Discussion Paper No. 990

WHAT ACCOUNT FOR THE DIFFERENCES IN RENT-PRICE RATIO AND TURNOVER RATE? A SEARCH-AND-MATCHING APPROACH

Daisy J. Huang
Charles Ka Yui Leung
Chung-Yi Tse

February 2017

The Institute of Social and Economic Research
Osaka University
6-1 Mihogaoka, Ibaraki, Osaka 567-0047, Japan

Title: What account for the differences in rent-price ratio and turnover rate? A search-and-matching approach

Author: Daisy J. Huang, Charles Ka Yui Leung and Chung-Yi Tse

This version: February, 2017

Abstract

We build an on-the-house-search model and show analytically that the rent-to-price ratio (or rental yield) and turnover rate, which are frequently used metrics for the housing market, are jointly determined in equilibrium. We therefore adopt a simultaneous equation approach on matched sale-rental pairs in our empirical investigation, as a housing unit cannot be owner-occupied and renter-occupied at the same time. Our empirical results confirm a higher turnover rate is associated with a lower rent-to-price ratio, as predicted by the model. Furthermore, our results suggest a form of "dichotomy" in the empirical determinants of rental yield and turnover at the real-estate-development (RED) level: the demographic structure, and past return performance affect its turnover rate, while popularity, human capital environment, mortgage burden, and long run rent growth determine the rental yield. No evidence of "thick market effect" is found. The robustness of our results are established through a series of tests. In addition to these findings, our tractable search-theoretic model, a ranking of more than 130 RED in Hong Kong based on the popularity index we construct, and the estimated brand-premium for different major real estate developers may also carry independent research and practical interests.

Keywords: housing rental yield, turnover rate, bootstrap, leave-one-out cross validation (LOOCV), matching estimator, 3 stages least squares (3SLS)

JEL Classification Numbers: D10, G10, R21

Acknowledgement: We are grateful to Been-Lon Chen, Nan-Kuang Chen, Man Cho, Gangzhi Fan, Shihe Fu, Richard Green, Cheuk Yin Ho, Mingming Jiang, Kamhon Kan, Yuichiro Kawaguchi, Shian-Yu Liao, Steve Malpezzi, Masao Ogaki, Tim Riddiough, Jay Sa-Aadu, Jim Shilling, Jingjing Tang, Wei Xiao, Haijian Zeng, seminar participants at the Asia-Pacific Real Estate Research Symposium (2015, Cheng Du), Global Real Estate Submit (2015, Washington, D.C.), Taiwan Economic Association meeting (2016, Taipei), City University of Hong Kong, Guangxi University, anonymous referees for helpful comments. Financial support from the City University of Hong Kong and University of Hong Kong are gratefully acknowledged. Part of this research is conducted when Leung is a visiting scholar of ISER, Osaka University, whose hospitality is gratefully acknowledged. Tommy Chan provides capable research assistance. The usual disclaimer applies.

Correspondence: Huang: Institute of Economics and Finance, Nanjing Audit University, No.86 Yushan Rd(W), Pukou, Nanjing, P.R.China, 211815, daisy@nau.edu.cn ; Leung, Department of Economics and Finance, City University of Hong Kong, Kowloon Tong, Hong Kong, kycleung@cityu.edu.hk ; Tse, School of Economics and Finance, University of Hong Kong, Pokfulam, Hong Kong, tsechung@econ.hku.hk .

1 Introduction

Rent-to-price ratio and turnover are both key elements in discussions of the housing market. One example is how the rent-to-price ratio, also known as the (gross) rental yield, is often used in the media and in the academic literature as a metric of the possibility that a "bubble" has formed.[1] Similarly, turnover, also known as the trading or transaction volume, is a commonly used metric for the "status" of the housing market.[2] However, the *equilibrium determination* of rent-to-price ratio is under-explored in empirical works and policy discussion, making it difficult to assess whether the observed rent-to-price ratio and turnover rate are "reasonable." To put it differently, what levels of the rent-to-price ratio and turnover rate are more likely to be associated with housing market mispricing?[3] This study uses a straightforward method to answer this question. First, we propose a simple equilibrium search-theoretic model, in which the rent-to-price ratio and the turnover rate are endogenously determined.

[1] Examples of using rent-to-price ratio as a measurement of the housing market abound in the media. For instance, in a widely used website, Global Property Guide (2010) states that "In the stock market a very popular rule of thumb involves the price-earnings ratio, which measures how high the company's net earnings are, in relation to the price of the stock.... It's the same in the housing market. What's generally viewed as reasonable is similar to what's considered reasonable in the stock market, although houses tend to be expected to yield slightly less, perhaps because a house's value depreciates less over time than the assets of a typical company. *The price/rent ratio (or gross rental yield) is the housing parallel to the price/earnings ratio*." (Italic added).

The rent-to-price ratio is also widely used in the academic literature. For example, in their investigation of possible mis-pricing of housing markets in China, Wu et al. (2012, p.533) states that "...The data show that *price-to-rent ratios* not only are high in these places, but they have increased sharply in the past few years. The *price-to-rent ratio* in Beijing increased by almost three-quarters just in the last three years, rising from 26.4 in 2007(1) to 45.9 in 2010(1). Hangzhou, Shanghai, and Shenzhen also have seen their *price-to-rent ratios* rise sharply to over 40.... Given what we think are reasonable assumptions about the other parameters determining user costs, it appears that home buyers are assuming quite large capital gains on their homes.... However, home prices do not always rise and certainly not consistently at the high rates recently experienced in China... one can conclude that house prices in major Chinese markets are quite risky, even though the data are too limited to make a blanket claim of mis-pricing." (Italic added). Similarly, in a research report by the staff of the Reserve Bank of Australia, Fox and Tulip (2014) also use a constant-quality measure of rental yield to measure the degree of over-valuation.

