

Measuring the Risk on Housing Investment in the Informal Sector: Theory and Evidence from Pune, India*

Mudit Kapoor, David le Blanc
The World Bank
Washington, DC 20433, USA

7th October 2004

1 Introduction

In this paper, we provide an economic framework to analyze investment in informal housing in developing countries. Squatting is a major form of housing tenure in urban centers of developing countries.¹ With rapid growth of cities in the developing world, informal housing has not only subsisted, but has indeed thrived.² It is estimated that in Pune, India, 40 percent of the population live in informal settlements.³ It has also become increasingly clear that slum dwelling units do not convey the entire picture of "informal housing". In countries where land markets are very distorted, slums (settlements consisting of shacks, without basic services and utilities) are only the most visible part of informal housing, the tip of the iceberg. Informal housing also consists of dwelling units with construction (concrete, bricks, etc.) similar to those in the formal sector. Moreover, some

*The research and data collection have been co-funded by a World Bank research program grant on "Urbanization and Quality of Life" and the UK DFID's Urban Knowledge Generation and Toolkits program. We are grateful to Somik Lall for providing us access to the data used here. *The findings, interpretations and conclusions expressed in this paper are those of the authors and do not necessarily represent the view of the World Bank, its Executive Directors, or countries they represent.*

¹See Friedman, Jimenez and Mayo (1988) for an overview.

²For example, in Morocco, it is estimated that the informal housing stock has grown at an annual rate of 4 percent between 1992 and 2002, faster than the formal stock. In that country, informal housing now accounts for 10 to 15 percent of the total housing stock, in spite of large programs of upgrading and resettlement in the past 20 years by successive Governments. In other countries, the informal to formal ratio in the housing stock is even higher.

³One important issue for cross-country comparisons is the definition of "slums" and "informal housing". The empirical section will clarify the definition of "informal housing" used in our analysis.

residents in informal units have characteristics very similar to those of households living in formal neighborhoods.⁴ In such cases, the main difference between a formal and informal dwelling is that the informal dwelling has been built outside administrative procedures, be it at the titling level (no title for the land parcel requested or obtained from the titling authority) or at the construction level (no building permit requested or obtained from the urban planning authority), and thus faces a risk of destruction. This in turn may have implications for supply of basic infrastructure (roads) and utilities (electricity, water).

These elements show that in many countries construction of informal dwellings, rather than being an exception, is considered by households and firms as a viable alternative to constructing in the formal sector. Therefore, in equilibrium, both forms of investment coexist, and policy actions on either sector have an influence on the other sector. Hence, in order to devise efficient housing policies in those countries, one has to understand how the two alternatives compare and what drives their relative attractiveness.⁵ Our model formalizes this observation. Formal and informal housing investment are distinguished in terms of costs and risks. Theoretically, we show that the premium for the formal sector should be higher for prices than for rents, because the difference between the two premiums is a direct measure of the risk associated with investment in the informal sector. Using household data from Pune, India, we estimate that the perceived risk on housing investment in the informal sector is equivalent to an annual destruction rate ranging between 1 and 2 percent. We use our approach to estimate investors' risk in the informal sector for other countries from results in previous papers.

Studies on informal housing have primarily focused on two issues. The first is the estimation of price or rent differentials between squatting and non-squatting areas (Jimenez, 1984, Jimenez, Friedman and Mayo, 1988). In these two papers, the focus is on the demand side of the housing market. Differentials in rents (or prices) between squatter and non-squatter units are attributed to eviction risk in the informal sector. The second ap-

⁴For example, in Morocco, the 1994 Census identified about 230,000 households living in slums (shacks), and more than 700,000 living in concrete informal dwellings.

⁵General equilibrium effects are mentioned in introduction by Lanjouw and Levy (2002), p 988, but not considered afterwards.

proach focuses on models of incomplete housing markets, where the incompleteness comes from the absence of formal property rights (Lanjouw and Levy, 2002), or from large transaction costs (Hoy and Jimenez, 1991). These models focus on interactions in the housing market, between buyers and sellers or between landlords and renters.⁶ However, there has been a great body of literature focussing on the effects of security of land tenure on investment in the rural sector, for example the effects of titling on investment (see Besley, 1995, Banerjee, Gertler and Ghatak, 2002), or the impact of expropriation risk on use of soil fertilizer in rural China (Jacoby, Li and Rozelle, 2002).

A disturbing point of demand-side models of tenure security as developed in the two papers above is the fact that renters and owners are not distinguished, and thus are assumed to face the same risks.⁷ Empirically, the two papers find that differentials between the formal and informal sectors were greater for prices than for rents, which cannot be explained by their underlying model.⁸ One reason for these findings must be that renters and owners in squatter dwellings do not face the same risks. Lanjouw and Levy (2002), focusing on formal and informal property rights, note that the benefits of having a title are manifold: having a title increases tenure security, which allows households to capture the benefits of durable investments in housing; it also lowers transaction costs by clarifying the nature of property rights; it allows households to use their parcel of land or their dwelling as collateral for bank loans. Finally, when urban land is titled, utility providers are able to charge for services, and municipal governments can use property taxes to finance infrastructure, resulting in better serviced neighborhoods and dwellings. It is clear that the first three items on this list concern owners, rather than renters. Thus, it seems reasonable to argue that the risk of building in the informal sector (be it called "eviction

⁶Another strand of the literature has focussed on estimating the costs and benefits of "site and services" programs for low-income housing (for example, Kaufman and Quigley, 1987).

⁷In their introduction, Friedman, Jimenez and Mayo (1991) themselves note: "The rent in a squatter area should be lower than rent in a formal area. One source of this difference is that a renter household must be compensated for the inconvenience of possibly having to move, if evicted from a squatter area. On the other hand, the price discount of owned units in squatter areas should compensate for the possible loss of the housing capital, in addition to moving cost".

⁸Jimenez (1984) found a 18 percent differential for the rents of rental units and 58 percent for the prices of owner-occupied units. In Friedman, Jimenez and Mayo (1988), the corresponding figures are 15 and 25 percent, respectively.

risk” as in the earlier literature on the subject, or broken down into different components as above) is mainly borne by the owner of the dwelling, and relates to the certainty of present and future cash flows from the investment. On the other hand, renters with high moving costs may care about physical eviction risk.

In this paper we clarify these issues. We show that appropriate data on rents and prices of dwelling units in the formal and informal sectors can be used to isolate two economic variables: the difference in rates of return on housing investment, which reflects the risks borne only by investors, and the premium on rents for the formal sector, which reflects conditions prevailing in rental markets. These two quantities have different economic meanings, and should therefore be analyzed separately. We consider a simple model of investment in the housing market, where investors can choose between two sectors: the formal sector, where physical investment faces no risk of destruction, and the informal sector, where investment, in each period, is subject to an exogenous risk of destruction. Construction costs differ between the two sectors. On the demand side, all households are renters.⁹ Renters shop for dwelling attributes and do not care about the sector (formal or informal) *per se*. There is no moving cost, and renters can choose freely between the formal and informal sector (no segregation in the housing market).

The implications of this simple model are the following. Free mobility of renters implies that rents in the two sectors must be equal. On the supply side, the higher risk faced by investors in the informal sector must be reflected in a higher return on equity invested. Since rents are equal in the two sectors, the values of houses will be lower in the informal sector. Returns on investment, measured by the rent-to-value ratio, will thus be higher in the informal sector. This result relies only on investment side considerations and is independent of assumptions on the households’ side. Thus it will hold even when there are mobility costs or when households have intrinsic preferences for living in the formal sector.

We use a survey conducted by the World Bank in the city of Pune, India. Among

⁹Owner-occupying households are seen both as investors and renters, insofar as they rent their dwellings to themselves.

other features, this survey had the peculiarity of asking the households, regardless of tenure status, questions about the market rent and value of their dwelling. Thus, for every dwelling in the sample, we have estimates of both the value and the market rent of the unit, given by the household actually occupying that unit. This allows us to calculate individual rates of return for each unit, without facing the typical selection bias problems, each house serving as its own control.¹⁰ Comparing the distributions of returns in the informal and formal sectors, we obtain the following results: (i) Rates of return are significantly higher in the informal sector, as predicted by the model. The mean rate of return in the formal sector is 6.7 percent, versus 8.1 percent in the informal sector. This corresponds to a 1.4 point average premium, or 22 percent premium in relative value. (ii) These figures imply a perceived risk on housing investment in the informal sector equivalent to an annual destruction rate ranging between 1 and 2 percent, depending on assumptions regarding the time horizon of housing investors. (iii) The two distributions of rates of return present highly idiosyncratic components, and traditional explanatory variables proxying either the strength of informal property right (Lanjouw and Levy, 2002), or lower perceived risks of eviction (Jimenez, 1984, Friedman, Jimenez and Mayo, 1988) have little power to explain the differences in individual rates of return between the formal and informal sector.

Our framework allows us to rationalize empirical findings of previous papers, which found that premiums for the formal sector, as measured by including a dummy for formal dwellings in otherwise classical hedonic equations, were always greater for prices than for rents. We show that the difference between the two premiums directly reflects the risks on housing investment in the informal sector and thus, should be positive. Moreover, from this difference in premiums we can estimate the risk faced by informal investors, as in our case. Calculations indicate that physical risk on informal investment was four to five times higher in Davao than in Manilla at the beginning of the 1980s.

Lastly, we try to understand what factors may drive the rental premium for the formal

¹⁰In a typical survey, the selection problem would arise from the fact that only the rents are known for the rental units, and the prices or estimated values for the dwelling occupied by their owners. A similar technique is used by Lanjouw and Levy (2002). In their paper, they used a survey where households living in untitled houses were asked to assess the value of their dwellings if it were titled.

sector. In Pune, India, this premium is 35 percent, after controlling for observable characteristics of the dwelling and of the neighborhood at our disposal. This figure is comparable to but slightly higher than those found in previous studies.

The remainder of the paper is organized as follows. In section 2, we present the basic model. Section 3 briefly presents our data. The main results of the paper are given in section 4. Section 5 presents some policy implications of the model. Section 6 presents conclusions.

2 The Model

We consider a simple model of investment in the housing market where investors can choose between the informal and formal sector. On the demand side there are households who shop for the attributes of dwellings. Land is used only to produce housing, and is not valued *per se* by households. There is only one type of house.

2.1 The Investor's Problem

Define:

ρ	opportunity cost of capital of the investor
β	$\equiv \frac{1}{1+\rho}$
d	exogenous destruction factor, as perceived by investors
δ	$\equiv 1 - d$
R_f	annual unit rent in the formal sector
R_i	annual unit rent in the informal sector
C_f	construction costs in the formal sector
C_i	construction costs in the informal sector.

Constructions costs in the formal (respectively informal) sector are assumed to be a convex function of the quantity of housing units produced in this sector, q_f (respectively q_i). In both sectors, investors maximize expected profit from the investment. To simplify,

we suppose that there is no inflation in the economy so that the rent of a dwelling stays the same forever. In the formal sector, cash flows include contemporaneous construction costs, $-C_f(q)$, and future rents totaling qR_f per period. Since investment in the formal sector benefits from security of tenure, cash flows generated by the investment last forever.

