be directly observed. The principal difference in the model we have laid out and Jimenez (1984) is that we allow for untitled households to strategically make investments in their untitled housing units so as to affect their realized threat of eviction. In addition, we assume that the eviction risk facing a household before they make their housing decision is a function of household characteristics,  $W$ , and the baseline exogenous threat of eviction,  $\pi_0$ . Certain household characteristics such as the gender of the head of household or their caste status, for example, may place those households at greater risk of being evicted. The baseline threat of eviction is assumed to be constant within a single urban housing market, although this may differ across municipalities reflecting differences in political, economic, social, cultural, and institutional environments.<sup>6</sup> If the owner of the untitled house is not evicted then the household realizes budget constraint  $m = n^U + r(\mathbf{x}^U, \mathbf{z}^U)$ . If the owner of the untitled house is evicted then the household realizes budget constraint  $m = n^E + r(\mathbf{x}^U, \mathbf{z}^U) + \phi + r(\mathbf{x}^E, \mathbf{z}^E) + p(\mathbf{x}^E, \mathbf{z}^E)$ , where  $r(\mathbf{x}^U, \mathbf{z}^U)$  is the value of the untitled house that is vacated,  $\phi > 0$  is a lump-sum cost of eviction, and  $r(\mathbf{x}^E, \mathbf{z}^E) + p(\mathbf{x}^E, \mathbf{z}^E)$  is the cost of obtaining a tenure secure housing unit (with title) once evicted.

When the household buys an untitled housing unit *ex ante*, they maximize  $\pi(\mathbf{z}^U, \mathbf{w}, \pi_0)u(n^E, \mathbf{x}^E, \mathbf{z}^E, \mathbf{w}) + (1 - \pi(\mathbf{z}^U, \mathbf{w}, \pi_0))u(n^U, \mathbf{x}^U, \mathbf{z}^U, \mathbf{w})$  subject to the two budget constraints conditional on the state of the world that is realized. This provides the following first-order conditions:

$$\begin{aligned} \mathbf{x}^E: u_{x^E} &= u_{N^E}(r_{x^E} + p_{x^E}), \\ \mathbf{x}^U: (1 - \pi)u_{x^U} &= r_{x^U}[\pi u_{N^E} + (1 - \pi)u_{N^U}], \\ \mathbf{z}^E: u_{z^E} &= u_{N^E}(r_{z^E} + p_{z^E}), \\ (1) \quad \mathbf{z}^U: (1 - \pi)u_{z^U} - r_{z^U}[\pi u_{N^E} + (1 - \pi)u_{N^U}] &= \pi_{z^U}(u^U - u^E), \end{aligned}$$

where  $\pi = \pi(\mathbf{z}^U, \mathbf{w}, \pi_0)$ ,  $u^U = u(n^U, \mathbf{x}^U, \mathbf{z}^U, \mathbf{w})$ , and  $u^E = u(n^E, \mathbf{x}^E, \mathbf{z}^E, \mathbf{w})$ .

The first two equations in (1) are analogous to the model developed by Jimenez (1984) in which eviction risk is exogenous to the household and cannot be altered by household investment. The first through third lines state that when the household is evicted they choose housing unit characteristics such that expected marginal benefits equal expected marginal costs.

The fourth line states that the household selects  $\mathbf{z}^U$  such that net expected benefits equal the marginal change in eviction risk from a marginal change in  $\mathbf{z}^U$ ,  $\pi_{z^U}$ , multiplied by the difference in realized utilities in the non-evicted and evicted states of the world,  $(u(U) - u(E))$ . Since  $u(U) > u(E)$ , when  $\pi_{z^U} > 0$  the right-hand side reflects an implicit tax and when  $\pi_{z^U} < 0$ , an implicit subsidy. Intuitively, when consuming more of a housing attribute that raises the risk of eviction, the household marginally selects less of it. Conversely, when a housing attribute lowers the threat of eviction, the household consumes more of it. Depending upon how housing attributes

<sup>5</sup> As noted by Besley (1995), the realization of eviction is analogous to being forced to pay a tax at random.

<sup>6</sup> Since  $\mathbf{z}$  can include a measure of the spatial location of the housing unit,  $\pi(\cdot)$  can also capture heterogeneity in baseline eviction risk *within* a municipality. This is also considered by Jimenez (1984), although our treatment is more general.

are combined in  $r(\cdot)$  and  $p(\cdot)$ , households can partially adapt to the *ex ante* eviction risk they face by making strategic investments in their housing unit.

The first-order conditions provide the *ex ante* indirect utility of the household when they purchase a non-titled housing unit:  $V^U(r^E + p^E, r^U, \phi, \pi_0, \mathbf{w})$ , where  $r^E = r(\mathbf{x}^E(r^E + p^E, r^U, \phi, \pi_0, \mathbf{w}), \mathbf{z}^E(r^E + p^E, r^U, \phi, \pi_0, \mathbf{w}))$ ,  $p^E = p(\mathbf{x}^E(r^E + p^E, r^U, \phi, \pi_0, \mathbf{w}), \mathbf{z}^E(r^E + p^E, r^U, \phi, \pi_0, \mathbf{w}))$ , and  $r^U = r(\mathbf{x}^U(r^E + p^E, r^U, \phi, \pi_0, \mathbf{w}), \mathbf{z}^U(r^E + p^E, r^U, \phi, \pi_0, \mathbf{w}))$ . *Ex ante*, a household decides to purchase a titled home when  $V^S(r^S + p^S, \mathbf{w}) > V^U(r^E + p^E, r^U, \phi, \pi_0, \mathbf{w})$  and purchases an untitled home otherwise. Therefore, the adaptive investments implied by the last equation of (1) are contingent upon whether the household decides to obtain a title *ex ante*, which depends upon the relative rental rates of alternative housing units, the baseline rate of eviction, household attributes, and household income. Those housing attributes which lower eviction risk are a substitute for obtaining a title and those housing attributes which raise eviction risk are complements to obtaining a title.

Following the discussion in Jimenez (1984) and the hedonic literature (Rosen, 1974), an equilibrium in the housing market requires:  $V_i^S(r_{ij}^S + p_{ij}^S, \mathbf{w}_i) = V_i^U(r_{ij}^E + p_{ij}^E, r_{ij}^U, \phi, \pi_0, \mathbf{w}_i)$  for all  $i$  households occupying housing unit  $j$ . The resulting equilibrium vector of rental values reflects the outer envelope of all households' utility maximizing indifference curves in the characteristics-price space, such that every household occupies a single housing unit.

The above analysis clearly illustrates how the baseline threat of eviction interacts with exogenous household characteristics to effectuate the distribution of housing unit characteristics across housing units in equilibrium. Obtaining a title is but one means of adapting to a household's eviction risk which depends upon the exogenous baseline threat of eviction and household characteristics. Alternately, households may make strategic improvements in their housing stock. Simply put, not only is the distribution of titled housing units endogenously determined in equilibrium as a consequence of the baseline threat of eviction and the distribution of household characteristics, but so too is the distribution of most housing unit characteristics.<sup>7</sup> Depending upon the marginal rental prices of various housing attributes (e.g. the cost of bricks or sheet roofing), the ways in which heterogeneous households respond to the baseline threat of eviction will yield different housing mixes in terms of both the attribute and title space. Therefore, there is no *a priori* reason to suspect that the distribution of household and housing unit attributes will be balanced across those households who possess a titled property and those who do not. For a subset of households occupying untitled housing units that are not evicted *ex post*,  $p(\mathbf{x}^U(\cdot), \mathbf{z}^U(\cdot))$  is unobserved. Thus, the parameters within  $p(\cdot)$  can be empirically estimated only for those housing units with observed characteristics  $\mathbf{x}^S(\cdot), \mathbf{z}^S(\cdot)$  or  $\mathbf{x}^E(\cdot), \mathbf{z}^E(\cdot)$  and these characteristics may have limited support relative to  $\mathbf{x}^U(\cdot), \mathbf{z}^U(\cdot)$ . This suggests that hedonic analyses in which titling is voluntarily assigned such as Jimenez (1984) and Friedman, Jimenez, and Mayo (1988) likely also suffer from specification and common support failure biases.