More recently, Green (2015) acknowledges that in the evaluation of whether a "bubble" exists or not, it is not enough to consider the price only; the rent should also be taken into consideration.

[2] For example, Financial Times (2015) reports that as the global transaction volume of commercial real estate hits a record in the last quarter of 2014, "fear of bubble" arises among analysts. Similar statements can be found in the academic literature as well. For example, in their survey of different classes of "bubble models", Brunnermeier and Oehmke (2013) write that "... heterogeneous-belief models generate the prediction that bubbles are associated with high *trading volume*, something that is often observed in practice.... " (Italic added). In the book coauthored by S. Gjerstad and Nobel laureate Vermon Smith (2014, p.21), it is written that "...before an asset bubble collapses, *trading volume* typically declines substantially. This occurred in the housing market..." (Italic added).

[3] Clearly, there are exceptions, such as Kashiwagi (2014). More discussion is to be followed.

We find, unsurprisingly, that the two variables are jointly determined at the equilibrium. This motivates us to take a simultaneous equations approach to identify the empirical determinants of both the rent-to-price ratio and turnover rate, taking into considerations that the two variables would interact in equilibrium. Our empirical framework can hence provide a benchmark for an estimate of "reasonable" rent-to-price ratio and turnover rate given the empirical determinants, which in turn constitute a yardstick for potential housing market mispricing. The details will be provided in later sections.

The search-theoretic approach to the study of real estate markets and urban economic issues has recently been the subject of much research (for example, Anglin, 2006; Anglin and Gao, 2011; Albrecht et al., 2007; Diaz and Jerez, 2013; Fisher et al., 2003; Genesove and Han, 2012; Krainer, 2001; Novy-Marx, 2009; Piazzesi and Schneider, 2009; Tse, 2011; Wheaton, 1990; Yavas, 1992a, 1992b). Many studies have examined equilibrium transaction price and trading volume, but the rental segment of the residential real estate market appears to be under-explored.[4] We take an initial step in this regard by developing a search and matching model that enables landlords to choose whether to sell or to rent their properties. Thus, the equilibrium rental yield and the turnover can then be derived endogenously. Changes in real estate development (RED) level variables are demonstrated to induce a negative correlation between rental yield and normed transaction. Our simple framework is tractable and can be further extended for other research projects. Hence, it may carry independent research interest.

Our second contribution is to propose a simultaneous equations system to account for endogenously correlated rental yield and turnover rate, instead of the usual ordinary least squares estimation.[5] We employ the Hong Kong data for our empirical testing based on the following reasons. First, Hong Kong provides a housing market in which the owner-occupied and rental segments are "integrated," in the sense that the same unit can be used for both types of housing. Owners can hence arbitrage for different opportunities and so the rent-to-price ratio is maintained at a level that reflects a "no-arbitrage" situation. The population density of Hong Kong is high and most people live in units in high-rise condominiums, which are for sale and for rental.[6] In contrast, the for-sale housing in the United States is typically detached houses, while rental housing is concentrated in the apartment buildings, many of which are owned by institutional investors. Thus, the two housing markets are virtually segregated, hence limiting the arbitrage opportunities for individual house owners. The standard rent-to-price ratio may therefore *not* be as informative (Glaeser and Gy-

---

[4] Xiao (2014), Sato and Xiao (2015) explore the relationship between house rent and labor market. Their models do not have housing price though.

[5] More specifically, we implement three-stage least squares estimation on a two equation system with rental yield and turnover rate treated as endogenous variables and on the same footing.

[6] Among others, see Leung and Tang (2015) for more discussion on the Hong Kong housing market.

ourko, 2007, Malpezzi, 2014, among others). We have access to RED level data in Hong Kong, which helps us to mitigate the potential aggregation bias.[7] More specifically, our sample covers 130 major private REDs in Hong Kong during the period from November 2011 to October 2012. This choice of sampling period enables us to include corresponding information about demographics, household income, and housing related expenditure from the official census.[8] The rent-to-price ratio and normed transactions are constructed out of the micro-transaction data. We also construct a "popularity" index based on a formative measurement model, and attempt to make it as composite as possible.[9] This index may therefore independently be of interest for academic research and practical business investment.

In addition to constructing the popularity index, we have several other major findings. First, our three-stage least squares (3SLS) estimation confirms our theoretical model and suggests *significantly negative effects of turnover and popularity on the rent-price ratio.* If the normalized transaction increases by one standard deviation, other things being equal, the associated rental yield would be 0.48 standard deviations lower. If the popularity index increases by one standard deviation, its rent-to-price ratio would be lower by almost 0.16 standard deviations, holding other variables constant. On the other hand, neither the estimate coefficient of the popularity index nor that of the rent-to-price ratio of the RED, is statistically significant in explaining the turnover rate.[10] Second, we *identify some (potential) empirical determinants of* the rent-to-price ratio and the turnover rate, controlling for the popularity index and the endogenous variables. As a demand shifter, increases in income should stimulate sales volume, but a higher income also means a higher opportunity cost of moving. Ex ante, it is not clear which will be the dominant force, but our empirical finding suggests it will be the latter. Third, we identify a form of *empirical "dichotomy" in the housing market.* The demographic structure, and past return performance of a real estate development tend to affect its turnover rate, while popularity, human capital environment, mortgage burden, and long-run rent growth have more influence on rental yield. Last, we obtain *direct evidence on the brand value ranking of nine main developers in Hong Kong* residential property development market. We also find *occupational exposure to real estate* has significant and positive impacts on the valuation ratio of rental yield, but not on trading intensity, which is consistent with the confirmed moral hazard effect in real estate brokerage.[11]

---

[7] See Hanushek et al. (1996) for more discussion, among others.