In the informal sector, cash flows include contemporaneous construction costs, $-C_i(q)$ and future rents, qR_i per period. But because of lack of formal land rights, the life span of the dwelling is now uncertain. In each period, the dwelling is subject to an exogenous probability of destruction d . When the unit is destroyed, the whole investment is lost and rent flows stop forever.

The expected discounted profit of investment in the formal sector is:

$$\pi_f(q) = -C_f(q) + q \sum_{n=0}^{\infty} R_f \beta^n = -C_f(q) + q \frac{R_f}{1-\beta}. \quad (1)$$

The last term of the RHS member of equation (1) is the value of the investment.

Denote

$$V_f \equiv \frac{R_f}{1-\beta} \quad (2)$$

the unit value of housing investment.

Denoting $\delta \equiv 1 - d$, the probability that the unit will *not* be destroyed in a given year, the expected discounted profit of investment in the informal sector can be written as:

$$\pi_i(q) = -C_i(q) + q \sum_{n=0}^{\infty} R_i (\beta\delta)^n = -C_i(q) + q \frac{R_i}{1-\beta\delta}. \quad (3)$$

Let the value of the investment V_i be defined in the same way as in the formal case. Maximization of profits in both sectors gives:

$$C'_f(q) = \frac{R_f}{1-\beta}, \quad C'_i(q) = \frac{R_i}{1-\beta\delta}. \quad (4)$$

Inverting those two relations, we get the formal and informal sector supply curves:

$$q_f^s = g_f\left(\frac{R_f}{1-\beta}\right), \quad q_i^s = g_i\left(\frac{R_i}{1-\beta\delta}\right). \quad (5)$$

Rates of return on housing investment in both sectors are calculated as $r_f = R_f/V_f$ in the formal sector, and $r_i = R_i/V_i$ in the informal sector. From (2), we get:

$$r_f = 1 - \beta, \quad r_i = 1 - \beta\delta \quad (6)$$

such that $r_f < r_i$.

2.2 The Households' Equilibrium Condition

Households are assumed to be renters and move freely, that is, there are no moving costs.¹¹ Moreover they do not care about the sector in which they live (i.e., formal or informal) *per se*.¹² Thus the equilibrium condition is simply:

$$R_f = R_i = R_e \quad (7)$$

2.3 Equilibrium Determination

Quantities and rents in equilibrium, R_e and q_e , are determined by equating supply and demand in the rental market. Total quantity of housing supplied equals:

$$q^s(R_e) = q_f^s(R_e) + q_i^s(R_e). \quad (8)$$

Demand for housing equals

$$q^d(R_e) = D(R_e). \quad (9)$$

¹¹In particular, there is no return to tenure length (for example, no accumulation of social capital).

¹²The two assumptions on households are made only for the clarity of the argument. They will be relaxed in the empirical section.

R_e and q_e are functions of the taste parameters, the technology parameters, the actualization factor of the investors and the perceived risk on informal investment d . Figure 1 shows how formal and informal quantities and rents are determined in equilibrium.

[INSERT FIGURE 1 HERE]

2.4 Comparative Statics

We are mainly interested in two key parameters : the investors' opportunity cost of capital ρ , reflected in β , and the exogenous risk of destruction d , reflected in δ . Formal derivations of the comparative statics of the model are relegated to Appendix 1.

2.4.1 Changing the Investors' Opportunity Cost of Capital

Looking at changes in the discount factor of investors β , we obtain :

$$\frac{\partial R_e}{\partial \beta} < 0, \quad \frac{\partial q_e}{\partial \beta} > 0, \quad \frac{\partial q_f}{\partial \beta} > 0, \quad \frac{\partial q_i}{\partial \beta} \leq 0.$$

These results are quite intuitive. As the investors' opportunity cost of capital rises, the discount factor β decreases, that is, future profits are discounted more heavily. This has two effects: i) on the whole, housing investment becomes less attractive vis-à-vis other investments, ii) future risks of destruction become less of a problem, and the difference between the two sectors diminishes. This causes informal investment to decrease less than formal investment. Overall, housing quantity supplied decreases, and equilibrium rents increase.

2.4.2 Changing the Investors' Perceived Risk of Destruction

We now look at the changes in the exogenous destruction risk, d . The key intuition of the model is the following. Since demand does not depend directly on d and supply depends negatively on it, it follows that the equilibrium quantity of housing is negatively related to d , whereas the equilibrium rent is an increasing function of d . Moreover, since formal supply is not affected by d and is upward sloping, a positive change in d results in higher

housing quantity in the formal sector, more than offset by a decrease in housing quantities in the informal sector. The reverse is true when a decrease in d , that is an increase in δ , is considered. Figure 1 illustrates the argument graphically. We obtain

$$\frac{\partial R_e}{\partial \delta} < 0, \quad \frac{\partial q_e}{\partial \delta} > 0, \quad \frac{\partial q_f}{\partial \delta} < 0, \quad \frac{\partial q_i}{\partial \delta} > 0.$$

Thus, lowering the destruction rate d (or raising δ) has the effect of lowering the equilibrium rent, augmenting the total quantity of housing, but lowering the quantity of formal housing. The intuition is simple: as d decreases and δ closes up to one, informal land becomes more and more indistinguishable from formal land. Since the cost differential does not change, investment in informal housing becomes more attractive with respect to investment in formal housing, thus raising the former and lowering the latter. The overall effect turns out to be positive, since the overall return on housing investment has increased.

3 The Data

The empirical analysis is based on data from a survey conducted by the World Bank between August and October 2002 in the city of Pune, India. The respondent sample comprises 2,849 households. The survey documents socio-demographic characteristics of households, characteristics of dwelling units, tenure status, accessibility to infrastructure facilities like water, electricity and transport. In addition, the survey also provides information on the valuation of the units by households themselves, in the form of the two following questions:

- Estimated present market price for a similar unit in this neighborhood,
- Estimated monthly rental value for a similar unit in this neighborhood.

The two questions are asked to all households irrespective of their tenure status. Thus for every dwelling we have estimates of both the rent and the price.

We now give some details on the way we constructed the main variables used in our analysis. The main focus of the paper is on the distinction between formal and informal housing. The Pune questionnaire included a question aimed at classifying the dwelling into the following categories: a) non notified squatter settlement, b) notified squatter settlement, c) resettlement, d) unauthorized colony, e) wadas,¹³ f) MHADA (Maharashtra Housing and Area Development Authority) plots, g) MHADA flats, h) cooperative housing, i) employer housing (government or private), j) private builders/colonies, k) Core city area, l) chawls,¹⁴ m) urban village. We define informal housing as the union of categories a), b) and d). Formal housing is defined as the union of all other categories.

Most previous studies in the field have based their empirical distinction between formal and informal housing on the existence of a formal title for the dwelling. In our dataset, questions about the existence of a legal title and its type were asked only to the subset of owners, which comprises 1,999 observations.¹⁵ That is the reason why we adopt another definition for formal and informal housing. Our definition allows us to keep the sample of renters (850 households) available for the analysis. The discrepancy between the two definitions can be examined on the subsample of owners. Only 54 in 872 (6.3 percent) dwellings classified in the "informal" category have a title, and only 15 in 1,127 (1.3 percent) dwellings classified as "formal" have no formal title. Considering this good correspondence between the two variables, we found it worthwhile to keep the whole sample of observations instead of working with only the subsample of owners. However, in the empirical analysis we compare our results to those obtained with the "title" variable on the subsample of owners to assess the robustness of our findings.

A key variable in the analysis is the rate of return on housing investment. Denoting R (respectively V) the estimate of the annual market rent (respectively price) of a dwelling, the rate of return for this dwelling is defined as $r = R/V$.¹⁶ However, rates of return could

¹³Wadas are old-style houses typically located in inner cities.

¹⁴Chawls are multistoried one-room tenements.

¹⁵This makes sense, because households renting their units need not know or even care about the existence of a title; nor should the precise nature of the title be known, at least in some cases.

¹⁶Strictly speaking, rates of return on housing should account for (expected) capital gains. However we have no data on capital gains ; besides, there is no reason to suppose that housing prices should evolve differently in the formal and informal sectors. In that case, taking only rental returns into account is

not be computed for every observation in our sample. For 35 observations, information was missing on rent only; for 101 observations, information was missing on price only; and 104 observations had missing values for both rent and price. Thus 240 observations were dropped, reducing the size of our sample to 2609.

Other variables used in the analysis need some explanation. We begin with dwelling unit characteristics. From the survey questions, dummies are constructed indicating the absence of sewerage, the absence of toilets in the unit, the condition of the roof and the walls, and water availability. Conditions of the roof and walls are defined from the predominant materials used in their construction. "Good condition" for roofs or walls means that brick, stone, or concrete have been used. Households were asked detailed questions about their water sources, the providers of water, the number of hours of availability of water from each source per day, connections costs, monthly payments, and distance and time to the water source when applicable. Appendix 2 gives some details on the answers to water module in our sample. From preliminary investigations, we constructed three types of variables to apprehend water availability:

- Dummies for primary water source. We isolated the three "formal" water sources, individual water connection, individual sub-connection, and community tap. We expect rents and prices to vary with different types of connection available.
- Dummy for secondary source of water.
- Availability in hours per day, interacted with the three dummies for formal water sources. We expect rents and prices to vary with water availability, though differently depending on the type of source. For example, we would expect a high availability of water to have more impact on rents for an individual connection than for a community tap.

With respect to household characteristics, we use dummies for the highest educational level completed by the household head. The first dummy groups people reporting no

enough to account for differences in returns between the two sectors.

education at all. The second group includes people reporting some non-formal schooling, some formal schooling or primary school. The third group includes middle, secondary and higher secondary school. The last group includes technical training, college or post graduate levels. We also use variables specific to the Indian context, and in particular to Pune. These concern religion, caste and language. Religion is represented with a dummy for hindu, which is the majority religion in Pune. Caste is represented by a dummy variable taking the value 1 if the household is from a scheduled caste (that is belongs to the lower caste level). Language is taken into account by a dummy taking the value 1 for people reporting Marathi as their mother tongue. A measure of per capita consumption, constructed from a report of fortnightly expenditures at the household level, is used. We also constructed measures of assets held by the households. Apart from homeownership, which is taken into account directly, the survey asked questions about possession and purchase price of 20 kinds of durable assets. Based on the average purchase value reported in the sample, the four most important assets are cars, two-wheelers, furniture and rugs, and TV/VCRs. The asset indicators we consider are dummies for possession of cars, two-wheelers, and TV/VCRs, and a dummy taking the value 1 when reported value of furniture exceeds 5000 rupees.

4 Results

4.1 Summary Statistics

Before reporting the main empirical findings, it is interesting to look at summary statistics of the main variables used in our empirical analysis (see Table 1). It is evident from Table 1, that informal and formal sectors differ sharply in terms of household and dwelling unit characteristics. Households living in informal sector, on average have more children and lower per capita consumption and more often belong to the scheduled (lower) caste. More than 50 percent of them have either no education or primary education, while in the formal category approximately 70 percent of households have secondary, higher or technical education. Differences in economic well being are further reflected in ownership

of assets like television, two wheelers, automobile and furniture. A larger fraction of formal households possess these assets.

Dwelling units in the two sectors also differ sharply. Only 16 percent of informal units have good quality roofs versus more than 50 percent in formal sector. Informal dwelling units are smaller and lack proper toilet facilities and sewerage more often. Interestingly, individual water connections are as frequent in the informal sector as in the formal sector. However, formal households rely more on individual sub-connections and less on community taps for water supply whereas the reverse is true for informal households. Lastly, in the informal sector, average rent is three times lower and average value four times lower than in the formal sector.