Moreover, since many housing attributes may complement or substitute for the receipt of a title and do so in highly non-linear ways, this will affect the ability to evaluate the benefits of

<sup>7</sup> Although we assume only a subset of housing unit characteristics directly affect the realized threat of eviction facing a household, depending upon the marginal rates of transformation between housing characteristics in  $\mathbf{x}$  and  $\mathbf{z}$  in  $r(\cdot)$  and  $p(\cdot)$ , so too will the distribution of characteristics in  $\mathbf{x}$  be affected in the realized housing market equilibrium.

receiving a title, even when the assignment of a title is plausibly exogenous. In the context we examine, the random receipt of a title also accompanies the provision of a new housing unit. If the attributes of these new housing units have limited overlap with the attributes of untitled units, we will likely observe a discrepancy between hedonic estimators that address common support failure bias such as PSM and estimators that do not such as OLS.

### 3. DATA AND SUMMARY STATISTICS

We focus on slum dwellers in Pune, India. Pune is the second largest city in Maharashtra, a state which has the highest per capita income and lowest levels of income disparity in India (ORGCC, 2011). In recent years, it has experienced rapid growth in its industrial, education, and information-technology sectors; this has been followed by exceptionally rapid population growth of 23.1 percent in 2011 over the previous decade (ORGCC, 2011). Consequently, the proportion of slums in Pune (32.5 percent) is slightly larger than the world average of urban slum dwellers (31.6 percent) (UN, 2003; Times, 2011). Slum or squatter settlements typically develop by encroaching on open lands, both public and private, including land prone to natural hazards. Residents typically live in small and poorly designed dwelling units with limited access to public services, especially water and adequate sanitation, and lack a title to the land occupied by their dwelling unit as well as the dwelling unit itself. Here we describe the distribution of the population as well as characteristics of the housing stock for a subsample of 767 households from the World Bank's Pune Household Survey (PHS) collected between August and October of 2002. The PHS is the primary dataset used in our analysis, although we geospatially link households from the PHS to several other geographic datasets from the World Bank.<sup>8</sup>

The PHS contains 2,849 randomly sampled households selected to be representative of the 430 km<sup>2</sup> metropolitan area administered by the Pune Municipal Corporation (PMC).<sup>9</sup> All households in the city are part of this sampling universe with the exception of institutional populations (e.g. prisons) or residents of military installations. Of the original sample, we focus on 1,071 informal settlements, of which 943 households are squatter settlements or slums and 128 are resettlements. Squatter settlements where the government has agreed to provide basic services are termed as "notified" slums, and are thus considerably less likely to be evicted. Notification typically precedes attempts to provide on-site slum upgrading and improvements in access to municipal services (i.e. slum improvements).<sup>10</sup> Eight hundred forty-six households or roughly nine-tenths of all squatter settlements surveyed are of this type. As slum dwellers frequently encroach upon areas that are already considerably congested, slum dwellers may be involuntarily resettled in response to large-scale municipal infrastructure programs, such as road, bridge, or railway expansion initiatives. Slum dwellers that have lived at their address for over five years and prior to 1995 qualify for resettlement, according to Maharashtra state regulations (Cronin and Guthrie, 2011).<sup>11</sup> Resettled slum dwellers are involuntarily relocated to a new location, in many instances to dwelling units of improved quality and design with better access to municipal services,

---

<sup>8</sup> For complete details on variable and sample construction, additional tests, and alternative specifications, see the Online Appendix available at: <http://www.joelrlandry.com>.

<sup>9</sup> The PMC is the governing authority in Pune. Pune is divided into 48 administrative wards each of which directly elect three members to the PMC. Principal executive authority in the PMC resides with the Municipal Commissioner which is appointed by the Maharashtra state government.

<sup>10</sup> We do not observe which slums receive improvements from the government, although we do observe housing and neighborhood characteristics which reflect these improvements and are therefore included as controls.

<sup>11</sup> According to the Pune Slum Census conducted by Sen and Hobson (2005), 82 percent of slum dwellers in Pune meet this criterion.

and are given a formal title to their new home.<sup>12</sup> Of the 846 notified slums, 723 have been living in their current location for ten years or longer. Likewise, for 128 resettlements this includes 107 housing units. The majority of these units own their own home, or 666 notified slums and 101 resettlements, respectively, leaving a final sample of 767 housing units. Given notification, it is not wholly surprising that the proportion of notified squatters that have been living in their current location for ten years or longer is quite large (85 percent), and slightly greater than the same for resettlement households (84 percent). This suggests a high degree of *de facto* or informal tenure security for both notified squatters and resettlement households in our sample.

### 3.1 Experimental Design

We exploit a quasi-experimental research design where notified slums form the “untreated” group and resettlements form the “treated” group. As we discuss below, treatment status is independent of observed household characteristics, and hence, plausibly exogenous. Our approach is quasi-experimental as we lack complete pre-treatment data on both the control and treatment groups. Our estimated treatment effect thus reflects the impact of titling to owners conditional on notification status. One concern to limiting the sample in this way is that our estimated treatment effects may not be valid generally if the remaining sample is not representative of all slum households. For instance, one might suspect that notified squatters may have been notified as a result of having better social networks or political connections, or that resettlements which have resided in their current residence for ten years or more remain to be surveyed because of attributes which make them less likely to move or possess greater tenure security. To determine whether these or related issues<sup>13</sup> are a concern for our analysis, we perform several balancing tests across a rich array of characteristics between the sub-samples included in our sample and those excluded from our final sample, the full results of which are reported in the Online Appendix. We find no meaningful statistically significant difference in these variables, either between notified and non-notified slums, between notified slums with occupancy of ten years or more, and notified slums with occupancy of less than ten years, or between resettlements with occupancy of ten years or more and resettlements with occupancy of less than ten years. This suggests that the slum dwellers remaining for evaluation are generally representative of the full population of slum dwellers, and thus our estimate of the impact of titling is likely valid more generally.

### 3.2 Summary Statistics

The first three columns of Table 1 present summary statistics for the full sample, the treatment (resettlements) group, and the control group, respectively. The final column reports the normalized differences between the treatment and control groups (Imbens and Wooldridge, 2009), and asterisks report the results of *t*-tests of the null hypothesis that there is no difference in means between the treatment and control groups. While there is very little difference in household characteristics between resettled households and notified squatters, there is considerable disparity

---

<sup>12</sup> There are likely drawbacks to resettlement as residents’ original residential location choice reflected a careful arbitrage between the amenities provided by the physical housing stock and other locational amenities, such as access to public schools and employment (Bento, Cropper, and Tacheuchi, 2008; Lall and Lundberg, 2008). We account for how these locational amenities may be capitalized into housing values by including a rich set of geographic controls.

<sup>13</sup> A related concern is that we only observe households at a single snapshot of time. Ideally, we would like to observe treatment and control households before and after resettlement. Those resettled households that remain to be evaluated after ten years may differ from those households that are not present in the sample in ways that may be correlated with the possible benefits of receiving a title. By evaluating whether there are statistically meaningful differences in observed, treatment-invariant household characteristics of resettled households of varying tenure lengths, we evaluate indirectly whether such selection is a concern for our analysis. See Online Appendix for additional discussion.

**Table 1: Summary Statistics**

Variable Name	Full Sample ( <i>n</i> =767) Mean	Resettle- ments ( <i>n</i> =101) Mean	Notified Squatters ( <i>n</i> =666) Mean	Normal. Diff.
<i>Housing Unit Characteristics:</i>				
Size of the housing plot (sq. feet)	370.65	678.05	324.19	0.30***
Total covered area of housing unit (sq. feet)	356.54	596.50	320.07	0.26***
Area occupied by the HH for living space (sq. feet)	262.60	370.24	246.27	0.35***
Num. of rooms in the HU (excl. bathrooms and kitchens)	2.04	3.35	1.85	0.31***
Year of construction	1985	1983	1985	-0.13
Estimated monthly rental value for a similar unit in this neighborhood (rupees)	887.29	1550.00	798.30	0.36***
% of HU that have poor quality floor	1.96	0.00	2.25	-0.15
% of HU that have high quality floor	68.19	83.17	65.92	0.27***
% of HU that have poor quality walls	0.26	0.00	0.30	-0.05
% of HU that have high quality walls	77.44	87.13	75.98	0.20
% of HU that have poor quality roof	2.22	0.99	2.40	-0.08
% of HU that have high quality roof	14.99	31.68	12.46	0.32***
% of HH that has a kitchen within the HU	99.22	100.00	99.10	0.09
% of HH that has a kitchen within the HU, separate room	28.55	44.55	26.13	0.27***
% of HH that has a toilet within the HU	17.86	59.41	11.56	0.63***
<i>Geographic Characteristics:</i>				
% of HU that fall within 50 meters of a river	2.22	0.99	2.40	-0.08
% of HU that fall within 500 meters of a primary school	63.36	88.11	59.61	0.44***
Number of slums within 200 meters of HU	9.31	3.13	10.24	-0.68***
Number of HU within 500 meters of HU	37.61	34.37	38.10	-0.13
Distance of HU to central business district (meters)	4151.59	4171.60	4148.56	-0.01
<i>Neighborhood Characteristics:</i>				
Hours per week that water is available	76.71	57.21	79.67	-0.31***
% of HU that have legal piped water	74.05	94.06	71.02	0.41***
% of HU with toilet discharge into a sewer system	76.40	86.14	77.92	0.20
% of HU with access to electricity via a metered connection	85.27	95.05	83.78	0.25***
% of HH who report roads that are in need of maintenance	94.65	91.09	95.20	-0.11*
% of HH that report that someone in their household has been a victim of a minor crime in the last three years	3.39	0.99	3.75	-0.13
% of HH that report that neighborhood is safe for women to walk around in after dark	76.66	93.07	74.17	0.35***