[8] The Census and Statistics Department of the Hong Kong government has published, the first time in Hong Kong history, RED level statistics for free public access after the 2011 Census.

[9] It takes into account all factors that are accessible to us and potentially relevant in housing purchase decision, such as affordability, location appeal, access to public transportation, developer reputation, architectural design, surrounding environment, school zone, building age, facility, etc.

[10] On the other hand, we do identify other variables that are significant in explaining the RED level turnover rate. Please see below.

[11] It is beyond the scope of this paper to discuss this literature. Among others, see Levitt and Syverson (2008), Rutherford et al. (2005).

Our paper can also be broadly considered as a study of the asset market liquidity (e.g. Amihud et al. 2013), as real estate is an asset traded with significant frictions. There are recent studies of the rent-to-price ratio that also use a search-theoretic framework. For instance, Kashiwagi (2014) considers a search theoretic model with two sub-markets ("locations"). Homeowners in a location receiving "separation shocks" are forced to sell their houses immediately, move to another location and become renters in the next period. Renters who do not receive separation shocks would attempt to purchase a house. In that sense, Kashiwagi's model may be closer to the *inter-regional mobility* discussed in the literature, i.e. a situation when economic agents are relocated among different states or cities that are far from each other. Hence, it is natural for Kashiwagi (2014) to calibrate his model to match the aggregate data of the US. In terms of modelling strategy, the assumption that homeowners need to sell their houses immediately and move to another location is analogous to the "job destruction model" developed in the seminal work of Mortensen and Pissarides (1994), where workers *need to* search for new jobs once they receive the separation shocks, because their jobs are "destroyed." Kashiwage also proves the existence and uniqueness of a symmetric steady state. Our model instead assumes that mismatched households can stay in their original homes until they find new ones, as houses are not "destroyed" after the separation shocks. It may therefore be applicable to the situation of *intra-regional mobility* or *intra-metropolitan mobility* discussed in the literature, which is also consistent with our empirical implementation of using micro-data from a single city.[12] In terms of modelling strategy, as agents in our model are allowed to stay in the houses even after receiving separation shock, our model has some parallel to the "on-the-job search" model in the labor market search literature. In those models, workers who are not satisfied with the current job-wage package would nevertheless stay in the current jobs until they find "better ones" (Burdett and Mortensen, 1998; Pissarides 1994). We also prove that a unique steady state exists in that environment. Halket and Pignatti (2013) also study the rent-to-price ratio in a search-theoretic housing market model. Their focus is whether rental housing is available in the local market and how the rent-to-price ratio will relate to the homeownership rate in equilibrium. They do not elaborate the relationship between the rent-to-price ratio and trading volume, which is one of the focuses of this paper. They also propose a competitive search model (CSM) while we adopt a random search model (RSM),[13] and it is well known that CSM and RSM can deliver very different conclusions on similar issues.[14] The data sets in these previous studies

[12] The literature on intra-regional or intra-metropolitan mobility has a long history. Among others, see Quigley and Weinberg (1977) for a review of the earlier literature, and Zax (1994) for a recent analysis. See also Green and Malpezzi (2003).

[13] For another example of CSM in real estate, see Leung and Zhang (2011), among others.

[14] For instance, see Moen and Rosen (2004), and the reference therein.

In the context of a housing market, a CSM envisions the situation where there are many different prices offered for identical for-sale units, and many different rents offered for identical rental units. It follows that there will be a distribution of rent-price ratios at the equilibrium. It may create are from the United States, while ours is from an Asian city, and we adopt a 3SLS estimation methodology. Our work should therefore be considered complementary to that in these other studies.

Recently, the liquidity of the housing market is studied empirically. For instance, Kwok and Tse (2006) investigate the price premia for easy-to-sell condominiums in one administrative district of Hong Kong in 2005. Using data from 24 major office markets in the U.S. from 1995 to 2010, Liu and Qian (2012) find a bust period driven illiquidity premium in the real estate expected returns. We complement these studies by adopting the simultaneous equations approach and by combining housing market data with household data (in this case, the 2011 Census of Hong Kong). The trade-off is that we are unable to examine the time series variations of the rent-to-price ratio, as the 2011 Census is a cross-sectional survey by nature. We leave this to future research.

Our findings are also connected to the literature on real estate capitalization rates, which focuses on four categories of cap rate determinants: macroeconomy (Ambrose and Nourse 1993; Evans 1990; Jud and Winkler 1995), property attributes (Peng 2013), investor characteristics, as well as regional market conditions including expected returns and rent growth rate (An and Deng 2009; Campbell et al., 2009; Plazzi et al., 2010), vacancy rate (Sivitanidou and Sivitanides 1999), investor sentiment (Clayton et al., 2009), and credit availability (Chervachidze and Wheaton 2013). We demonstrate in the context of the Hong Kong market that the property's physical attractiveness, earnings potential, and household characteristics all matter in the determination of housing rent to price ratio. However, short-term housing price appreciation rate does not seem to be significant.

The organization of the remainder of this paper is straightforward. The next section presents a simple theoretical model. We then describe our empirical methodology, which is followed by a discussion of the data. Empirical results are then presented, and the concluding remarks are given in the last section.