4.2 Analysis of Rates of Return on Housing

The main implication from our investment model is that rates of return on housing investment should be higher in the informal sector than in the formal sector. This can be checked directly from our data, which allow us to compute individual rates of return for every dwelling unit in the sample. Starting from our sample of 2,609 observations, we drop 47 observations for which the rate of return is less than .005 or greater than .4. In the rest of the empirical analysis, we use this restricted sample. The distributions of rates of returns in the two sectors are shown in Figure 2. The two distributions look alike, the informal distribution being a rightward mean-shifted version of the formal one. Regressing rates of return on a dummy for formal sector, we obtain that the average rate of return is 6.7 percent in the formal sector and 8.1 percent in the informal sector, and the difference is statistically significant (Table 2a). Term money market rates at the time of the survey were around 6 to 7 percent. The averages we get thus seem consistent with *a priori* reasonable values of investors' opportunity cost if capital.¹⁷

To assess the robustness of this conclusion, we estimate the following equation,

¹⁷Recall that expected capital gains are not included in these estimations of housing returns, so that the actual value of these returns is higher than the one we computed.

$$r = c + \alpha F + \beta O + \gamma(F \times O) + \varepsilon, \quad (10)$$

where r denotes the rate of return, F is the formal sector dummy and O is a dummy indicating home ownership. Recall that questions about rent and price of the dwelling are asked to every respondent, irrespective of their tenure status. If owners and renters differ systematically in their valuations of rents and prices,¹⁸ results from a simple regression of rates of return on the formal dummy will reflect a composition effect, since self-reported owners are more frequent in the informal sector. Results of estimation are reported in Table 2a. The first column reports results with the restriction $\beta = \gamma = 0$, the second column reports results with $\gamma = 0$, and in the third column we report results without any restrictions on parameters. From these regressions, we conclude that r is approximately 22 percent higher in the informal category. As a supplementary robustness check, we use an alternative specification where we use log value of r as a dependent variable. Results, reported in Table 2b, are very similar to our previous specification. We also ran regression (10) on the subsample of owners, with F denoting the existence of a legal title for the dwelling. Results, presented in Appendix 3, are once again very similar to those in table 2b.

4.3 Estimating Implied Destruction Risks on Informal Housing

Getting back to the theoretical model, we can relate the difference in rates of return on housing in the two sectors to the two parameters governing that difference in the model, i.e. the exogenous destruction risk and the investors' discount factor. For a given value of the investor's opportunity cost of capital (OCC) ρ , there is a one-to-one mapping between the differential in rates of returns the model and the destruction rate d . Indeed, from formula (6), we get:

¹⁸For example, owners could tend to overstate their unit's value because of idiosyncratic match between them and their unit, and to underestimate the rent because they do not account for the risk of renting to another household. If this were the case, computed rates of returns would be systematically lower for owners, *ceteris paribus*.

$$d = \rho \left(\frac{r_i}{r_f} - 1 \right)$$

Since $\left(\frac{r_i}{r_f} - 1\right) \approx \ln\left(\frac{r_i}{r_f}\right)$, we can directly approximate d by multiplying the investors' OCC by the differential in rates of return, that is, 22 percent. Exact calculation gives the following results. When the OCC of investors is 4 percent, the perceived risk of destruction is only 1 percent. For an OCC of 5 percent (respectively 6 percent), it is 1.2 percent (respectively 1.5 percent). This is the first measure of this risk we are aware of. Presenting the results this way allows one to become aware that although price differentials between the two sectors are quite high, the implied perceived destruction risk is low.

One of the reasons for this low estimate lies in the infinite investment horizon considered in the model. We also consider alternative values for the investment horizon of 20 and 50 years.¹⁹ These values seem good bounds for a realistic horizon (from Table 1, it can be seen that dwellers in the informal sector have been in their dwelling for 24 years). When the investment horizon is restricted to 50 years, implied perceived risk of destruction rises. With a 20 year horizon, the implied destruction rates range between 2 to 3 percent, which is fairly high. To understand these results, recall that the differential in rates of return is governed by the "life expectancy" of housing investment in the two sectors. Thus, a very small risk of destruction can result in huge price differences between the two markets because future flows are heavily discounted. However, when the rate of destruction is held constant and the horizon of investment diminishes in both sectors, only cash flows closer to present time remain, and these flows are those which face less uncertainty in the informal sector. Thus, the shorter the horizon, the larger the destruction rate has to be for a given value of the differential in rates of return.

¹⁹Rates of return in the two sectors are now given by $r_f = \frac{1-\beta}{1-\beta^n}$, $r_i = \frac{1-\beta\delta}{1-(\beta\delta)^n}$, where n is the time horizon of investment.

4.4 Explaining Individual Rates of Return

We now try to explain the distribution of rates of return within each sector. From the theoretical model, in the formal sector rates of return should be related to exogenous factors affecting the investor's opportunity cost of capital, whereas in the informal sector, rates of return should also depend on factors affecting the risk to informal investment.

Opportunity cost of capital could vary with household characteristics. A priori shifters include variables describing the position of the household in the life cycle (age, family type), variables that could be linked to access to financial markets (income, assets, education), and variables that could reflect segregation, insofar as these prevent households from accessing differentiated forms of investment (sex, caste, religion).

Concerning the risk to investment in the informal sector, previous literature gives us three directions to investigate. The first one concerns the existence of informal property rights, and their strength. Lanjouw and Levy (2002) make a distinction between transferable and non-transferable informal property rights. Transferable property rights may include title, the extent to which property boundaries in the neighborhood are settled, and the extent to which any particular household has claims to a particular piece of property. Non-transferable property rights are household characteristics that may contribute to its authority or indicate that the household has greater access to other mechanisms of property rights enforcements. Proxies for these variables used by Lanjouw and Levy include income and assets, education, single-headed female households, and years of residence in the dwelling.

The second argument, made by Jimenez (1984), Friedman, Jimenez and Mayo (1988), is that, due to heterogeneous tastes, households locate in places presenting various levels of risk. These authors run hedonic price regressions for urban areas of the Philippines to assess the premium given to dwelling values by property rights. Relating those to household characteristics,²⁰ they conclude that richer and smaller households are willing to pay more for security than poorer or larger households. Although the set of variables

²⁰Age of the head of household, household size, and a proxy for permanent income.

considered in the two cases are close, the underlying reasons to consider them diverge. Jimenez (1984), Friedman, Jimenez and Mayo (1988) argue that unobservable tastes could be correlated to exogenous variation in community level risk of eviction, whereas Lanjouw and Levy (2002) consider a model where the households characteristics *change* the level of eviction.

The third direction concerns finding variables relating to the neighborhood of dwellings which could reflect the investor’s perception of risk on housing investment. This approach has been suggested by Jimenez (1984), Friedman, Jimenez and Mayo (1988), although they look at housing demand, rather than housing investment. In our model, housing investment *will* vary with the level of risk (which is exogenous), so that house types should differ in areas presenting different levels of risk.²¹ Thus, house characteristics cannot be used as exogenous shifters. We try to use only variables exogenous to the investment problem.

From this brief discussion, it follows that household characteristics might reflect differences in opportunity cost of capital, as well as differences in risk. Thus, they can be expected to affect rates of return in both sectors. On the other hand, exogenous neighborhood characteristics are related only to perception of risk on informal investment; thus, they should explain rates of return in the informal sector only.

We first relate rates of return in both sectors to household characteristics. We run the following regression

$$\ln r = c + \alpha F + Y\beta + \eta, \tag{11}$$

where c is an intercept, F is the dummy variable for the formal sector, and Y is a vector of household characteristics. Results are reported in Table 4. Household variables do not have much explanatory power in the pooled regression. Running separate regressions for the two sectors does not improve the explanatory power. R-squares actually drop as

²¹Another interpretation could be that higher investments are aimed at altering the level of risk for that particular dwelling, because for example local authorities will hesitate more to destroy concrete, multi-storey buildings than shacks.

compared to the pooled regression. In the informal sector, the only variable significant at the 1 percent level is the dummy for born in Pune, which decreases the risk. This variable is not significant in the formal sector regression. Education variables are significant at the 10 percent level and affect rates negatively in the informal sector, whereas they have a positive sign in the formal sector regression (only the dummy for primary education being significant at 10 percent). Log value of per capita consumption also shows the same pattern between the two sectors, although neither of the coefficient are significant. We interpret these results in the following way. On one hand, higher education and income signal greater access to financial markets, and therefore higher opportunity cost of capital. Hence the positive sign in the formal sector regression. On the other hand, these variables also signal higher ability to defend one's rights in the informal sector (non-transferable property rights in the sense of Lanjouw and Levy, 2002). Hence the negative sign in the informal sector regression. Thus, as expected, the same variables may influence rates of return in more than one way.

Next, we relate rates of return in both sectors to neighborhood characteristics that potentially influence perception of risk by investors. We try to use only variables exogenous to the investment problem. In particular we look at sanitation facilities, time to nearest bus stop and accessibility to community tap for water. The reason for including these variables is that these services are provided by local municipalities. Moreover, it must be noted that in Pune none of these facilities are financed by local or property taxes. Provision of these facilities could therefore reflect governments commitments in recognizing informal settlements.

In addition, we also control for the spatial location of the dwelling in the form of the ward where it is located. At the time of the survey the city of Pune was divided into 48 wards (boroughs) serving as electoral districts and also for administrative and representative purposes. Risks to informal investment could differ spatially, either because of commitments of the government (unobserved to the econometrician) in recognizing particular informal settlements, of variations in political influence of informal settlements across

wards, or because land planning and zoning at the city level have different implications for the tolerance of municipal authorities to informal settlements across the city.

In particular, we run the following regression

$$\ln r = c + \alpha F + Z\beta + w_i + \zeta, \quad (12)$$

where c is a constant term, F is a dummy variable for formal sector, Z are variables describing provision of services by local authorities, and w_i are ward dummies. Results are reported in Table 5. In the first column, we do not include ward dummies. We find that inclusion of dummies significantly improves the explanatory power in terms of R -squares. Another interesting feature of this result is that rates of return are 19 percent and 22 percent higher in the informal sector with exclusion and inclusion of dummies for wards respectively. This is similar to what we have found in regression (10). As expected, we find that provision of these services have no explanatory power when we run a separate regression for the formal sector. In the informal sector, point estimates suggests that provision of community tap for water supply raises the rate of return, implying a higher probability of destruction. Therefore, informal settlements which rely on community tap for water are more vulnerable in the sense that they face a higher risk of destruction. An interpretation for this result could be that if the government were committed in recognizing a particular informal settlement, households living in that settlement would rely more on individual connections, or some form of piped water supply.