Notes: HH denotes households, whereas HU denotes housing units. Slums include both notified and non-notified settlements. Fourth column reports normalized differences (Imbens and Wooldridge, 2009). \*, \*\*, and \*\*\* denotes statistical significance at the 10 percent, 5 percent, and 1 percent levels, respectively, for unpaired t-tests that evaluate the null hypothesis that there is no difference in means between the treatment and control groups.

in housing characteristics.

For example, resettled households have an additional 1.3 rooms to occupy and an additional

123.9 square feet of living space compared to notified squatters. Resettlements are also more likely to have good walls, floor, and a roof than notified slums as well as a kitchen in a separate room and a toilet in the house. Disparities between resettlement and notified slums in terms of geographic and neighborhood characteristics are also large. Resettlements are slightly less densely concentrated and have better access to primary schools than notified slums. Moreover, they are two-thirds less likely to be concentrated next to other slums or to be located near rivers which are known to flood.

Although resettlements have better access to sewage, electricity, and water, the hours per week that water is available is considerably less than it is for notified slums. The latter is likely due to the high geographic variability in water supply in Pune, suggesting that resettlements may be located in higher elevations than notified slums (Bapat and Agarwal, 2003). Other indicators of municipal service quality, such as perceptions of road maintenance and crime levels, suggest that resettled households have more positive perceptions than notified squatters.

Part of the differences between resettlements and notified slums extend from the fact that resettlement is accompanied by slum improvements, whereas notified slums are housing units that have received various slum improvements as well as those that have received none. Normalized differences are in many cases large and we reject the null hypothesis of no difference in means between treatment and control groups for many of the variables discussed above. Consequently, these differences portend what is likely to be shallow common support between resettlements and notified slums which will have important implications for identifying the impact of titling on property values.

#### 4. OLS HEDONIC

Hedonic estimation provides a natural framework for assessing the impact of a title on housing values, as it relies on the fact that houses are composite goods whose value is a combination of the values of the individual attributes that compose the house, as well as the amenities and services available in the area (Lancaster, 1966; Rosen, 1974). Since the effects of titling should be capitalized into property values, we can capture this effect when estimating the hedonic equation by including a dummy for titling alongside other housing and neighborhood attributes. The model we estimate is given by:

$$(2) \quad y_{iw} = \alpha + \beta t_{iw} + \gamma \mathbf{x}_{iw} + \delta_w + u_{iw}$$

where  $y_{iw}$  is the natural log of the estimated monthly rental value of housing unit  $i$  in administrative ward  $w$ ,<sup>14</sup>  $t_{iw}$  is a dummy equal to one if a housing unit possesses a title,  $\mathbf{x}_{iw}$  is a vector of covariates,  $\delta_w$  are administrative ward fixed effects,  $u_{iw}$  is the error term, and  $\alpha$ ,  $\beta$ , and  $\gamma$  are the coefficients to be estimated.  $\delta_w$  captures unobserved differences between administrative wards that are invariant across housing units residing in the same ward.

Included in  $\mathbf{x}_{iw}$  are many of the housing unit, geographic, and neighborhood attributes whose summary statistics are reported in Table 1, including: the natural log of the area used for living space, the number of notified and non-notified slums within 200 meters of a housing unit, the number of total housing units within 500 meters of a housing unit, the distance to the central business district, the natural log of the number of hours per week for which water is available, and

<sup>14</sup> We report results for a semi-log specification because it is easier to interpret model parameters, and because the results are very close to those from an optimal Box-Cox specification, which are provided in the Online Appendix.

dummies for whether a housing unit has a kitchen in a separate room, has a toilet inside the unit, has high quality floors (mosaic/floor tiles), high quality walls (burnt brick walls), has a low quality roof (corrugated metal or asbestos sheeting), has legal piped water, has toilet discharge into a sewer system, has access to electricity via a metered connection, lies within 500 meters of a river or other open water source, lies within 500 meters of a primary school, reports roads that are in need of maintenance, reports that they know someone in their household who was a victim of a minor crime in the last three years, and reports that the neighborhood is safe for women to walk around in after dark. The rich set of neighborhood and geographical variables that we include allow us to control for how resettlement may reflect different locational amenities and attributes as households are moved across the city that are independent of the receipt of a title.

Our preferred dependent variable is the natural log of each household's self-reported estimate of the monthly rental value for "a similar unit in the neighborhood",<sup>15</sup> which we interpret as the household's marginal willingness-to-pay for the vector of dwelling unit and neighborhood characteristics that are included on the right-hand side of Equation (2). Thus, the coefficient on  $t_{iw}$  can be interpreted as a household's willingness-to-pay for a title, which should be capitalized into the rental rate.<sup>16</sup>

Table 2 presents OLS estimates for the full sample. The first model in column one is the base model, which is the outcome variable of interest regressed solely against the titling dummy. Column two includes the titling dummy as well as all housing unit characteristics. Column three includes the titling dummy and the housing unit and geographic covariates. Finally, column four includes the titling dummy and all housing unit, geographic, and neighborhood characteristics. Robust standard errors clustered at the ward level are reported for each model. As shown in the Online Appendix, multicollinearity and spatial autocorrelation are not a concern for our models.

Under our OLS hedonic specification, a household's willingness-to-pay for a title is considerable. As additional covariates are included, the coefficient on titling remains statistically significant, with the presence of a title corresponding to a 20.2 percent increase in the estimated monthly rental value of the housing unit in the full model. This is considerably smaller than the 58 percent effect that Jimenez (1984) finds, but closer to the 23 percent effect found by Friedman, Jimenez, and Mayo, (1988) and the 23.5 percent effect identified by Lanjouw and Levy (2002). That our estimate is lower than these previous estimates may be explained by our ability to deal with the endogenous selection of titling<sup>17</sup> which these other papers either do not address (Jimenez, 1984; and Friedman, Jimenez, and Mayo, 1988), or only address partially (Lanjouw and Levy, 2002).

<sup>15</sup> The PHS also provides a self-reported estimate of the current market price for a similar unit in the neighborhood. Since both measures are self-reported, we do not have especially strong priors for either measure. Other authors have used the monthly rental price (see e.g., Lall and Lundberg, 2008), and so we report results using that variable for consistency. As shown in the Online Appendix, we find no statistically significant difference in the mean rent to price ratio between titled and untitled housing units included in our sample and thus do not believe that the choice of dependent variable is consequential for the analysis that follows. For completeness, we also provide complete results for all models using the estimated current market price in the Online Appendix. While the estimated coefficients tend to be larger in magnitude, the exposition and conclusions are similar to that which follows.

<sup>16</sup> More precisely, this capitalization reflects the present value of the discounted stream of future net benefits that are expected to accrue to the housing unit owner as a result of having a title.

<sup>17</sup> A remaining concern one may have in this regard is that successful or avoided resettlement depends upon other un-observables, such as political power. In the Online Appendix we perform a comprehensive evaluation of our claim that titling assignment is plausibly exogenous. We find no compelling evidence of endogeneity of titling. This is unsurprising given that our analysis is restricted to slum dwellers who are very poor and who likely possess limited access to the political process in general. Consequently, resettlement in response to large-scale public works projects is likely to be dictated by the siting needs of the public works project, or, in the case of resettlement due to flooding concerns, by proximity to vulnerable waterways.