2 Theoretical Analysis

2.1 Model setup

2.1.1 Basic Environment

In our theoretical model, time runs continuously from zero to infinity. The population of households is normalized to unity, of which a fraction α demand to reside in for-sale

an issue as our dataset only has one rent-price ratio for each RED. The theoretical model and the empirical part would not match. See also Leung et al. (2006) for a related study.

housing and the remaining fraction $(1-\alpha)$ demand rental housing.[15] The housing stock is exogenously given at some value H, assumed to exceed the population of households at unity, each unit of which can be used for either for-sale or rental housing.[16]

A household in either for-sale or rental housing who finds their current dwelling suited to their needs is referred to as a matched household as opposed to a mismatched household – one who desires to move because the current dwelling no longer meets the household's needs. We model the distinction between matched and mismatched households by assuming that the former, but not the latter, receive a flow utility $\upsilon > 0$ per time unit. Matched households are subjected to periodic moving shocks, after which they become mismatched with their current dwellings. The shocks are meant to model events such as job relocation, separation, the loss of family members, etc., that necessitate the need to change housing. Because such shocks typically occur independent of the state of the housing market, it is natural to treat them as exogenous to the market. Formally, we model the occurrence of the shocks as Poisson processes with an arrival rate equal to some value we denote as δ.[17] An intuitive interpretation is that a given household would be hit by a moving shock at an instantaneous probability δ over each unit of time.

When an owner-occupier is mismatched, the household remains in the old house while searching for a new house to buy. The household vacates the old house once the household finds a new house to move into, but not before.[18] Similarly, when a rental household is mismatched, the household remains in the old house while searching for a new house to rent. Once a new house is found, the household breaks the lease and leaves the old house. An owner of a vacant house chooses between offering his property for sale and for rent – whichever option offers the higher expected discounted returns.

[15] In this paper, we assume that renters and owner-occupiers would not change their identities. This assumption has two major merits. First, Jovanovic and Rosenthal (1988) show that in an anonymous game (of which market equilibrium is a special case) where agents are not allowed to change their identities, the existence and certain properties of equilibrium are guaranteed. Second, endogenizing the tenure choice (i.e., rent or own) in a frictional housing market is by no means a simple problem. For a flavor of the complexity involved, see Chambers et al. (2009) which study households' tenure choices, with price and rent endogenously determined, in a life-cycle model without search friction. We believe that the goal of the theoretical analysis in this paper – to construct the simplest model possible in which the transaction volume and the rent-to-price ratio are determined simultaneously in equilibrium – is better served with a model in which tenure choices are exogenously given.

[16] See also Leung and Tse (2016), in which the stock of owner-occupied houses is assumed to be less than the population of households demanding such housing whereby a fraction of those households must stay in rental housing temporarily at each moment in time.

[17] The assumption of exogenous separation shock is the "standard" assumption in the labor market search literature, which can be traced back to Mortensen and Pissarides (1994), among others.

[18] See also Leung and Tse (2015) for the case where mismatched households need to sell their houses before they can search for new houses.

2.1.2 Matching rate and Market Tightness

The model housing market is a decentralized market in which imperfectly informed buyers and sellers engage in time-consuming search for trading partners, the outcome of which, as in Pissarides (2000), is modeled by a matching function. In particular, we assume that if there are $\widetilde{N}_S$ would-be buyers and N_{V_S} vacant houses for sale, there would be $M_S\left(\widetilde{N}_S, N_{V_S}\right)$ bilateral meetings between pairs of buyers and sellers over a unit of time, where $M_S(.,.)$ is an increasing, concave, and constant returns function.[19] For an individual buyer, a meeting with a seller comes about at the Poisson arrival rate $M_S\left(\widetilde{N}_S, N_{V_S}\right)/\widetilde{N}_S$, whereas the arrival rate of meetings for an individual seller is $M_S\left(\widetilde{N}_S, N_{V_S}\right)/N_{V_S}$. Now, $M_S\left(\widetilde{N}_S, N_{V_S}\right) = N_{V_S} M_S\left(\frac{\widetilde{N}_S}{N_{V_S}}, 1\right)$, given that $M_S(.,.)$ is constant returns to scale. Then, we can write the meeting rate for sellers as $\eta_S(\theta_S) \equiv M_S\left(\widetilde{N}_S, N_{V_S}\right)/N_{V_S} = M_S(\theta_S, 1)$, where $\theta_S \equiv \widetilde{N}_S/N_{V_S}$ denotes the ratio of the measures of buyers to sellers, or what is known as the market tightness, in the for-sale market. The meeting rate for buyers $\mu_S(\theta_S) \equiv M_S\left(\widetilde{N}_S, N_{V_S}\right)/\widetilde{N}_S$ can be shown to be equal to $\eta_S(\theta_S)/\theta_S$. With the matching function increasing and concave in both arguments,

$$\frac{\partial \eta_S(\theta_S)}{\partial \theta_S} > 0, \quad \frac{\partial^2 \eta_S(\theta_S)}{\partial \theta_S^2} < 0, \quad \frac{\partial \mu_S(\theta_S)}{\partial \theta_S} < 0, \quad \frac{\partial^2 \mu_S(\theta_S)}{\partial \theta_S^2} > 0. \tag{1}$$

The upshot of all this is that the meeting rates for buyers and sellers are functions of the tightness of the market only, independent of the scale of the market due to the constant-returns assumption on the matching function. Intuitively, in a tighter market as manifested by a larger θ_S, there are more buyers for each seller and as a result, it would be easier for an individual seller to meet a potential buyer and the converse holds for an individual buyer.

A bilateral meeting does not guarantee that the house in question is a good match for the would-be buyer though. We assume that the would-be buyer indeed finds the house a good match with probability $q \in (0,1)$.[20] Then, the instantaneous probability (Poisson arrival rate) at which a buyer finds a suitable house to buy is

[19] For instance, Blanchard and Diamond (1990) show that the aggregate matching function in the labor market in the U.S. is indeed subject to constant returns to scale. A few subsequent studies have substantiated the finding in various different contexts. See Petrongolo and Pissardes (2001), among others, for a review of the literature.