4.5 Interpreting Results from Previous Literature

The remaining of the empirical part of the paper tries to make the link between our approach and empirical findings of Jimenez (1984), Friedman, Jimenez and Mayo (1988). Taking a demand-side approach in which renters and owners are not distinguished, they run hedonic (log) price and (log) rent regressions for urban areas of the Philippines to assess the premium given to dwelling values by property rights. Jimenez (1984) looking at data from Davao, 1979, found a 18 percent differential for the rents of rental units and

58 percent for the prices of owner-occupied units. In Friedman, Jimenez and Mayo (1988), the corresponding figures are 15 and 25 percent, respectively for Manilla, 1983.²² From our model, these differences between owners and renters are readily explained and can be used to obtain to separate the risks faced by investors from a risk faced by renters. Indeed, suppose that the differences between the formal and informal sectors in log prices and log rents are constant, so that the hedonic rent and price equations in log forms can be written:

$$\begin{cases} \ln R = X\gamma_1 + C_{1i} + (\Delta R_f)F + \varepsilon_1 \\ \ln V = X\gamma_2 + C_{2i} + (\Delta V_f)F + \varepsilon_2 \end{cases} \quad (13)$$

where F denotes the dummy for formal sector, the C 's are the constants for the informal sector in the hedonic regressions, and X is a vector of housing characteristics (which does not include the constant). Thus ΔR_f is the observed rent premium for the formal sector, and ΔV_f is the observed price premium for the formal sector. Differencing the two equations, and using the fact that the rate of return on housing, r , is simply equal to R/V , we get :

$$\ln r = X(\gamma_1 - \gamma_2) + (C_{1i} - C_{2i}) + (\Delta R_f - \Delta V_f)F + \varepsilon_1 - \varepsilon_2 \quad (14)$$

The difference in log rates of return associated to the formal sector is the coefficient of the dummy variable F . This equals

$$\Delta r = \Delta R_f - \Delta V_f \quad (15)$$

It turns out to be simply a "difference in difference" estimator, i.e. the difference between the premiums for the formal sector on (log) prices and (log) rents. We can now give a straightforward interpretation of previous findings. Since risks are higher in the informal sector, returns in this sector should be higher if one is to observe any investment

²²Lanjouw and Levy (2002) report an overall effect of title on properties' value of 23.5 percent but do not give the corresponding figure for rents.

in that sector. Thus the coefficient Δr should be negative. From the expression of Δr , it follows that *the average price premium for the formal sector should be higher than the average rent premium.*

Note, however, that in the context of our model, equation (14) contains endogenous variables on the right-hand side, since investment (which would correspond to the X 's) and rate of return are two endogenous variables, determined by the same exogenous variable, the destruction rate d . Thus, Δr cannot be interpreted as a causal effect of F on r . However, following the same empirical approach for Pune, India, we find that the X 's are very weakly correlated to rates of returns. Besides, the differential between the formal and informal sector estimated this way is $\Delta r = 20$ percent, very close to our previous estimations. Extrapolating to the aforementioned studies, we conclude that this technique may be used to obtain a measure of the investor's premium. Results from estimation of equations (13) are presented in Table 6. Results conform to intuition, and the explanatory power of variables lies in the upper range of what is usually obtained in similar case: the R-square for rents is .62, that for prices is .63.

Most of the explanatory variables have the expected signs. Firstly, both rents and prices are higher in formal category, which is reflected by the positive coefficient on the formal dummy. Residents are willing to pay higher rents if the dwelling unit has more rooms, kitchen, if there is proper sewerage, good toilet facilities, closer to public transport facilities measured in terms of time to nearest bus stop. Moreover, households are willing to pay higher rents for increased hours of water supply irrespective of the source. These results are very similar when we look at sale prices instead of rents. However, it is interesting to note that households' willingness to pay in terms of rents are lower if they have to invest in individual water connection or get an individual sub-connection for water supply. Furthermore, rents are much lower if it relies on community tap as a source for water supply. The results are similar when we look at prices but they are not significant.

Our investment model can be used to estimate a destruction risk for Davao, 1979, and Manilla, 1983. From Jimenez' (1984) results for Davao, the figure corresponding to the

investor's premium would be $58 - 18 = 40$ percent, whereas from Friedman et al. (1988), the risk in Manilla there would be $25 - 15 = 10$ percent. Results from the computation of annual destruction risks implied by the model are gathered in Table 7. From those results, one would conclude that destruction risks were about four times higher in Davao than in Manilla. However, investors' opportunity cost of capital may have been different in the two cities.

4.6 Why are Rents in the Informal Sector so Low ?

Lastly, we come back to the question of the difference in rents in the two sectors. All the studies mentioned above found significant differences in the rents in the two sectors, when observable characteristics of the dwellings are accounted for in a rent hedonic equation. From Table 6, our results show a difference of 35 percent in the rents of similar dwellings in the two sectors. In our theoretical model, equality of rents relies on two assumptions: free mobility of the households, and households having no intrinsic preference for the formal of the informal sector. When the first assumption is relaxed, renters, having mobility costs, may be concerned by the eviction risk. Thus, as in Jimenez (1984), they will be willing to pay less for a unit in the informal sector than in the formal sector. As regards the latter assumption, common sense would suggest that households have a higher valuation for dwellings in the formal sector. However, there is no direct evidence of this, as far as we know. Other factors could account for this difference. The first one is the influence of all variables not observed by the econometrician that may have an impact on rents, such as unobserved quality of dwelling units, access to basic utilities, distance to amenities, transport facilities, etc. If those are on average worse in informal neighborhoods than in formal ones, an hedonic equation including a dummy for informal sector will show a negative coefficient for this variable. A second type of factors relate to the legal, regulatory and judiciary environment of the rental sector. For example, if the legal environment is judged too biased in favor of renters, as appears common in developing countries, landlords in the informal sector will face lower risks and rents will be lower in the informal sector.

Lastly, taxes (for example, tax on rental income, property tax, utility tax, etc.) which typically apply only to formal dwellings may result in equilibrium rent differentials between the two sectors.

5 Discussion: Policy Issues

Our model shows that investments in the formal and informal sectors will be determined mainly by the investors' outside options (opportunity cost of capital), and by the risk on investment in the informal sector. There is little a government can do about the investors' opportunity cost of capital. On the other hand, governments can use the risk on informal investment to influence the attractiveness of the informal sector in one way or another. Recall that δ is the *perceived* risk on informal investment, that is, the risk factor used by investors to discount future income flows.²³ Thus, this factor is driven more by what investors think the government will do, rather than by what the government effectively does. One of the policy options available to the government to change the perception of risk attached to informal land is to make an explicit or implicit commitment not to destroy informal settlements.²⁴ On the other hand, a government can undertake highly visible actions to erase informal settlements, as has been the case in 1999 in Dhaka, Bangladesh (BBC, 1999). By such actions the government can actually raise the perceived risk of investing in informal settlements, whereas the actual risk (as could be measured by the ratio of informal dwellings destroyed in a given year over the total stock) is low. The model highlights the dilemma faced by policy makers in many developing countries. Public commitment not to enforce property rights will foster informal settlements.²⁵ However, such a commitment increases total housing investment, which is a desirable outcome in case of severe housing shortage.

²³This risk incorporates different sources, for example: risk of demolition, other risks stemming from insecurity of tenure (see Lanjouw and Levy, 2002).

²⁴This has been observed in a number of countries, especially in Latin America (Payne, 2002).

²⁵However, another interpretation of our model can be in terms of quality instead of quantity of investment. Following this interpretation, increase in investment security in the informal sector results in increased investment by households in their dwellings, for example, improving walls, floors, adding stories, as observed empirically (see Payne, 2002).

Our model, however, suggests that the best way to reduce investment in the informal sector is to make formal investment more attractive, by lowering the costs associated with it. Direct costs include administrative, time and money costs of land titling and compliance with zoning and building regulations. Casual data from developing countries suggest these costs are significant by any means. Indirect costs include all the legal and regulatory barriers to an efficient land market (the main of which concerns the land titling system), which typically result in high price of formal serviced land. This alternative approach to dealing with informal development has been recognized by many governments in the last decade (see for example Payne, 2002). This paper provides a simple rationale for this type of policy intervention. Figure 3 diagrammatically illustrates the equilibrium changes resulting from reduction of costs in the formal sector. Total quantity and formal quantity of housing increase, whereas quantity of informal housing and equilibrium rent decrease.

However, some caveats apply to these results. Our basic model is designed for a better understanding of risks faced by investors in the formal and informal housing markets. In particular, we do not go into the details of discussing the cost functions in the two sectors. Results from the previous sections are thus valid only in the case of exogenous land prices. This setup may be realistic to describe land markets in developing countries where all the land belongs to the State.²⁶ In this context, the price of land can be thought of as a policy variable, which can be used by the Government to achieve policy objectives. In those countries landless people tend to invade portions of public land and build houses for themselves, beginning with shacks and upgrading their dwellings over time. In this context, there are few reasons to suppose that suppliers act on a competitive basis, or that investments in the two sectors should have equal returns. However, when there exists a market for formal land, investors can invest in both markets, and will look for the sector offering the higher returns. In equilibrium, the price of formal land will be determined

²⁶This was a common case in developing countries in the seventies. In this case, the State sell formal land to developers (often, these were State-owned companies operating non competitively, as in Morocco, Jordan, and many countries of Africa). The price of land sold to developers was set by the State, and examples show that actual prices resulted more from *ad hoc* decisions than economic reasoning.

endogenously and will reflect differences in investment risks between formal and informal sectors. As risk on informal housing investment decreases, informal land becomes a closer substitute to formal land, and the price differential between formal and informal land will decrease.

6 Conclusions

This work was motivated by the simple evidence that in many developing countries construction of informal dwellings, rather than being an exception, is a viable alternative to constructing in the formal sector. Our contribution can be summarized as follows. We show that appropriate data on rents and prices of dwelling units in the formal and informal sectors can be used to decompose the premiums for the formal sector in prices and rents in two parts: the difference in rates of return on housing investment, which reflects the risks borne only by investors, and the formal premium on rents, which reflects conditions prevailing in rental markets and tastes of households. Those two quantities have different economic meanings, and should be analyzed separately. When risk on informal investment is modeled as an exogenous annual destruction rate, estimates of the destruction rates can be obtained very simply from the observed rate premium on informal investment and the investor's opportunity cost of capital.

On the empirical side, we show that rates of return on housing investment are indeed higher in the informal sector than in the formal sector in the city of Pune. The implied risk as perceived by investors in the housing sector is estimated to be equivalent to an annual destruction rate ranging between 1 and 2 percent. The two distributions of rates of returns present highly idiosyncratic components, and traditional explanatory variables proxying either the strength of informal property rights or lower perceived risk of eviction have little power to explain the differences between individual rates of return.

Our approach, by separating (temporal) risks faced by investors from (instantaneous) risks faced by renters and other factors affecting the rental market differently in the for-

mal and informal sectors, sheds light on previous studies, which consistently found that premiums for the formal sector were greater for prices than for rents. We show that the difference between the two premiums precisely reflects the risk faced by investors.

Consistent with previous studies, we find the difference in rents between the formal and informal sectors to be important. Once observed characteristics of the dwellings and of the neighborhood are controlled for, there is still a 35 percent gap between the two sectors. This figure is comparable, although larger, than those found by Jimenez (1984) and Friedman, Jimenez and Mayo (1988) for the Philippines. This difference may be explained only by unobservable characteristics of the dwellings, as well as by the different ways in which housing characteristics interact in the formal and informal sectors. However, since the variables we use in our hedonic regressions are similar to those used in previous papers, one can suppose that this high premium reflects idiosyncratic conditions of the rental market in the informal sector in Pune. Exploiting datasets from other cities in India could give us indications whether this high premium is particular to Pune city, or if it reflects conditions prevailing in the rental markets throughout India. Further empirical studies are needed to better understand what drives those differences across countries or regions.