**Table 2: OLS Hedonic Estimates**

	(1)	(2)	(3)	(4)
	ln(Rent)	ln(Rent)	ln(Rent)	ln(Rent)
Legal title	0.631 (0.115)***	0.233 (0.083)***	0.211 (0.089)**	0.202 (0.091)**
ln(Area occupied by the HH for living space)		0.378 (0.044)***	0.378 (0.044)***	0.366 (0.044)***
HU has a kitchen within the unit, in a separate room		0.081 (0.05)	0.071 (0.052)	0.07 (0.051)
HU has a toilet within the unit		0.285 (0.072)***	0.276 (0.074)***	0.255 (0.077)***
HU has high quality floor		0.119 (0.059)*	0.125 (0.059)**	0.13 (0.056)**
HU has high quality walls		0.147 (0.058)**	0.152 (0.059)**	0.112 (0.060)*
HU has average quality roof		-0.209 (0.043)***	-0.21 (0.045)***	-0.213 (0.045)***
HU has legal piped water		0.039 (0.047)	0.029 (0.048)	0.01 (0.049)
HU falls within 50m of a river			0.004 (0.119)	0.013 (0.12)
HU falls within 500m of a primary school			-0.092 (0.053)*	-0.108 (0.054)*
Number of slums within 200m of HU			-0.008 (0.004)**	-0.007 (0.004)*
Distance of HU to central business district (1,000m)			0.003 (0.019)	0.011 (0.02)
Number of HU within 500m of HU			0.001 (0.002)	0.001 (0.002)
ln(hours per week that water is available)				0.026 (0.026)
HH reports that roads are in need of maintenance				-0.26 (0.103)**
HH reports household has been victim of a minor crime				-0.061 (0.107)
HH reports that neighborhood is safe for women				0.088 (0.058)
HU with toilet discharge into a sewer system				0.037 (0.046)
HU with access to electricity via a metered connection				0.169 (0.057)***
Ward fixed effects?	Yes	Yes	Yes	Yes
Observations	767	767	767	767
R2	0.19	0.45	0.46	0.47

Notes: HH denotes households, whereas HU denotes housing units. Slums include both notified and non-notified settlements. \*, \*\*, and \*\*\* denotes statistical significance at the 10%, 5%, and 1% levels, respectively. Standard errors clustered by ward are reported in parentheses.

Most of the signs on the remaining covariates are of the expected direction. Of those coefficients whose signs are opposite of what we would expect, only primary schools within 500 meters of the housing unit is statistically significant at the 10 percent level. One possible explanation is that disamenities located near schools exceed the positive benefit of living next to a school. Seeing that the households in our sample are typically older households given their ten years or greater tenure status, the benefit of living next to a primary school is likely very small.

## 5. PROPENSITY SCORE MATCHING

There are two major shortcomings in relying on the conventional OLS hedonic to provide unbiased estimates of the impact of titling: misspecification of the functional form (specification bias) and the reliance of observations that are not on the common support (common support failure bias). Propensity score matching (PSM) methods first proposed by Rosenbaum and Rubin (1983) can generate more accurate estimates of the impacts of titling than conventional hedonic methods by dealing with these issues explicitly.

### 5.1 Propensity Score Matching Model

The coefficient of interest that we seek to identify using PSM is the average treatment effect on the treated or  $\Delta^{TT}$ . Here  $\Delta^{TT}$  corresponds to the average willingness-to-pay for a title *for those households that received a title* which is given by:

$$(3) \quad \Delta^{TT} = E(y_{iw}^1 - y_{iw}^0 | t_{iw} = 1),$$

where  $y_{iw}^1$  is the value of a housing unit with a title,  $y_{iw}^0$  is the value of a housing unit without a title, and  $t_{iw} = 1$  indicates a housing unit that received a title. Given our quasi-experimental setting, if a housing unit receives a title we only observe  $E(y_{iw}^1 | t_{iw} = 1)$  and not  $E(y_{iw}^0 | t_{iw} = 1)$ . PSM estimators provide a proxy for  $E(y_{iw}^0 | t_{iw} = 1)$ ,  $E(y_{iw}^0 | \hat{p}(\mathbf{x}_{iw}))$ , by pairing each treated observation with one or more observationally equivalent non-treated observations on the basis of the estimated propensity score,  $\hat{p}(\mathbf{x}_{iw})$ . The proxy we utilize uses all control observations weighted using the Epanechnikov kernel (Frölich, 2004),  $\omega_j(\cdot)$ :  $E(y_{iw}^0 | \hat{p}(\mathbf{x}_{iw})) = \sum_{j=1}^J \omega_j(\hat{p}(\mathbf{x}_{iw}), \hat{p}(\mathbf{x}_{jw})) y_{jw}^0$ , where  $\hat{p}(\mathbf{x}_{iw})$  is the estimated propensity score of resettlement  $i$ ,  $\hat{p}(\mathbf{x}_{jw})$  is the estimated propensity score of notified slum  $j$ , and  $y_{jw}^0$  is the observed value of notified slum  $j$ .<sup>18</sup> We estimate  $\hat{p}(\mathbf{x}_{iw})$  using a probit model that includes all of the covariates included in column four of the OLS specification. In addition, following common practice in the literature, we also include squares and cubits of all the continuous covariates in  $\mathbf{x}_{iw}$  to add flexibility. Table A8 in the Online Appendix presents the output for this model. The PSM estimator addresses specification bias through this flexible semi-parametric probit and the kernel weights.

The PSM estimator handles common support bias in two ways. First, treatment observations that fall off of the common support are excluded completely. Second, the kernel estimator assigns weights for each untreated observation inversely to the strength of the common support. The smaller the bandwidth, generally the better it deals with these issues, although this will often lead to fewer matches. The key benefit of the PSM estimator relative to the OLS estimator is that it avoids parametrically extrapolating a treatment effect when the common support is weak. Through the estimated propensity score, the PSM estimator compares the housing values of resettlements that have received slum improvements with notified slums that have received those improvements as well. Likewise, it compares resettlements that did not receive slum improvements with notified squatters of the same. Since the propensity score transforms heterogeneity across multiple covariates into a single dimensional measure, the per-treatment unit

<sup>18</sup> The matching method we use follows the standard procedure described in detail in Heckman and Robb (1986), Heckman, Ichimura, and Todd, (1997), and Heckman et al. (1998). For reviews of PSM methods see Imbens and Rubin (2015), Guo and Fraser (2015), Caliendo and Kopeinig (2008), and Stuart (2010). Abadie and Imbens (2011) suggest that kernel estimators may be more efficient than other matching estimators. Black and Smith (2004) test three competing matching estimators, finding that the Epanechnikov kernel performs the best.

impacts of titling are likely to be non-linear with respect to the propensity score, conditional on the common support available for identification. Consequently, the average impact,  $\Delta^{TT}$ , is likely to be lower than that identified using the OLS estimator if those propensity scores for which common support is weak are more likely to be tenure insecure.

Ichimura and Todd (2007) suggest selecting an optimal bandwidth using Least Squares Cross Validation (LSCV). However, when the common support is weak, the optimal bandwidth that results may be quite large. Large bandwidths introduce excessive smoothing bias by ‘washing out’ conditioning on covariates.<sup>19</sup> For these reasons, we report PSM estimates at two small bandwidths, 0.05 and 0.10, as well as a bandwidth of 0.80, which was determined using LSCV. A bandwidth of 0.05 was chosen because it is the bandwidth at which the smoothing bias is smallest and yet still able to pass standard balancing tests for propensity scores below 0.75. We also report estimates using a bandwidth of 0.10 to demonstrate how rapidly large bandwidths introduce smoothing bias. Two other estimators that rely on PSM can also address smoothing bias: a trimmed PSM estimator that discards observations using the rule of thumb proposed by Crump et al. (2009), and a post-PSM, kernel weighted OLS estimator that uses the sample remaining after matching.<sup>20</sup> We report results for both of these estimators at all three bandwidths. Finally, to assist with comparison to our earlier OLS estimates, we also report an unweighted OLS estimator based on the sample remaining after matching for each bandwidth.

## 5.2 The Distribution of Propensity Scores

In Figure 1, we examine the common support condition by plotting histograms of the estimated propensity scores for both treated (titled) and untreated (no title) dwellings. The top histogram corresponds to dwellings that received a title, while the bottom histogram corresponds to dwellings that did not. The horizontal axis defines intervals of the propensity score, and the height of each bar on the vertical axis indicates the fraction of the relevant sample with scores in the corresponding interval. The histograms clearly show that the support is weak between treated and untreated dwellings. Three-quarters of untreated dwellings lie below the 5<sup>th</sup> percentile of treated dwellings. Moreover, the mean propensity score of treated observations is 0.73, while the mean for the control observations is 0.08. We reject the null hypothesis of equality of means between the propensity scores for two groups at a significance level of 1 percent.