[20] Here we assume that the "match quality" is either 1 or 0, with respective probabilities q and $1-q$. A more general approach is to assume that the match quality is a continuous random variable having positive support over a range of values as in Pissardies (2000, chapter 6). In this case, there will be a reservation match quality above which the would-be buyer would find acceptable in equilibrium. The generalization would add the new endogenous variable – reservation match quality – to be determined in equilibrium, which is the counterpart to the reservation job matching quality in the labor market search literature. This generalization is useful in the latter literature in studying how the reservation match quality changes over the business cycle and how the changes

$q\mu_S(\theta_S)$, whereas the instantaneous probability (Poisson arrival rate) at which a seller successfully sells is $q\eta_S(\theta_S)$.

The same search and matching framework is imposed on the rental market.[21] In particular, over a unit of time, there would be $M_R\left(\widetilde{N}_R, N_{V_R}\right)$ bilateral meetings between pairs of individuals from the two sides of the market, where $M_R(.,.)$ is an increasing, concave, and constant returns to scale matching function and $\widetilde{N}_R$ and N_{V_R} the respective measures of rental households who are on the market and owners of for-rent vacant houses. Then, the instantaneous probability (Poisson arrival rate) at which a landlord successfully finds a tenant is $q\eta_R(\theta_R)$ and the instantaneous probability (Poisson arrival rate) at which a rental household successfully finds a house is $q\mu_R(\theta_R)$, where $\eta_R(\theta_R) \equiv M_R(\theta_R, 1)$, $\mu_R(\theta_R) = \eta_R(\theta_R)/\theta_R$, and $\theta_R = \widetilde{N}_R/N_{V_R}$ is the tightness of the rental market. Given the assumptions on $M_R(.,.)$, we have, analogous to $\eta_S(\theta_S)$ and $\mu_S(\theta_S)$ for the for-sale market,

$$\frac{\partial\eta_R(\theta_R)}{\partial\theta_R} > 0, \quad \frac{\partial^2\eta_R(\theta_R)}{\partial\theta_R^2} < 0, \quad \frac{\partial\mu_R(\theta_R)}{\partial\theta_R} < 0, \quad \frac{\partial^2\mu_R(\theta_R)}{\partial\theta_R^2} > 0. \tag{2}$$

2.1.3 Expected Discounted Payoff, Price, and Market Rental

A key element of our model is how owners of vacant houses decide between offering their properties for sale and for rent. To proceed, we first assume that all actors in the housing market are risk neutral. Second, we should restrict attentions to studying steady-state equilibrium of the housing market in which the measures of households in various states are unchanging through time.

Consider first a matched household in an owner-occupied house. Over a unit a time, the household enjoys the flow utility υ while staying in a matched house. But then, the match can be broken at each moment in time at an instantaneous probability δ, after which the given household becomes a mismatched household. A forward-looking risk-neutral matched household then has expected discounted payoff W satisfying,

$$\rho W = \upsilon + \delta\left(\widetilde{W} - W\right), \tag{3}$$

where ρ is the household's discount rate and $\widetilde{W}$ the expected discounted payoff of a mismatched household, and that $\widetilde{W} - W$ can be interpreted as the capital loss the household suffers at the moment of becoming mismatched. Had we not assumed that the market is in a steady state, W could vary over time, in which case we

help amplify the effects of technology shocks on employment and output. In the present analysis, we study the steady state of a housing market not subjected to any random shocks. Adding "stochastic matching" to the present model complicates the analysis while not helping bring out any additional insights.

[21] The virtue of imposing similar search and matching frameworks for the two markets is that we can be certain that the theoretical results to follow are not due to any possible difference in the meeting technologies assumed for the two markets.

need to append the time derivative term $\dot{W} \equiv \partial W/\partial t$ to the RHS of (3) for any possible expected capital gains or losses. The same caveat applies to all "asset-pricing" equations to be defined in the following.

A mismatched household in an owner-occupied house *no longer enjoys* any positive flow utility and is now on the market, searching for a new house to buy. The household meets a seller and finds the house for sale a suitable match at the instantaneous probability $q\mu(\theta_S)$, after which a trade between the buyer and the seller takes place at a price we denote as p, the determination of which will be spelled out in the following. Thereafter, the household moves into the new house, while becoming a matched household, and vacates the old house. In all, the expected discounted payoff of a mismatched household $\widetilde{W}$ satisfies,

$$\rho\widetilde{W} = q\mu_S(\theta_S)\left(W - \widetilde{W} - p + V\right), \tag{4}$$

where V is the expected discounted payoff the household earns from the vacant house that they just leave behind.

The respective expected discounted payoffs of matched and mismatched rental households can be defined analogously as follows,

$$\rho U = v - r + \delta\left(\widetilde{U} - U\right), \tag{5}$$

$$\rho\widetilde{U} = -r + q\mu_R(\theta_R)\left(U - \widetilde{U}\right), \tag{6}$$

where r is the endogenously determined flow rental payment. Notice that the mismatched rental household is obliged to continue paying the same rent for the house even though the house no longer matches their needs as we assume that recontracting between the landlord and the tenant is not feasible.

An important assumption in the above is that the matched owner-occupier and the renter derive the same flow utility v from staying in a matched house. The assumption ensures that the equilibrium rent-to-price ratio would be free of any effects due to any assumed difference in utility households enjoy from the two types of housing.