We believe that our very simple investment model can be extended in many fruitful ways. An obvious extension consists in going beyond our literary discussion in section 4 to consider endogenous land prices. Other insights could be obtained by introducing heterogeneity in the degree of uncertainty on informal investment (that is more categories of informal land), or in the opportunity cost of capital (to reflect the difference between private households with no or limited access to alternative portfolio choices and firms with potentially high returns on financial investments). Consider the latter case. If one finds that firms (supposed to have higher opportunity cost of capital) only invest in the formal sector, this would mean that returns in the informal sector are not high enough to compensate for the higher risk on investment there. Methods similar to ours would thus underestimate the eviction risk.

References

- [1] Banerjee, A. V., P. Gertler, and M. Ghatak, 2002, "Empowerment and Efficiency: The Economics of a Tenancy Reform", *Journal of Political Economy*, 110(2), pp.239-280.
- [2] Besley, T., 1995, "Property Rights and Investment Incentives: Theory and Evidence from Ghana", *Journal of Political Economy*, 103(5), pp.903-937.
- [3] British Broadcasting Corporation, 1999, "World, South Asia: Dhaka court rules on slum evictions", http://bbc.co.uk/2/hi/south_asia/428052.stm.
- [4] Friedman, J., E. Jimenez, and S. K. Mayo, "The Demand for Secure Tenure in Developing Countries", *Journal of Urban Economics* 29, 1988, pp.185-98.
- [5] Hoy, M. and E. Jimenez, "Squatter Rights and Urban Development: An Economic Perspective", *Economica*, 58 (229), February 1991, pp.74-92.
- [6] Jacoby, H., G. Li, and S. Rozelle, 2002, "Hazards of Expropriation: Tenure Insecurity and Investment in Rural China", *American Economic Review*, December 2002, pp.1420-47.
- [7] Jimenez, E., "Tenure Security and Urban Squatting", *Review of Economics and Statistics*, Vol.66, November 1984, pp.556-67.
- [8] Lanjouw, Jean O. and Philip I. Levy. "Untitled: A Study of Formal and Informal Property Rights in Urban Ecuador." *The Economic Journal*, 112 (October), 986-1019.
- [9] Payne, Geoffrey, Ed. 2002. *Land rights and innovation: improving tenure security for the urban poor*. London, ITDG.

7 Appendix

7.1 APPENDIX 1: Comparative Statics

Totally differencing the first order conditions in (4) and equilibrium equation in (7) with respect to β , we get:

$$\begin{aligned} dq_i + dq_f &= D'(R)dR, \\ \frac{Rd\beta}{(1-\beta)^2} + \frac{dR}{(1-\beta)} &= C_f''(q_f)dq_f, \\ \frac{R\delta d\beta}{(1-\beta\delta)^2} + \frac{dR}{(1-\beta\delta)} &= C_i''(q_i)dq_i. \end{aligned} \tag{16}$$

In matrix form, we get:

$$\underbrace{\begin{bmatrix} 1 & 0 & \frac{(-C_i''(q_i))^{-1}}{(1-\beta\delta)} \\ 0 & 1 & \frac{(-C_f''(q_f))^{-1}}{(1-\beta)} \\ 1 & 1 & -D'(R) \end{bmatrix}}_A \begin{bmatrix} dq_i \\ dq_f \\ dR \end{bmatrix} = \begin{bmatrix} \frac{(C_i''(q_i))^{-1}\delta}{(1-\beta\delta)^2} \\ \frac{(C_f''(q_f))^{-1}}{(1-\beta)^2} \\ 0 \end{bmatrix} R d\beta, \tag{17}$$

$$\det(A) = -D'(R) + \frac{(C_i''(q_i))^{-1}}{(1-\beta\delta)} + \frac{(C_f''(q_f))^{-1}}{(1-\beta)} > 0.$$

We have

$$A^{-1} = \frac{1}{\det(A)} \begin{bmatrix} -D'(R) + \frac{(C_f''(q_f))^{-1}}{(1-\beta)} & \frac{(-C_i''(q_i))^{-1}}{(1-\beta\delta)} & \frac{(C_i''(q_i))^{-1}}{(1-\beta\delta)} \\ \frac{(-C_f''(q_f))^{-1}}{(1-\beta)} & -D'(R) + \frac{(C_i''(q_i))^{-1}}{(1-\beta\delta)} & \frac{(C_f''(q_f))^{-1}}{(1-\beta)} \\ -1 & -1 & 1 \end{bmatrix},$$

so that:

$$\begin{aligned}
\frac{dR}{R} &= \frac{d\beta}{\det(A)} \left[-\frac{(C_i''(q_i))^{-1}\delta}{(1-\beta\delta)^2} - \frac{(C_f''(q_f))^{-1}}{(1-\beta)^2} \right] \leq 0, \\
\frac{dq_i}{R} &= \frac{d\beta}{\det(A)} \left[\underbrace{-D'(R)\frac{(C_i''(q_i))^{-1}\delta}{(1-\beta\delta)^2}}_{>0} + \underbrace{\frac{(C_i''(q_i))^{-1}(C_f''(q_f))^{-1}}{(1-\beta\delta)^2(1-\beta)^2}(\delta-1)}_{<0} \right] \leq 0 \quad (18) \\
\frac{dq_f}{R} &= \frac{d\beta}{\det(A)} \left[\underbrace{-D'(R)\frac{(C_f''(q_f))^{-1}}{(1-\beta)^2}}_{>0} + \underbrace{\frac{(C_i''(q_i))^{-1}(C_f''(q_f))^{-1}}{(1-\beta\delta)^2(1-\beta)^2}(1-\delta)}_{>0} \right] \geq 0.
\end{aligned}$$

When d is small ($\delta \approx 1$), the expression for $\frac{dq_i}{R}$ reduces to

$$\frac{dq_i}{R} = \frac{d\beta}{\det(A)} \left[-D'(R)\frac{(C_i''(q_i))^{-1}\delta}{(1-\beta\delta)^2} \right] \geq 0,$$

so the response of informal supply is positive.

Next we do a similar analysis with changes in δ , or perceived risk of destruction. Differencing the first order conditions and equilibrium equations with respect to δ , we get:

$$\begin{aligned}
dq_i + dq_f &= D'(R)dR, \\
\frac{dR}{(1-\beta)} &= C_f''(q_f)dq_f, \\
\frac{R\beta d\delta}{(1-\beta\delta)^2} + \frac{dR}{(1-\beta\delta)} &= C_i''(q_i)dq_i.
\end{aligned} \quad (19)$$

Proceeding as above, we obtain:

$$\begin{aligned}
\frac{dR}{R} &= \frac{d\delta}{\det(A)} \left[-\frac{(C_i''(q_i))^{-1}\beta}{(1-\beta\delta)^2} \right] \leq 0, \\
\frac{dq_i}{R} &= \frac{d\delta}{\det(A)} \left[\underbrace{-D'(R)\frac{(C_i''(q_i))^{-1}\beta}{(1-\beta\delta)^2}}_{>0} + \underbrace{\frac{(C_i''(q_i))^{-1}\beta(C_f''(q_f))^{-1}}{(1-\beta\delta)^2(1-\beta)}}_{>0} \right] \geq 0, \quad (20) \\
\frac{dq_f}{R} &= \frac{d\delta}{\det(A)} \left[-\frac{(C_i''(q_i))^{-1}\beta(C_f''(q_f))^{-1}}{(1-\beta\delta)^2(1-\beta)} \right] \leq 0,
\end{aligned}$$

where $\det(A)$ is the same as before.

7.2 APPENDIX 2: Water Data

All households in our sample have access to one source of water. 94 percent of households have access to "formal" water sources, in the form of individual water connection (55 percent), individual sub-connection (21 percent), or community tap (18 percent). All households in those three categories are provided water by the Pune Municipal Corporation (PMC). The remnant of households (6 percent) access water, either by illegal water connection, by someone else's connection, community wells or hand pumps, etc. Concerning availability of water, the distribution of the hours of availability per day presents two humps (see figure 4), the main one around 6 hours per day and a smaller one around 24 hours per day. Some households having access to formal water supply also have access to a second water source, such as individual hand pumps and tubewells or mini water systems. Most households having a secondary water source report low availability per day of the main source, which suggests a causality effect. Secondary water sources are a means of circumventing this low availability: 83 percent of users of secondary sources report a 24 hours per day availability for this source.

7.3 APPENDIX 3: Regression of rate of return on the subsample of owners, with "title" as the explanatory variable

See Table 8

7.4 APPENDIX 4: Rents and prices hedonic equations for formal and informal sectors

See Tables 9 and 10

Figure 1: Equilibrium determination and effect of increase in risk of housing investment in informal sector