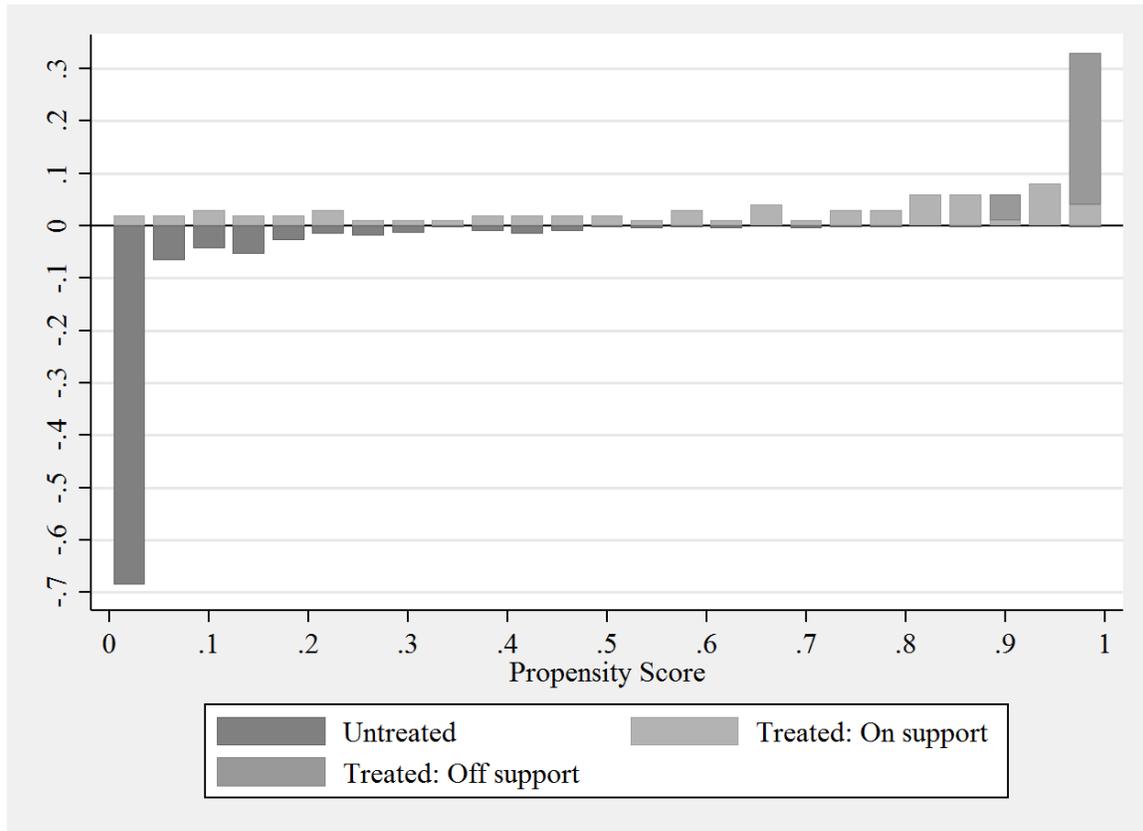
Identification under PSM relies on matching a small number of resettlements with decent quality housing and neighborhood characteristics to a large number of notified slums with decent to low quality housing and neighborhood characteristics. For high values of the probability of participation, we are relying on a small number of untreated observations to construct the counterfactual for a large number of treated dwellings. While the support condition does not fail completely, it holds so weakly that the PSM estimates end up having high variances. This suggests

---

<sup>19</sup> Larger bandwidths induce more smoothing bias as observations less similar to each treated observation receive greater weight in constructing the estimated counterfactuals. As the bandwidth approaches infinity, one is effectively using all controls as an equal match for each treatment. The resulting estimates of the average treatment effect on the treated converge to the coefficient identified from running the baseline OLS model (e.g. column one in Table 2), for the subsample remaining for matching.

<sup>20</sup> Imbens and Rubin (2015) propose an alternate method that involves stratifying the sample into blocks on the basis of the propensity score and then running regressions for each stratum in order to estimate treatment effects. We have attempted this stratification, but based upon the criterion they propose, we are unable to split the sample even once because too few treatments remain in the bottom strata and too few controls in the upper strata. Consequently, we also attempted an alternate stratification based upon these limitations, the details of which are provided in the Online Appendix. Results using this method are similar to the central findings reported below.

**Figure 1: Frequency Distribution of Treated and Untreated, On and Off Support for Epanechnikov Kernel with Bandwidth of 0.05**



Note: Treated off support identifies treatment observations for which there is no common support, and thus dropped from PSM when calculating the average treatment effect on the treated.

that the inability to properly account for the weakness of the common support condition is what may be driving the statistically significant effect from titling identified by the OLS estimator, even though titling is plausibly exogenously assigned.

### 5.3 Propensity Score Matching Estimates

Table 3 presents our PSM estimates with bootstrapped standard errors reported over 2,000 draws. The first row presents the PSM estimate at different bandwidths of the Epanechnikov kernel estimator: 0.05, 0.10, and 0.80 in columns one, two, and three, respectively. Using a bandwidth of 0.05 (column one), in which smoothing bias is the smallest, we find a very small willingness-to-pay for a title of 13.2 percent of a dwelling unit’s estimated monthly rental value (panel one). For comparison, the second row provides the preferred OLS estimator holding fixed the observations used to identify the PSM estimate.<sup>21</sup> Our preferred PSM estimate is three-fifths the OLS estimate

<sup>21</sup> The probit estimates propensity scores for 498 observations. This is less than the 767 observations that enter the OLS model because observations that fall in wards without variation in title status are dropped from the probit. These should be thought of as observations that are dropped due to the failure of the common support condition (because they perfectly predict the binary outcome). Observations for which the probit estimates a zero value for the propensity score include 32 observations and these are also dropped from the matching analysis. All of these are from the control group. When we enter the matching, we have a total of 466 observations, of these, 366 are controls and 100 are treatments. Furthermore, the matching estimator itself drops treatments that fall off of the common support (34 treatment observations when bandwidth is 0.05, and 29 treatment observations when the bandwidth is 0.10 or 0.80).

**Table 3. Propensity Score Matching Estimates**

	(1) ln(Rent) Epanechnikov kernel, BW=0.05	(2) ln(Rent) Epanechnikov kernel, BW=0.10	(3) ln(Rent) Epanechnikov kernel, BW=0.80
Propensity score matching estimate	0.132 (0.124) [n=432]	0.154 (0.131) [n=437]	0.398 (0.107)*** [n=437]
OLS estimates without weights, on support	0.218 (0.112)*	0.209 (0.109)*	0.209 (0.109)*
OLS estimates with weights, on support	0.078 (0.124)	0.105 (0.121)	0.133 (0.097)
Propensity score matching with trimming	0.157 (0.170) [n=125]	0.177 (0.168) [n=125]	0.222 (0.125)* [n=125]

Notes: BW denotes bandwidth. \*, \*\*, and \*\*\* denotes statistical significance at 10 percent, 5 percent, and 1 percent levels, respectively. Parentheses denote standard errors and square brackets denote the total number of observations remaining after the probit and PSM (those that fall on support), and thus effectively the number of observations retained for estimation. Full results of the probit are provided in Table A8 the Online Appendix. OLS estimates include the covariates from model (4) in Table 2. For the propensity score matching estimates, bootstrapped standard errors are reported using 2,000 replications. For the non-weighted OLS estimates, standard errors clustered at the ward level are reported. For the weighted, post-matching OLS estimates, bootstrapped standard errors are reported based on 2,000 replications. OLS estimates without weights, on support present an equivalent comparison using the same observations that are used in the matching estimation. Propensity score matching with trimming discards observations following the rule of thumb proposed by Crump et al (2009).

of 21.8 percent, which is 1.6 percentage points greater than our OLS estimate using the full sample. While the OLS estimate remains statistically significant (although now at the 10 percent level), our preferred PSM estimate is not. The finding of a small, statistically insignificant impact from titling when using the PSM estimator is a key result of this paper.