Thus far, we have described the expected discounted payoffs of households on the "demand side." of the two markets. We next turn to the "supply side." Consider first the expected discounted returns of putting up a vacant house for sale. Before the house is sold, it does not generate any flow payoff for the owner. The reward comes in the form of the sale price p at the time of sale, which takes place at each moment in time at the instantaneous probability $q\eta_S(\theta_S)$. The expected discounted returns of holding a vacant house for sale V_S thus satisfies,

$$\rho V_S = q\eta_S(\theta_S)(p - V_S), \tag{7}$$

where $p - V_S$ can be interpreted as the "capital gain" accruing to the owner from a successful sale. The expected discounted returns of putting up a vacant house for rent V_R can be defined similarly,

$$\rho V_R = q\eta_R(\theta_R)(R - V_R), \tag{8}$$

where R denotes the expected discounted returns of having a house rented to a matched rental household. To complete the definition, we next turn to the determination of R. While the house is occupied by the rental household, the owner collects the flow rental r per time unit. And then, at the instantaneous probability δ, the tenant is hit by a moving shock, after which it will be on the market searching for a new house to rent. Hence,

$$\rho R = r + \delta\left(\widetilde{R} - R\right), \tag{9}$$

where $\widetilde{R}$ denotes the expected discounted returns of having a house rented to a mismatched household who will break the lease as soon as the household succeeds in finding a new house to move into. At this point then, the house becomes vacant and the owner begins collecting the expected discounted returns of such a house. Thus, $\widetilde{R}$ satisfies,[22]

$$\rho\widetilde{R} = r + q\mu_R(\theta_R)\left(V - \widetilde{R}\right). \tag{10}$$

We have yet to specify how the sale price p and the market rental r are determined. In a perfectly competitive market with fully informed participants on both sides of the market, the market price should be at the level that clears the market. In a frictional market with imperfectly informed buyers and sellers, the market need not and usually would not be cleared at any given moment. Moreover, unlike the buyers and sellers in a perfectly competitive market who can trade with just anyone in the market at any moment in time, when a pair of potential trading partners meet each other in the environment that a search and matching model describes, they may only trade with each another as search frictions prevent them from contacting others. There is then a "bilateral monopoly," so to speak, in which the terms of trade may be determined by either the seller committing to some sale price beforehand, the buyer committing to some purchase price beforehand, or by the two bargaining over the surplus of the trade. Assuming commitment by either the buyer or the seller is ad hoc and the theoretically more appealing approach is that the terms of trade are reached via the two sides bargaining over the trade surplus. In Nash bargaining in game theory,[23] each side earns one-half of the surplus of trade. In the context of the present model, p and r satisfy the following two conditions, respectively,

$$W - \widetilde{\widetilde{W}} - p + V = p - V_S, \tag{11}$$

[22]Insofar as $\widetilde{R}$ depends on V, given that once a mismatched tenant finds a new house, he/she would move out and the landlord would be left with a vacant house, whose value is V, R ultimately also depends on V. In this regard, our model differs significantly from the models in Xiao (2014) and Sato and Xiao (2015), where the search friction is only present in the labor market but not in the rental housing market.

[23]See Rubinstein (1990) for example.

$$R - V_R = U - \widetilde{U}. \tag{12}$$

In (11), the LHS gives the surplus the buyer earns from becoming a matched household, whereas the RHS gives the surplus the owner of the for-sale vacant house earns in selling the house. The bargaining concludes with the sale price p set at the level that equates the surpluses. A similar interpretation applies to (12).

The final step is to analyze how an owner of a vacant house chooses between putting up the house for sale and for rent. In maximizing expected discounted returns, the owner chooses putting up the house for sale instead of for rent if and only if $V_S > V_R$. Then, the value of a vacant house is given by,

$$V = \max\{V_S, V_R\}.$$

Now, if $V_S > V_R$, all vacant houses will be for sale, whereas if the inequality holds in reverse, all vacant houses will be for rent. Either scenario cannot be equilibrium with a fixed fraction of households demanding for-sale and the rest demanding rental housing. This means that

$$V = V_S = V_R \tag{13}$$

must hold in equilibrium.

At this point, we can solve (3)-(13) for all the expected discounted returns, the sale price, and the market rental – W, $\widetilde{W}$, U, $\widetilde{U}$, V_S, V_R, R, $\widetilde{R}$, p, r – in terms of the two market tightness θ_S and θ_R. The equation which is of particular interest in this exercise is the solution of (13), given by

$$V_S = V_R \Rightarrow \frac{\eta_S(\theta_S)}{2(\rho+\delta) + q\mu_S(\theta_S)} = \frac{\eta_R(\theta_R)}{\rho+\delta+q\mu_R(\theta_R)}, \tag{A}$$

which is simply a no-arbitrage condition for how the two market tightness should be related to ensure that owners of vacant houses earn the same expected returns between putting up their properties for sale and for rent.

2.1.4 Stock-flow Equations

We next turn to restrictions on the measures of matched and mismatched households and the stock of vacant houses in the two markets that should hold in market equilibrium. First, with a fraction of the population α demanding for-sale housing, any member of which must either be staying in a matched or a mismatched house at each moment in time,

$$N_S + \widetilde{N}_S = \alpha, \tag{14}$$

where N_S denotes the measure of matched and $\widetilde{N}_S$ the measure of mismatched households in for-sale housing. A similar relation holds in the rental market,

$$N_R + \widetilde{N}_R = 1 - \alpha, \tag{15}$$

where N_R and $\widetilde{N}_R$ are the respective counterparts to N_S and $\widetilde{N}_S$ for the rental market.