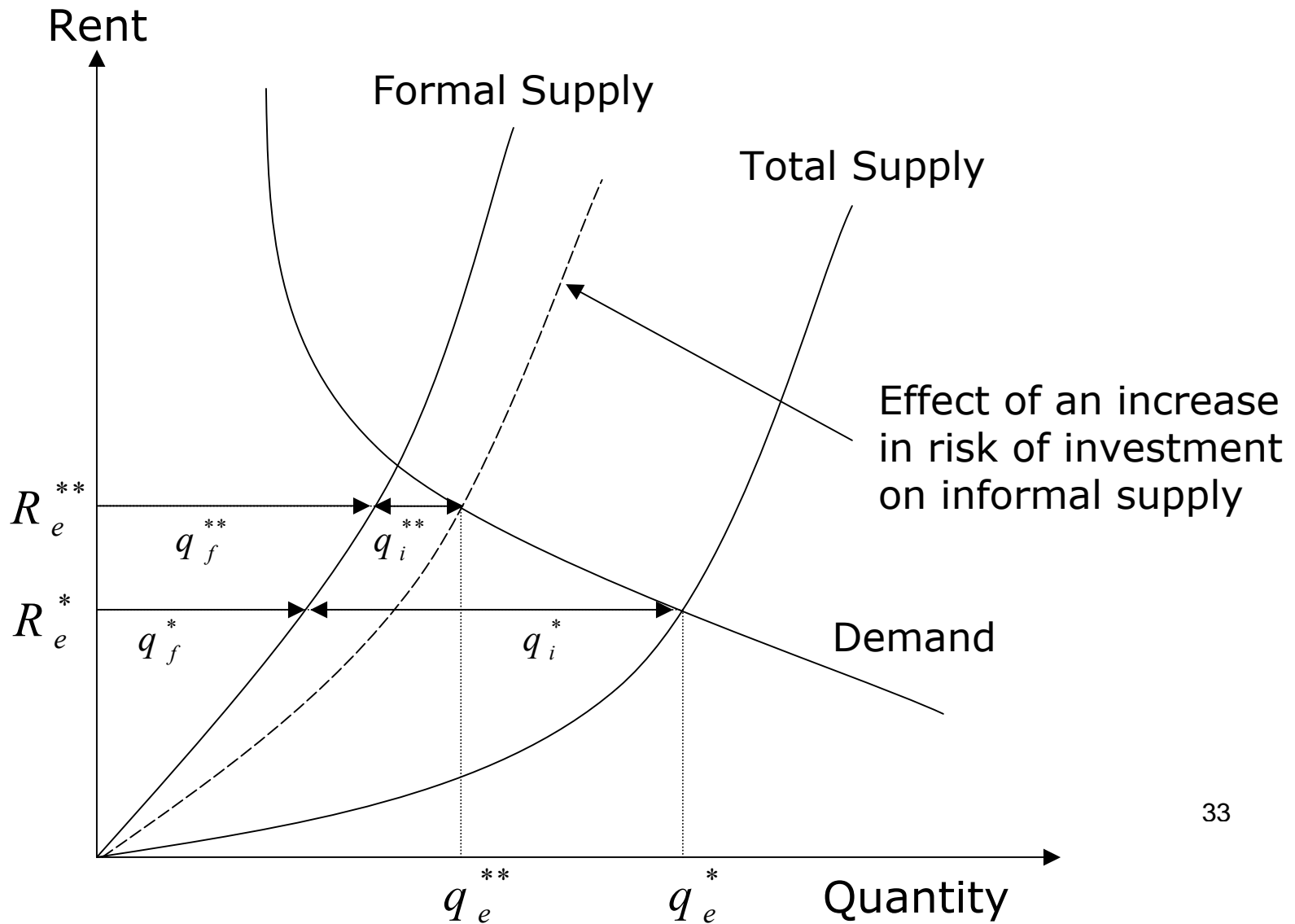


Figure 2: Kernel density of Rate of return for Informal and Formal Category

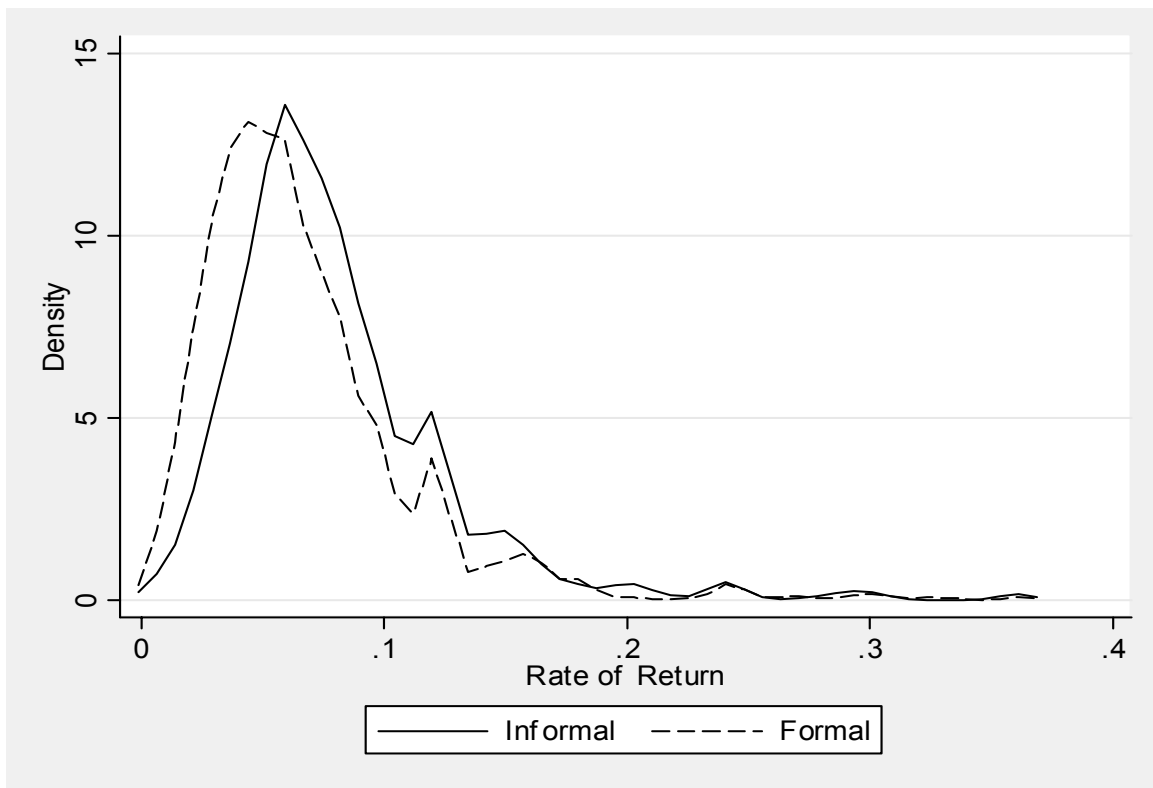


Figure 3: Effect of a decrease in formal production costs

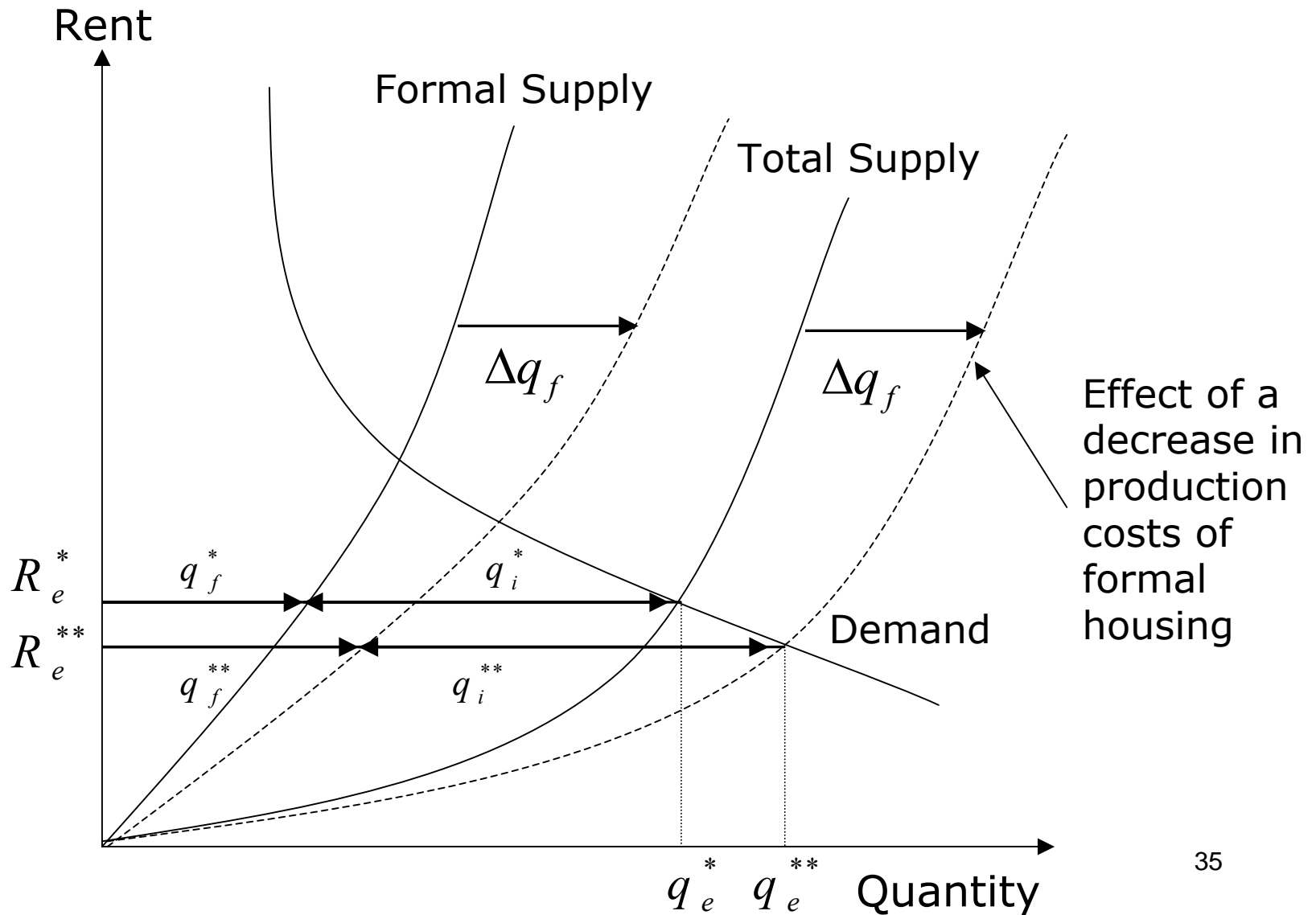


Figure 4: Kernel density of hours per day of water availability

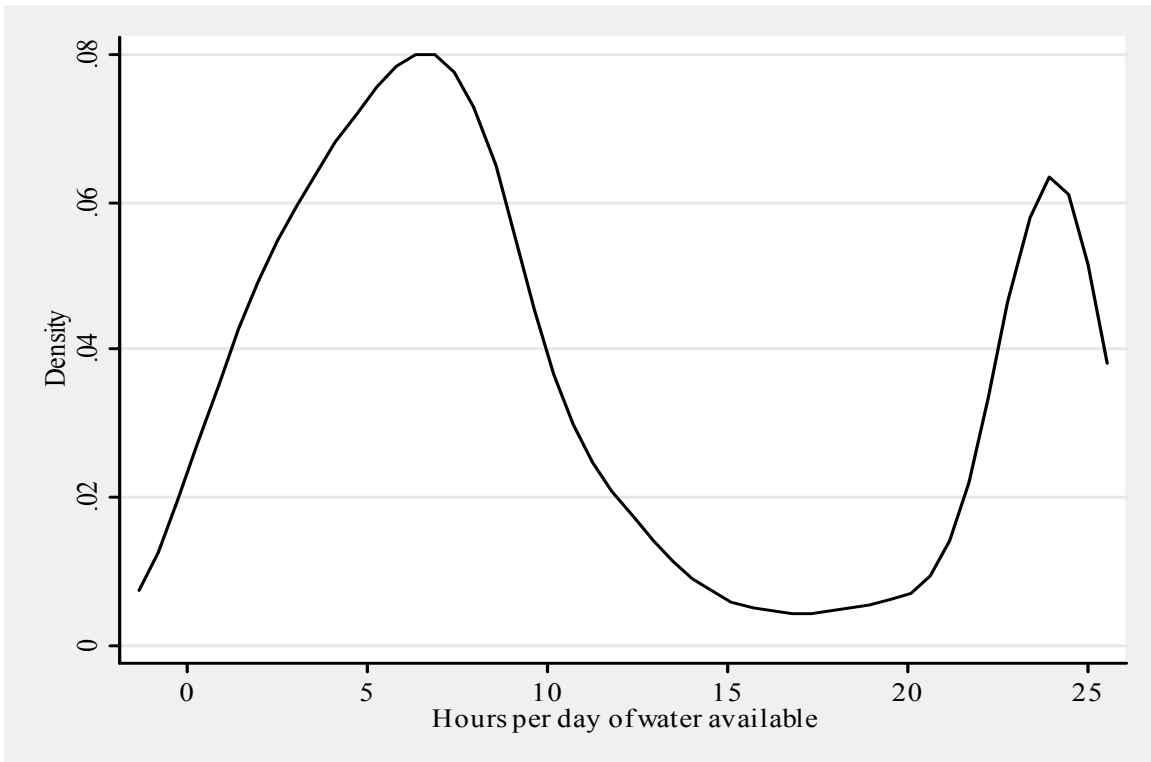


Figure 5: Decomposition of formal/informal premiums on rents and prices

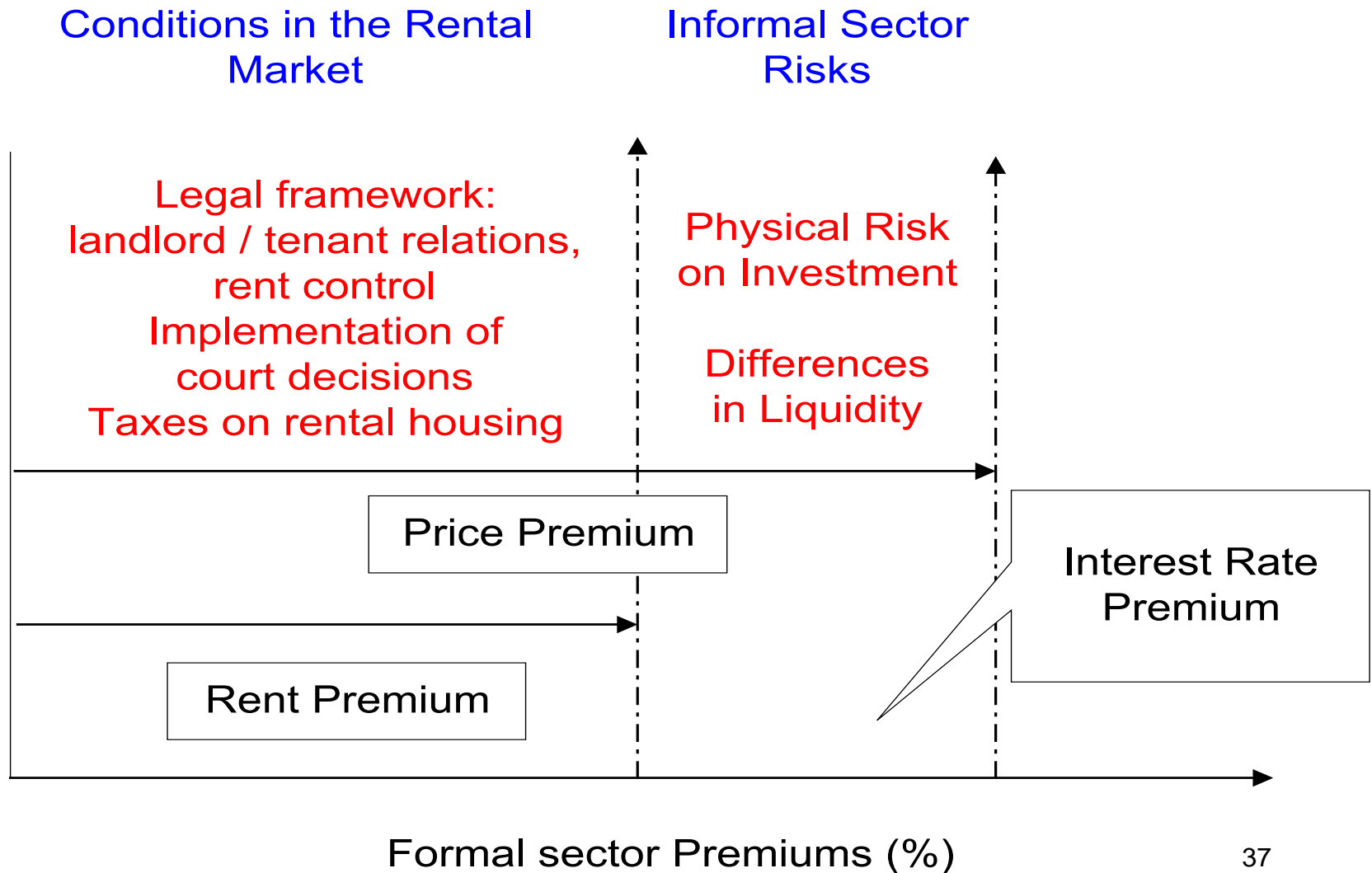


TABLE 1

Details and Summary Statistics of the Variables for Informal and Formal Category

Variable	Informal			Formal		
	Mean	Standard Deviation	Sample Size	Mean	Standard Deviation	Sample Size
A. Household Characteristics						
Hindu	.85	.36	1003	.88	.33	1846
Scheduled Caste	.51	.50	1003	.28	.45	1846
Marathi Speakers	.82	.38	1003	.81	.39	1846
Age of the head of household	46	12.24	1002	50.41	13.24	1846
Number of children in household	1.63	1.44	1003	1.11	1.21	1846
Size of household	5.35	2.05	1003	4.88	2.10	1846
Female headed household	.12	.32	1003	.12	.33	1846
Log value of per capita consumption	9.48	.52	1003	9.98	.62	1846
Households' with head greater than 60	.12	.33	1003	.22	.41	1846
Distance to work (in meters)	5725	8307	731	7104	12432	1269
Time to work (in minutes)	24	17.1	731	25.5	25.14	1269
No education	.27	.45	1003	.11	.31	1846
Primary education	.31	.46	1003	.18	.39	1846
Secondary education	.37	.48	1003	.44	.50	1846
Higher or technical education	.04	.20	1003	.27	.44	1846
Born in Pune city	.63	.48	1003	.70	.46	1846
Years in House	24.13	18.32	1003	30.6	29.5	1846
Ownership of television	.85	.36	1003	.94	.24	1846
Ownership of two wheeler	.26	.44	1003	.54	.50	1846
Ownership of automobile	.04	.19	1003	.13	.33	1846
Ownership of furniture (> Rupees 5000)	.36	.48	1003	.67	.47	1846
B. Dwelling Unit Characteristics						
Quality of roof (0 = bad, 1 = good)	.16	.37	1003	.52	.50	1846
Quality of walls (0 = bad, 1 = good)	.76	.43	1003	.94	.24	1846
Quality of exterior of house (0 = bad, 1 = good)	.25	.43	1003	.49	.50	1846
Total area (in square feet)	351.99	513.06	989	1478.9	2102	1229
Total space of residence (in square feet)	267.50	247.30	1003	504.5	508.34	1846
Number of rooms	1.71	.87	1000	2.32	1.23	1825
Space per room (in square feet)	152	84.40	1000	201.60	116.7	1825
Kitchen (0 = no, 1 = yes)	.27	.45	1003	.58	.49	1846
Quality of toilet (0 = good, 1 = bad)	.85	.36	1003	.41	.49	1846
No sewerage (0 = yes, 1 = no)	.29	.45	1003	.18	.39	1846
Own standing house (0 = no, 1 = yes)	.07	.25	1003	.12	.32	1846
Attached house (0 = no, 1 = yes)	.93	.26	1003	.58	.49	1846
Time to nearest bus stop (in minutes)	9.82	6.33	1001	7.46	4.97	1842
Secondary water connection	.04	.19	1003	.07	.25	1846
Individual water connection	.58	.49	1003	.54	.50	1846
Individual sub-connection for water	.075	.26	1003	.28	.45	1846
Community tap for water	.28	.45	1003	.12	.32	1846
Hours per day of water from individual connection	5.67	7.50	1003	5.03	7.13	1846
Hours per day of water from individual sub-connection	.85	3.82	1003	3.01	6.44	1846
Hours per day of water from community tap	3.52	7.32	1003	1.06	3.82	1846
Ownership of house (0 = no, 1 = yes)	.87	.34	1003	.61	.49	1846
House used as workplace (0 = no, 1 = yes)	.10	.31	1003	.12	.32	1846
Legal title of ownership (0 = no, 1 = yes)	.062	.241	872 ¹	.99	.11	1127 ²
Monthly rent (in rupees)	883	1470	974	2,851	4884	1736
Sale price (in rupees)	1,67,224	378835	946	6,92,364	1260882	1698