The impact of titling increases from 13.2 percent to 15.4 percent when the bandwidth is 0.10 (column two). This estimate remains statistically insignificant, although the standard error improves slightly. The PSM estimate generated using the LSCV bandwidth of 0.80 reported in column three suggests a large, statistically significant effect from titling of 39.8 percent. However, as the bandwidth of the kernel estimator increases, so too does smoothing bias. For this reason, in the third row we linearly recondition on the covariates included in our preferred OLS model using the kernel weights used for PSM.

According to the kernel weighted OLS estimator, titling increases property values by 7.8 percent, 10.5 percent, and 13.3 percent, for the 0.05, 0.10, and 0.80 bandwidth cases, respectively. Consistent with our central result, none of these estimates are statistically significant. All three estimates are lower than their PSM counterparts. This suggests that even at the lowest bandwidth smoothing bias is a serious challenge. Since a bandwidth of 0.80 does a better job of fitting observations on the common support, the standard error on the kernel weighted OLS is the smallest for this case, although it remains insignificant. Thus, even when variance is at its lowest under the optimal LSCV bandwidth, the impact of titling remains small and not significant. In the last three rows of Table 3 we also report the PSM estimates using the trimmed sample suggested by Crump

et al. (2009). These estimates align closely with our weighted OLS estimates with estimates of 15.7 percent, 17.7 percent, and 22.2 percent for the three respective bandwidths, although they are again not significant.

A selective comparison of our PSM estimates for our central bandwidth of 0.05 to the OLS estimates provided in Table 2 and Table 3, allow us to decompose the various sources of bias that emerge when using the OLS hedonic estimator to identify the impact of titling relative to PSM.<sup>22</sup> Our kernel-weighted OLS estimate uses the same sample and weights as our central PSM estimate, but our kernel-weighted OLS estimate re-imposes the linear functional form. Thus, the difference between these two estimates of -5.4 percent ( $=7.8-13.2$ ) provides a lower bound of the specification bias that results because the conventional OLS hedonic imposes a linear functional form.<sup>23</sup> The conventional OLS hedonic for the full sample imposes the same functional form as the kernel-weighted OLS, but the kernel-weighted OLS estimate excludes treatments and controls that fall off the common support and weighs control observations based on their common support. Consequently, the difference between these two estimates of 12.4 percent ( $=20.2-7.8$ ) reflects the common support failure bias that emerges because the OLS estimator parametrically identifies a treatment effect when the common support is weak. In absolute terms, common support failure bias is more than twice as large as the specification bias. A similar result can be observed from comparing the variance across both of these channels. Taken together, this suggests that the finding of a statistically significant impact from titling when using the OLS estimator is largely driven by the failure of the OLS estimator to account for the failure of the common support.

#### 5.4 Tenure Security and Possible Heterogeneous Treatment Effects from Titling

The failure of the common support suggests that while the average benefit from receiving a title may be zero across the full sample, the benefit of receiving a title need not be zero for all households, reflecting differences in unobserved tenure security. Consequently, titling may still merit consideration in the context evaluated if it is deployed in a way that targets the most tenure insecure and when it is likely to complement other slum intervention strategies. We present an exploratory analysis of these issues next.

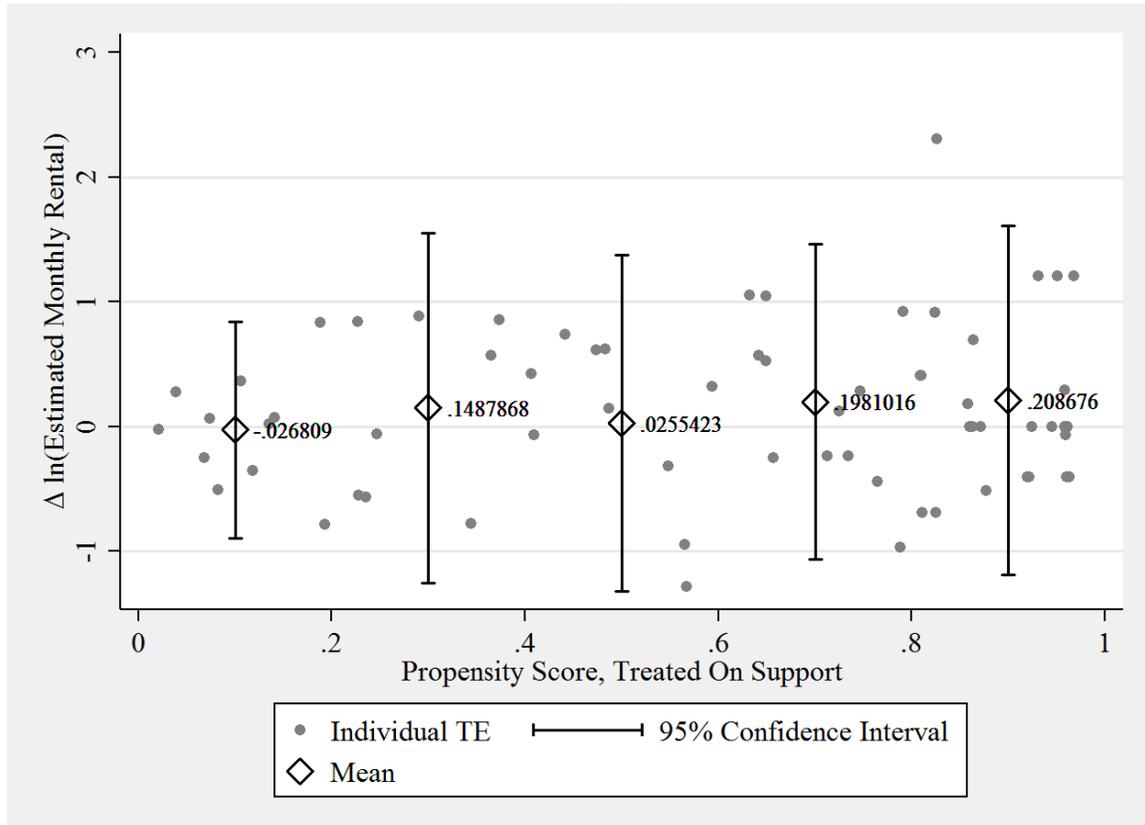
Figure 2 presents *individual treatment effects on the treated* (ITT) from our preferred PSM specification. Bars and means summarize these effects by propensity score quintile. Means by quintile are clearly non-linear with the first and third quintiles reporting means very close to zero and quintiles 2, 4, and 5 reporting means that bound our central OLS estimate. The mean ITTs are not monotonic across the full range of propensity scores reflecting the fact that the propensity score is a univariate index of a multi-dimensional covariate space. The spread in ITT is generally increasing with the propensity score, reflecting weaker balance between the treatment and control groups at higher bins of the propensity score.

Bins with a low mean ITT reflect matches between control dwellings that are more likely to have received slum improvements from the government and like treatment dwellings as well as matches between control dwellings that are otherwise similar to treatment dwellings in terms of observable characteristics. Such matches are likely to reflect a smaller capitalization of titling into

<sup>22</sup> For additional discussion of these two sources of bias in relation to OLS and matching, see Imbens (2015); in relation to the evaluation of anti-poverty programs using OLS and matching, see Ravallion (2008).

<sup>23</sup> This is a lower bound since the weights themselves reflect semi-parametric conditioning using the probit and the PSM estimate still reflects some smoothing bias which is corrected for by the kernel OLS estimate. Thus, the true magnitude of the specification bias from the semi-log OLS model is likely larger.

**Figure 2: Individual Treatment Effect on the Treated for Epanechnikov Kernel with Bandwidth of 0.05**



Notes: Points depict individual treatment effects on the treated for those treatments on support after PSM. Bars and means summarize individual treatment effects on the treated by propensity score quintile.

property values, as those households relied upon most heavily in the control group likely possess greater tenure security. Conversely, bins with a high mean ITT are more likely to reflect matches between control dwellings that did not receive slum improvements and similar treatment dwellings. As our treatment group contains more of the former matches than the latter, the estimate of the impact of titling under PSM is substantially lower than under the OLS estimate.