Since each household must stay in one and only one house at each moment in time, the measure of occupied houses is simply equal to the population of households at unity. With the market endowed with a stock of $H > 1$ houses, the stock of vacant houses is then just equal to $H-1$. Write N_{V_S} and N_{V_R}, respectively, for the measures of vacant houses for sale and for rent. Then,

$$N_{V_S} + N_{V_R} = H - 1. \tag{16}$$

The last set of conditions describe how the measures of matched households in the two market evolve over time. In particular, over a unit of time, the measure of matched households in for-sale housing falls by the measure of those who are hit by moving shocks but increases by the measure of the successful buyers among mismatched households in the interim. Then,

$$\dot{N}_S = q\mu_S(\theta_S)\widetilde{N}_S - \delta N_S.$$

A similar equation holds for the rental market,

$$\dot{N}_R = q\mu_R(\theta_R)\widetilde{N}_R - \delta N_R.$$

In this paper, we shall restrict the analysis to steady-state equilibrium in which the measures of matched households in each market are unchanging over time, whereby $\dot{N}_S = \dot{N}_R = 0$.[24] The two equations above then specialize to, respectively,

$$\delta N_S = q\mu_S(\theta_S)\widetilde{N}_S, \tag{17}$$

$$\delta N_R = q\mu_R(\theta_R)\widetilde{N}_R. \tag{18}$$

Now, together with the definitions $\theta_S = \widetilde{N}_S/N_{V_S}$, $\theta_R = \widetilde{N}_R/N_{V_R}$, $\mu_S(\theta_S) = \eta(\theta_S)/\theta_S$, and $\mu_R(\theta_R) = \eta(\theta_R)/\theta_R$, we can solve (14)-(18) for N_S, $\widetilde{N}_S$, N_R, $\widetilde{N}_R$, N_{V_S} and N_{V_R} as functions of θ_S and θ_R. Of particular interest in this exercise is the solution of (16), given as

$$N_{V_S} + N_{V_R} = H - 1 \Rightarrow \frac{\delta(1-\alpha)}{q\eta_R(\theta_R) + \delta\theta_R} + \frac{\delta\alpha}{q\eta_S(\theta_S) + \delta\theta_S} = H - 1, \tag{E}$$

[24]The steady-state equilibrium can be interpreted as the long-run equilibrium that the market should approach over time from any initial state. Take the equation for $\dot{N}_S$ for example. Suppose that the RHS is positive initially. Then, N_S should increase over time while $\widetilde{N}_S$ should decline as more and more mismatched households find new houses to buy. Eventually, the two flows that make up the RHS of the equation should be equalized, at which point the market reaches the steady state. We did not explore the transition dynamics as we believe that the major lessons that concern us hold both in and off the steady state. Also, since our dataset is cross-sectional and as such is not suited to test the implications from the transition dynamics.

which describes how the two market tightness have to be related to equate the sum of the vacancies in the two markets to the stock of vacant houses in existence. If we think of N_{V_S} and N_{V_R} as the respective "demand for vacancies" in the two markets and $H-1$ as the "supply of vacancies", (E) can be interpreted as the counterpart to the market clearing condition in a model of a frictionless market.

2.2 Equilibrium

We now have two equations, (A) and (E), in two unknowns, θ_S and θ_R, the solution of which constitutes the equilibrium of the model housing market.

Proposition 1 *A unique equilibrium exists.*

Equation (A) defines an implicit function for θ_R in terms of θ_S that begins at the origin and then is increasing everywhere as depicted in Figure 1. Intuitively, when the for-sale market is tighter, vacant houses are sold faster on average. Putting up a house for sale then should yield greater expected discounted returns. For the no-arbitrage condition to continue to hold, the expected discounted returns of putting up a house for rent have to increase in tandem, which comes about when the rental market is tighter too. Equation (E) defines another implicit function for θ_R in terms of θ_S which is everywhere decreasing, as depicted in Figure 1. The interpretation is that in a tighter for-sale market, there will be faster sale on average, leaving behind fewer vacant houses for sale. In this case, there have to be more vacant houses for rent for the supply of vacant houses to be just equal to $H-1$. There can be more rental vacant houses on the market only if the houses are rented out more slowly on average due to a smaller instantaneous probability of lease as in a slower market with a smaller θ_R.

It should be clear from Figure 1 that the two functions (A) and (E) must intersect once and only once, meaning that there is a *unique* pair of $\{\theta_S, \theta_R\}$ that solves the two equations.[25] Once the two market tightness are pinned down, the various equilibrium expected returns, the sale price, the market rental, and the measures of households and vacancies in the two markets are known given that they can all be shown to be functions of $\{\theta_S, \theta_R\}$.

Before proceeding further, we would like to argue that the two equilibrium conditions (A) and (E) are both rather compelling. If condition (A) fails to hold, then no vacant houses would be made available for sale or for rent, depending on whether V_S

[25] It is well known that even a simple search and matching model can display instability and indeterminacy, such as in Krause and Lubik (2010). The Proposition here shows that the equilibrium is unique and hence rules out indeterminacy. The empirical analysis to follow is by and large motivated by the model's predictions on how the sale volume and the rent-to-price ratio are determined simultaneously and vary in response to changes in the environment. A model having multiple equilibria, as a rule, has no definite predictions about how the endogenous variables may vary when the environment changes. Among others, see Jovanovic (1989) for more discussion.