Rate of return on investment in housing ³	.098	.23	943	.080	.16	1666

¹ Legal title, only for owners of dwelling units.² Legal title, only for owners of dwelling units.³ Rate of return on investment in housing is calculated as (Monthly rent × 12) / Sale price.

TABLE 2a
Ordinary Least Square (OLS) Regressions of the Rate of Return

Variable	Regression		
	(1)	(2)	(3)
Formal Dummy	-0.014 (0.002)**	-0.016 (0.002)**	-0.015 (0.005)**
Own House Dummy		-0.009 (0.002)**	-0.007 (0.004)
Own House * Formal			-0.002 (0.005)
Intercept	0.081 (0.001)**	0.088 (0.002)**	0.087 (0.004)**
Observations	2562	2562	2562
R-squared	0.02	0.03	0.03

NOTE. - The dependent variable is the rate of return on housing. In the empirical analysis we dropped observations for which rate of return is greater than 0.4 or less than 0.05. In total 47 observations were dropped.

Standard errors are in parenthesis. * significant at 5%; ** significant at 1%

TABLE 2b
OLS Regression of Log Value of Rate of Return

Variable	Regression		
	(1)	(2)	(3)
Formal Dummy	-0.237 (0.024)**	-0.257 (0.025)**	-0.204 (0.060)**
Own House Dummy		-0.082 (0.027)**	-0.032 (0.059)
Own House * Formal			-0.064 (0.066)
Intercept	-2.656 (0.020)**	-2.584 (0.031)**	-2.628 (0.055)**
Observations	2562	2562	2562
R-squared	0.04	0.04	0.04

NOTE. - The dependent variable is the log value of rate of return on housing. In the empirical analysis we dropped observations for which rate of return is greater than 0.4 or less than 0.05. In total 47 observations were dropped.

Standard errors are in parenthesis. * significant at 5%; ** significant at 1%

TABLE 3

Opportunity cost of capital, Investment horizon, and Implied destruction rate (%)

Investor's Opportunity cost of capital	Implied annual destruction rate (%)		
	Horizon = 20 years	Horizon = 50 years	Infinite Horizon
0.01	2.2	0.9	0.2
0.03	2.4	1.1	0.7
0.04	2.5	1.3	0.9
0.05	2.6	1.4	1.1
0.06	2.7	1.5	1.3
0.07	2.8	1.7	1.5
0.08	3.0	1.9	1.8
0.10	3.2	2.3	2.2

TABLE 4
OLS Regression Explaining the Rates of Return: Household Characteristics

Variable	Regression		
	All Households	Informal	Formal
Formal	-0.183 (0.030)**		
Own House	-0.059 (0.028)*	-0.025 (0.055)	-0.07 (0.034)*
Hindu	-0.067 (0.040)	-0.024 (0.059)	-0.095 (0.054)
Scheduled caste	0.036 (0.026)	0.008 (0.037)	0.051 (0.036)
Marathi speaker	0.024 (0.0350)	-0.023 (0.0540)	0.053 (0.0460)
Age of head of household	-0.002 (0.001)*	-0.002 (0.002)	-0.002 (0.001)
Number of Children	-0.001 (0.013)	-0.019 (0.018)	0.009 (0.017)
Size of Household	0.003 (0.009)	0.01 (0.014)	-0.001 (0.011)
Female headed household	-0.03 (0.039)	-0.015 (0.058)	-0.031 (0.051)
Log value of per capita consumption	0.013 (0.031)	-0.059 (0.049)	0.055 (0.040)
Primary education	-0.015 (0.039)	-0.089 (0.048)	0.105 (0.062)
Secondary education	-0.066 (0.038)	-0.104 (0.050)*	0.015 (0.058)
Higher or technical education	-0.055 (0.049)	-0.033 (0.098)	0.013 (0.066)
Years in house	-0.003 (0.0001)**	-0.001 (0.001)	-0.003 (0.001)**
Born in Pune city	-0.022 (0.026)	-0.123 (0.038)**	0.044 (0.036)
Ownership of television	-0.064 (0.043)	-0.081 (0.053)	-0.049 (0.067)
Ownership of two wheeler	-0.064 (0.029)*	-0.001 (0.046)	-0.095 (0.037)*
Ownership of automobile	-0.084 (0.046)	-0.157 (0.099)	-0.089 (0.054)
Ownership of Furniture (> Rupees 5000)	0.036 (0.028)	0.059 (0.041)	0.024 (0.037)
Intercept	-2.445 (0.306)**	-1.745 (0.487)**	-3.118 (0.404)**
Observations	2561	928	1633
R-squared	0.06	0.04	0.04