This graphical analysis suggests that, with respect to our observed covariates, there may exist heterogeneous treatment effects from the receipt of a title. Crump et al. (2008) develop formal non-parametric tests to evaluate whether there is treatment effect heterogeneity with respect to the average treatment effect. We perform similar tests to evaluate treatment effect heterogeneity with respect to the average treatment effect on the treated. Relatedly, Caliendo and Kopeinig (2008) suggest performing *t*-tests to evaluate the characteristics of the treatment observations that remain on support under PSM to those that fall off of support under PSM. As detailed in the Online Appendix, the results of both sets of tests suggests that treatment effect heterogeneity is present in our data. Figure 2 suggests four broad segments which we use to evaluate heterogeneity in ITT (from low to high propensity scores): a 'zero' average ITT segment 1 including quintile 1 (mean ITT of -2.7 percent), a 'non-zero' average ITT segment 2 including quintile 2 (mean ITT of 14.9 percent), a 'zero' average ITT segment 3 including quintile 3 (mean ITT of 2.6 percent), and a 'non-zero' average ITT segment 4 including quintiles 4 and 5 (mean ITT of 20.3 percent).

We performed several *t*-tests to evaluate differences in means between all pairs of zero and

non-zero average ITT segments (segments 1 and 2, 1 and 4, 3 and 2, and 3 and 4).<sup>24</sup> We considered both observed pre-treatment household characteristics as well as the full set of covariates included in our central OLS specification. There are several important and consistent differences between the zero and non-zero average ITT segments which provide insight as to the possible role that titling may provide as a complement or substitute for other slum interventions and as suggested by our conceptual model.

Across all segment comparisons considered, we find that households in non-zero average ITT segments are located in areas with higher quality roads (statistically significant for 2 out of 4 comparisons), are located next to fewer slums (3 of 4), are more likely to have access to legal piped water (2 of 4), have slightly larger households, and are less likely to have high quality, brick walls. This suggests that titling may be a substitute for interventions that directly upgrade the structural integrity of housing units such as wall improvements, and that other interventions such as providing access to legal piped water, may be complements to titling.

In addition, we find that households in segment 2 are more likely to be in a scheduled caste, have a head of household with primary school education or higher, live in denser neighborhoods, are more likely to have a metal roof, and have less water availability than households in segments 1 or 3. While this is in line with our initial observations, this also weakly suggests that titling may provide greater gains to minority groups who may have less tenure security than majority groups. Households in segment 4 are more likely to have a head of household who is older, occupy more living space, are less likely to have a kitchen in a separate room in their house, are more likely to have mosaic floors, and are more likely to have sewage access than households in segments 1 or 3. This is again largely consistent with our initial observations.

### 5.5 Additional Discussion

While our data does not allow us to explicitly discern those slums that received slum improvements by the government from those households that made similar improvements themselves, our finding of a small, statistically insignificant effect from titling under PSM suggests that slum improvements may already enhance tenure security such that providing a formal title provides little additional benefit. This is further confirmed by our analysis of treatment effect heterogeneity on the treated in which certain variables which one could reasonably surmise correspond to greater tenure security are correlated with no treatment effect from receiving a title. In terms of the literature in this area, this suggests an assurance effect from receiving a title that is likely to be small when observable characteristics themselves imply some degree of assurance on their own. In particular, upgrading housing units to be structurally secure with good quality, brick walls, or ensuring dedicated cooking spaces will likely provide some assurance.<sup>25</sup>

Broadly speaking, our finding of a small assurance effect from titling coincides with

<sup>24</sup> A nascent and growing literature examines precisely treatment effect heterogeneity using flexible machine learning and non-parametric approaches, see Athey and Imbens (2015) and Imai and Ratkovic (2013). We also conducted the 'data-driven' analysis suggested by Imai and Ratkovic (2013). Many of the same covariate trends we identify using their ranking approach we confirm in the analysis that follows which is easier to interpret.

<sup>25</sup> Our analysis may also suggest that titling could achieve greater impacts when targeted at disadvantaged minority populations. We believe some caution is in order in regards to drawing this conclusion. While we believe that we have made a convincing case regarding the exogenous assignment of titling, we are not able to make the same case in regards to slum upgrading. If disadvantaged households were less likely to receive slum upgrades when they received a title, than this larger treatment effect may be due to their greater physical vulnerability because of lower quality housing or less secure neighborhoods rather than due to greater vulnerability due to social exclusion. If this is the case, then the policy prescription may not be to provide more titles, but more slum upgrades to historically disadvantaged populations.

previous literature in this area which has found a greater effect from titling in locales where tenure security is weak (Jimenez, 1984, Friedman, Jimenez, and Mayo, 1988; and Lanjouw and Levy, 2002). That said, our identification strategy is better able to account for heterogeneity in observables correlated with unobserved determinants of tenure security, as well as address other un-observables which may impact property values and are also correlated with receiving a title. We have shown that this is critical for our ability to properly identify the impacts of titling in Pune in which slum improvements by the government are considerable. It is not clear whether our identification strategy, if applied to other contexts, would yield as dramatic a disparity in the impacts of titling (relative to conventional OLS hedonic estimates). Other contexts have different pre-existing distributions of observables correlated with tenure security, and so the non-linearity in titling impacts identified here may be more or less important; although PSM makes this transparent in ways that OLS does not. That said, there is suggestive evidence in the summary statistics reported in Jimenez (1984) and the logistic regressions estimated in Friedman, Jimenez, and Mayo (1988) that there may be meaningful differences between titled and untitled households in the contexts examined by those authors across numerous housing unit and household attributes.<sup>26</sup> If so, as our conceptual model has demonstrated and has been confirmed empirically in the context of Pune, to the extent that there is significant heterogeneity in observables correlated with tenure security or un-observables correlated with both property values and titling, then PSM will likely yield more accurate estimates of the capitalization of titling into property values.

## 6. CONCLUSION

Exploiting a quasi-experimental design where titling is exogenously assigned, this paper has examined the implications of titling on the property values of slum dwellers in Pune, India. Using a conventional OLS hedonic estimator, we find a statistically significant effect from titling on housing values of 20.2 percent of the estimated monthly rental value of a housing unit—an effect comparable to that found elsewhere in the literature. In sharp contrast, propensity score matching identifies a much smaller impact from titling of only 13.2 percent, which is no longer statistically significant. Therefore, we find no effect of titling on average on property values. This is the main result of this paper and is robust to our kernel-weighted OLS regression and trimmed PSM estimates. We further show that common support failure bias from using the conventional OLS hedonic dominates specification bias by over two orders of magnitude, although both sources of bias are sizeable. Thus, common support failure bias is likely what drives the finding of a statistically significant impact from titling under the conventional OLS hedonic. We also find evidence that the impact of receiving a title on those who actually received one may vary across several observables that are likely correlated with tenure security. Collectively, these findings suggest that our main result of a small, statistically insignificant impact from titling emerges because slum upgrading and improvements that access municipal services by the government or by households themselves likely already improve tenure security to such an extent that titling has little to no additional impact on property values.

These findings have important implications for public policy interventions that aim to improve the welfare of slum dwellers. As shown here, when tenure security is strong and credit markets are weak, then the provision of a formal title has little additional impact on welfare as capitalized into property values. That being said, we do not perform a comprehensive benefit-cost

---

<sup>26</sup> Lanjouw and Levy (2002) do not report housing attributes by titling status, although they do provide some indication that household characteristics may affect the expected impacts of titling, including whether the head of household is male, which aligns with our findings here.

assessment of titling programs or other interventions, and therefore, titling may still be the preferred intervention strategy once other welfare impacts are taken into account. Our analysis suggests that the formalization of property rights through the provision of a legal title may be less important than the direct targeting of tenure security through other interventions such as on-site slum upgrading and enhancing and expanding access to municipal services and infrastructure. Our analysis suggests that interventions that enhance the structural integrity of housing units are substitutes for titling; although others, such as improving legal piped water access, are complementary.

An additional caveat concerns our results. We are unable to distinguish between slum improvements that result due to government intervention or self-upgrading. Moreover, slum improvements that result from government intervention are likely not random, although we believe that titling is. Consequently, we are unable to directly examine the capitalization of slum upgrading into property values and this how this compares to titling as an intervention strategy. Moreover, while our heterogeneity analysis provides some initial suggestive evidence of the trade-offs between titling and other slum interventions, future work should directly compare alternate intervention strategies and combinations of intervention strategies using randomized control experiments involving titling and alternative forms of slum upgrading. Such experiments should be designed in a way that shed additional insights into the underlying mechanisms by which titling is likely to be meaningful, and that account for the *de facto* rights enjoyed by communities in different political, social, and institutional contexts.