NOTE. - The dependent variable is the log value of rate of return. In the empirical analysis we dropped observations for which rents or sale price were missing. Furthermore, we have trimmed the data by dropping observations for which rate of return is greater than 0.4 and less than .005. Standard errors are in parenthesis. * significant at 5%; ** significant at 1%

TABLE 5
 OLS Regression Explaining the Rates of Return: Exogenous Characteristics

Variable	Regression			
	All Households		Informal	Formal
Formal	-0.227 (0.024)**	-0.193 (0.028)**		
No sewerage	0.02 (0.028)	-0.002 (0.032)	-0.006 (0.047)	0.013 (0.046)
Time to nearest bus stop	0 (0.002)	0 (0.002)	-0.004 (0.003)	0.001 (0.004)
Community tap for water	0.057 (0.033)	0.08 (0.034)*	0.119 (0.043)**	0.054 (0.054)
Intercept	-2.676 (0.028)**	-2.708 (0.094)**		
Ward Dummies	No	Yes	Yes	Yes
Observations	2556	2556	927	1629
R-squared	0.04	0.09	0.1	0.08

NOTE. - The dependent variable is the log value of rate of return. In the empirical analysis we dropped observations for which rents or sale price were missing. Furthermore, we have trimmed the data by dropping observations for which rate of return is greater than 0.4 and less than .005. Standard errors are in parenthesis. * significant at 5%; ** significant at 1%

TABLE 6
OLS (hedonic) Regression for Rents and Prices

Variable	Regression	
	Rent ^a	Price ^b
Formal	0.367 (0.031)**	0.566 (0.037)**
Quality of roof	0.057 (0.036)	0.037 (0.043)
Quality of walls	0.086 (0.040)*	0.17 (0.049)**
Quality of exterior	0.299 (0.032)**	0.256 (0.034)**
Total space of residence	0.0001 (0.0001)	0 (0.0001)
Number of rooms	0.225 (0.030)**	0.315 (0.035)**
Space per room	0.001 (0.000)**	0.001 (0.000)**
Kitchen	0.199 (0.035)**	0.246 (0.040)**
House used as a workplace	0.197 (0.044)**	0.198 (0.050)**
Quality of toilet	-0.23 (0.040)**	-0.205 (0.044)**
No sewerage	-0.147 (0.031)**	-0.17 (0.033)**
Own standing house	0.021 (0.057)	0.075 (0.066)
Attached house	-0.123 (0.041)**	-0.107 (0.049)*
Time to nearest bus stop	-0.009 (0.002)**	-0.008 (0.003)**
Secondary water connection	0.1 (0.045)*	0.176 (0.057)**
Individual water connection	-0.112 (0.043)**	-0.001 (0.053)
Individual sub-connection for water	-0.114 (0.056)*	-0.05 (0.067)
Community tap for water	-0.215 (0.058)**	-0.131 (0.068)
Hours per day of water from individual connection	0.011 (0.002)**	0.015 (0.003)**
Hours per day of water from individual sub-connection	0.012 (0.003)**	0.016 (0.004)**
Hours per day of water from Community tap	0.009 (0.003)**	0.01 (0.004)*
Intercept	6.134 (0.097)**	10.851 (0.119)**
Observations	2535	2535
R-squared	0.62	0.63

NOTE. - a) The dependent variable is the log value of rents b) the dependent variables is the log value of prices. In the empirical analysis we dropped those observations for which rents or sale price were missing. Furthermore, we have trimmed the data by dropping observations for which rate of return is greater than 0.4 and less than .005.

Standard errors are in parenthesis. * significant at 5%; ** significant at 1%

TABLE 7

Opportunity cost of capital, Investment horizon, and Implied annual destruction rate (%) for Manilla, 1983 and Davao, 1979

Investor's Opportunity cost of capital	Implied annual destruction rate (%)					
	Davao (Log Premium = 0.4)			Manilla (Log Premium = 0.1)		
	Horizon = 20 years	50 years	Infinite Horizon	20 years	50 years	Infinite Horizon
0.01	4.7	1.7	0.5	1.0	0.28	0.10
0.03	5.0	2.4	1.5	1.2	0.32	0.20
0.04	5.2	2.7	2.0	1.2	0.35	0.28
0.05	5.5	3.0	2.5	1.3	0.40	0.30
0.06	5.7	3.4	3.0	1.4	0.43	0.36
0.07	6.0	3.7	3.5	1.5	0.46	0.42

TABLE 8
 Ordinary Least Square (OLS) Regressions of the Rate of Return

Variable	Regression	
	(1) ^a	(2) ^b
Legal Title Dummy	-0.017 (0.002)**	-0.278 (0.027)**
Intercept	0.081 (0.001)**	-2.648 (0.020)**
Observations	1849	1849
R-squared	0.04	0.06

NOTE. - a) The dependent variable is the rate of return on housing, b) the dependent variable is the log value of rate of return. In the empirical analysis we dropped observations for which rate of return is greater than 0.4 or less than 0.05. In total 47 observations were dropped. Standard errors are in parenthesis. * significant at 5%; ** significant at 1%

TABLE 9
OLS Regression Explaining Rents: Dwelling unit characteristics

Variable	Regression		
	All Households	Informal	Formal
Formal	0.367 (0.031)**		
Quality of roof	0.057 (0.036)	0.202 (0.049)**	0.01 (0.047)
Quality of walls	0.086 (0.040)*	0.153 (0.040)**	0.04 (0.084)
Quality of exterior	0.299 (0.032)**	0.18 (0.045)**	0.352 (0.041)**
Total space of residence	0.0001 (0.0001)	0.0001 (0.0002)	0.0001 (0.0001)
Number of rooms	0.225 (0.030)**	0.247 (0.048)**	0.231 (0.036)**
Space per room	0.001 (0.000)**	0.001 (0.001)*	0.001 (0.000)*
Kitchen	0.199 (0.035)**	0.122 (0.046)**	0.234 (0.048)**
House used as a workplace	0.197 (0.044)**	0.13 (0.063)*	0.227 (0.057)**
Quality of toilet	-0.23 (0.040)**	-0.261 (0.061)**	-0.213 (0.050)**
No sewerage	-0.147 (0.031)**	-0.091 (0.040)*	-0.166 (0.045)**
Own standing house	0.021 (0.057)	-0.737 (0.510)	0.056 (0.067)
Attached house	-0.123 (0.041)**	-0.764 (0.503)	-0.112 (0.046)*
Time to nearest bus stop	-0.009 (0.002)**	0 -0.003	-0.015 (0.003)**
Secondary water connection	0.1 (0.045)*	0.185 (0.097)	0.073 (0.051)
Individual water connection	-0.112 (0.043)**	-0.12 (0.063)	-0.087 (0.056)
Individual sub-connection for water	-0.114 (0.056)*	-0.049 (0.120)	-0.063 (0.065)
Community tap for water	-0.215 (0.058)**	-0.187 (0.072)*	-0.281 (0.090)**
Hours per day of water from individual connection	0.011 (0.002)**	-0.001 (0.003)	0.018 (0.003)**
Hours per day of water from individual sub-connection	0.012 (0.003)**	0.006 (0.011)	0.012 (0.003)**
Hours per day of water from Community tap	0.009 (0.003)**	0 (0.003)	0.019 (0.008)*
Intercept	6.134 (0.097)**	6.786 (0.511)**	6.462 (0.144)**
Observations	2535	924	1611
R-squared	0.62	0.44	0.55

NOTE. - The dependent variable is the log value of rents. In the empirical analysis we dropped observations for which rents or sale price were missing. Furthermore, we have trimmed the data by dropping observations for which rate of return is greater than 0.4 and less than .005. Standard errors are in parenthesis. * significant at 5%; ** significant at 1%

TABLE 10
OLS Regression Explaining Prices: Dwelling unit characteristics

Variable	Regression		
	All Households	Informal	Formal
Formal	0.566 (0.037)**		
Quality of roof	0.037 (0.043)	0.172 (0.059)**	0.004 (0.056)
Quality of walls	0.17 (0.049)**	0.35 (0.051)**	-0.076 (0.101)
Quality of exterior	0.256 (0.034)**	0.206 (0.045)**	0.275 (0.046)**
Total space of residence	0 (0.0001)	-0.0003 (0.0003)	0 (0.0002)
Number of rooms	0.315 (0.035)**	0.285 (0.055)**	0.323 (0.045)**
Space per room	0.001 (0.000)**	0.002 (0.001)**	0.001 (0.001)*
Kitchen	0.246 (0.040)**	0.174 (0.053)**	0.28 (0.054)**
House used as a workplace	0.198 (0.050)**	0.11 (0.071)	0.226 (0.065)**
Quality of toilet	-0.205 (0.044)**	-0.31 (0.065)**	-0.174 (0.054)**
No sewerage	-0.17 (0.033)**	-0.093 (0.046)*	-0.193 (0.047)**
Own standing house	0.075 (0.066)	-1.035 (0.621)	0.1 (0.074)
Attached house	-0.107 (0.049)*	-1.048 (0.614)	-0.126 (0.053)*
Time to nearest bus stop	-0.008 (0.003)**	0 (0.003)	-0.014 (0.004)**
Secondary water connection	0.176 (0.057)**	0.266 (0.120)*	0.149 (0.064)*
Individual water connection	-0.001 (0.053)	-0.031 (0.080)	0.015 (0.069)
Individual sub-connection for water	-0.05 (0.067)	-0.02 (0.151)	0.014 (0.077)
Community tap for water	-0.131 (0.068)	-0.17 (0.090)	-0.173 (0.108)
Hours per day of water from individual connection	0.015 (0.003)**	0.003 (0.003)	0.023 (0.004)**
Hours per day of water from individual sub-connection	0.016 (0.004)**	0.015 (0.009)	0.015 (0.004)**
Hours per day of water from Community tap	0.01 (0.004)*	0.003 (0.004)	0.02 (0.0103)
Intercept	10.851 (0.119)**	11.799 (0.623)**	11.622 (0.177)**
Observations	2535	924	1611
R-squared	0.63	0.5	0.52

NOTE. - The dependent variable is the log value of prices. In the empirical analysis we have dropped observations for which rents or sale price were missing. Furthermore, we have trimmed the data by dropping observations for which rate of return is greater than 0.4 and less than .005. Standard errors are in parenthesis. * significant at 5%; ** significant at 1%