## REFERENCES

- Abadie, Alberto and Guido W. Imbens. 2011. "Bias-Corrected Matching Estimators for Average Treatment Effects." *Journal of Business and Economic Statistics*, 29(1), 1–11.
- Abadie, Alberto and Guido W. Imbens. 2008. "On the Failure of the Bootstrap for Matching Estimators." *Econometrica*, 76(6), 1537–1558.
- Alchian, Armen and Harold Demsetz. 1973. "The Property Rights Paradigm." *Journal of Economic History*, 33(1), 16–27.
- Athey, Susan and Guido W. Imbens. 2015. "Machine Learning Methods for Estimating Heterogeneous Causal Effects." Mimeo, July 20.
- Bapat, Meera and Indu Agarwal. 2003. "Our Needs, Our Priorities: Women and Men from the Slums in Mumbai and Pune Talk about their Needs for Water and Sanitation." *Environment & Urbanization*, 15, 71–86.
- Bento, Antonio M., Maureen Cropper, and Akie Tacheuchi. 2008. "Measuring the Welfare Effects of Slum Improvement Programs: The Case of Mumbai." *Journal of Urban Economics*, 64(1), 65–84.
- Besley, Tim. 1995. "Property Rights and Investments Incentives: Theory and Evidence from Ghana." *Journal of Political Economy*, 103, 903–37.
- Black, Dan and Jeffrey A. Smith. 2004. "How Robust is the Evidence on the Effects of College Quality? Evidence from Matching." *Journal of Econometrics*, 121(1–2), 99–124.
- Brasselle, Anne-Sophie, Frederic Gaspart, and Jean-Philippe Platteau. 2002. "Land Tenure Security and Investment Incentives: Puzzling Evidence from Burkina Faso." *Journal of Development Economics*, 67(2), 373–418.

- Caliendo, Marco and Sabine Kopeinig. 2008. "Some Practical Guidance for the Implementation of Propensity Score Matching." *Journal of Economic Surveys*, 22(1), 31–72.
- Cronin, Victoria and Peter Guthrie. 2011. "Community-led Resettlement: From a Flood Affected Slum to a New Society in Pune, India." *Environmental Hazards*, 10(3–4), 310–326.
- Cropper, Maureen L., LeLand B. Deck, and Kenneth E. McConnell. 1988. "On the Choice of Functional Form for Hedonic Price Functions." *The Review of Economics and Statistics*, 70(4), 668–675.
- Crump, Richard K., V. Joseph Hotz, Guido W. Imbens, and Oscar A. Mitnik. 2009. "Dealing with Limited Overlap in Estimation of Average Treatment Effects." *Biometrika*, 96(1), 187–199.
- Crump, Richard K., V. Joseph Hotz, Guido W. Imbens, and Oscar A. Mitnik. 2008. "Nonparametric Tests for Treatment Effect Heterogeneity." *The Review of Economics and Statistics*, 90(3), 389–405.
- Demsetz, Harold. 1967. "Toward a Theory of Property Rights." *American Economic Review*, 57(2), 347–59.
- De Soto, Hernando. 2000. *The Mystery of Capital: Why Capitalism Triumphs in the West and Fails Everywhere Else*. New York: Basic Books.
- Feder, Gershon and Akihiko Nishio. 1999. "The Benefits of Land Registration and Titling: Economic and Social Perspectives." *Land Use Policy*, 15, 25–43.
- Feder, Gershon, Tongroj Onchan, Yongyuth Chalamwong, and Chira Hongladarom. 1988. *Land Policies and Farm Productivity in Thailand*. Baltimore: Johns Hopkins University Press.
- Field, Erica and Maximo Torero. 2008. "Do Property Titles Increase Credit Access Among the Urban Poor? Evidence from a Nationwide Titling Program." Mimeo, Duke University.
- Friedman, Joseph, Emmanuel Jimenez, and Stephen K. Mayo. 1988. "The Demand for Tenure Security in Developing Countries." *Journal of Development Economics*, 29, 185–98.
- Frölich, Markus. 2004. "Finite-Sample Properties of Propensity Score Matching and Weighting Estimators." *The Review of Economics and Statistics*, 86(1), 77–90.
- Galiani, Sebastian and Ernesto Schargrodsy. 2010. "Property Rights for the Poor: Effects of Land Titling." *Journal of Public Economics*, 94(1), 700–729.
- Goodman, Alan C. 1978. "Hedonic Prices, Price Indices and Housing Markets." *Journal of Urban Economics*, 5(4): 471–484.
- Guo, Shenyang and Mark W. Fraser. 2015. *Propensity Score Analysis, Statistical Methods and Applications*. 2<sup>nd</sup> ed. Los Angeles: Sage Publications.
- Heckman, James J., Hidehiko Ichimura, Jeffrey A. Smith, and Petra E. Todd. 1998. "Characterizing Selection Bias Using Experimental Data." *Econometrica*, 66, 1017–1098.
- Heckman, James J., Hidehiko Ichimura, and Petra E. Todd. 1997. "Matching as an Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Programme." *The Review of Economic Studies*, 64, 605–654.
- Heckman, James J. and Richard Robb. 1986. "Alternative Methods for Solving the Problem of

- Selection Bias in Evaluating the Impact of Treatments on Outcomes.” *Drawing Inferences from Self-Selected Samples*. Wainer, Howard, ed. New York: Springer-Verlag.
- Ichimura, Hidehiko and Petra E. Todd. 2007. “Chapter 74: Implementing Nonparametric and Semiparametric Estimators.” *Handbook of Econometrics, Volume 6B*. James J. Heckman and Edward E. Leamer, editors. Amsterdam: North Holland.
- Imai, Kosuke and Marc Ratkovic. 2013. “Estimating Treatment Effect Heterogeneity in Randomized Program Evaluation.” *Annals of Applied Statistics*, 7(1), 443–470.
- Imbens, Guido W. 2015. “Matching Methods in Practice: Three Examples.” *The Journal of Human Resources*, 50(2), 373–419.
- Imbens, Guido W. and Donald B. Rubin. 2015. *Causal Inference for Statistics, Social, and Biomedical Sciences: An Introduction*. New York City: Cambridge University Press.
- Imbens, Guido W. and Jeffrey M. Wooldridge. 2009. “Recent Developments in the Econometrics of Program Evaluation.” *Journal of Economic Literature*, 47(1), 5–86.
- Jimenez, Emmanuel. 1984. “Tenure Security and Urban Squatting.” *The Review of Economics and Statistics*, 66(4), 556–67.
- Kuminoff, Nicholai V., Christopher F. Parmeter, and Jaren C. Pope. 2010. “Which Hedonic Models Can We Trust to Recover the Marginal Willingness to Pay for Environmental Amenities?” *Journal of Environmental Economics and Management*, 60(3): 145–160.
- Lall, Somik V. and Mattias Lundberg. 2008. “What are Public Services Worth, and to Whom? Non-parametric Estimation of Capitalization in Pune.” *Journal of Housing Economics*, 17(1), 34–64.
- Lancaster, Kelvin J. 1966. “A New Approach to Consumer Theory.” *Journal of Political Economy*, 74(2), 132–157.
- Lanjouw, Jean and Philip Levy. 2002. “Untitled: A Study of Formal and Informal Property Rights in Urban Ecuador.” *Economic Journal*, 112, 986–1019.
- Office of the Registrar General and Census Commissioner (ORGCC), India. 2011. *Census of India 2011*. Accessed, August 2016. <http://censusindia.gov.in/>.
- Ravallion, Martin. 2008. “Evaluating Anti-Poverty Programs.” Chapter 59, *Handbook of Development Economics, Volume 4*. T. Paul Schultz and John Strauss, editors. Amsterdam: North Holland.
- Rosen, Sherwin. 1974. “Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition,” *Journal of Political Economy*, 82(1), 34–55.
- Rosenbaum, Paul R. 1997. “The Role of a Second Control Group in an Observational Study.” *Statistical Science*, 2, 292–316.
- Rosenbaum, Paul R., and Donald B. Rubin. 1983. “The Central Role of the Propensity Score in Observational Studies for Causal Effects.” *Biometrika*, 70, 41–55.
- Sen, Srinanda and Jane Hobson. 2005. *The Pune Slum Census*. Shelter Associates. Pune, India.
- Stuart, Elizabeth A. 2010. “Matching Methods for Causal Inference: A Review and Look Forward.” *Statistical Science*, 25(1), 1–21.

Times News Network (Times). 2011. "32.5% Population of City Lives in Slums." *The Times of India*. January 19, 2011.

United Nations. 2003. Human Settlements Program. *Global Report on Human Settlements 2003: The Challenge of Slums*. London: Earthscan Publications Ltd.