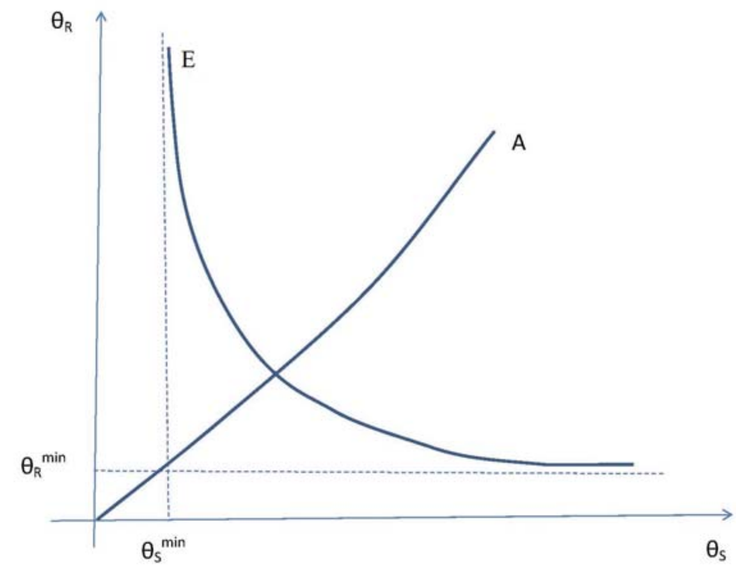


Figure 1: Existence and uniqueness of equilibrium

falls below or exceeds V_R. In particular, suppose $V_S < V_R$ initially; then the tightness of the for-sale market should reach infinity as all owners of vacant houses shun the market. But then if the market were that tight, the sale price p and then the expected discounted returns of putting up a house for sale V_S should rise significantly to overtake the initially greater expected returns V_R. This kind of oscillation, if it ever takes place at all, should eventually settle down when $V_S = V_R$ just holds. Condition (E) which stipulates that the sum of the vacancies in the two market $N_{V_S} + N_{V_R}$ is equal to the stock of vacant houses $H - 1$ is a straightforward adding-up constraint, the violation of which gives rise to a logical inconsistency of the first order.

2.3 Empirical Implications

Our main goal in studying a model of a frictional housing market is to analyze how the transaction volume in the for-sale market and the rent-to-price ratio are determined simultaneously and how the two may be correlated in the time series and in the cross section. At this point, it helps to simplify to assume that $\eta_S(.) = \eta_R(.)$; i.e., there is the same matching function for the for-sale and the rental markets. This is a reasonable assumption, as there should not be any reason to believe the matching technologies in the for-sale market and the rental market should differ in any fundamental manners given that real estate agencies and their online platforms almost always serve both markets.

2.3.1 Transaction Volume

In the for-sale market, there is a stock of N_{V_S} vacant houses on the market at each moment in time, each one of which is sold at the instantaneous probability $q\eta_S(\theta_S)$. Then, over a unit of time, there would be

$$T_S = q\eta_S(\theta_S) N_{V_S}. \tag{19}$$

houses sold in the market.

Proposition 2 *The transaction volume in the for-sale market T_S is increasing in H, α and q.*

Proof. Solving (14)-(18) for N_{V_S} and then substituting into (19),

$$T_S = \frac{\alpha\delta q\eta_S(\theta_S)}{q\eta_S(\theta_S) + \delta\theta_S}, \tag{20}$$

which is increasing in α and q but decreasing in θ_S.

First note that (A) holds only if $\theta_S > \theta_R$ for $\eta_R(\theta) = \eta_S(\theta)$. Then, an increase in α lowers the LHS of (E), which must be followed by decreases in θ_S and/or θ_R for the condition to be restored. An increase in H must similarly be met by decreases in θ_S and/or θ_R for (E) to continue to hold. By (A), the two market tightness goes up and down together. Then an increase in H or α must be met by a decline in θ_S. Thus, T_S is increasing in α by the direct effect of a larger α and the indirect effect of a smaller θ_S, whereas T_S is increasing in H by the indirect effect of a smaller θ_S.

An increase in q lowers the LHS of (E), meaning that there has be to a smaller θ_R for each θ_S for the condition to hold, as depicted in Figure 2. It is straightforward to verify that with a larger q, there is a larger θ_R for each θ_S for (A) as depicted in Figure 2, given that $\theta_S > \theta_R$ in equilibrium. Now, either tendency gives rise to a smaller θ_S as can be seen in Figure 2. T_S, as given in (20), then increases by the direct effect of a larger q and the indirect effect of a smaller θ_S. ■

That there should be a greater volume of sales in a market with a larger housing stock and where a larger fraction of households demand for-sale housing is perhaps not surprising. Holding H and α constant, the Proposition says that there would also be a greater sale volume when there is a higher matching probability q with which more meetings between pairs of buyers and sellers result in successful sales. But notice that when houses are sold more quickly, there can only be fewer vacant houses left behind, other things equal. Yet, with the housing stock H given, there must be the same stock of vacant houses on the market. The apparent contradiction is resolved once it is recognized that when mismatched households spend less time finding suitable new matches, they also vacate their old houses at a faster rate, refurbishing the market with vacant houses at a similarly faster rate to help maintain the same stock of vacant

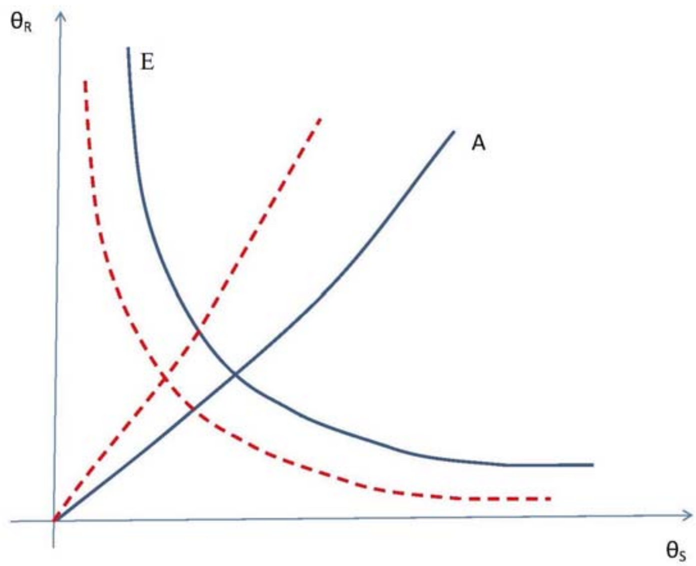


Figure 2: Comparative steady states of q

houses on the market. In all, with a greater matching probability q, more houses are put on the market for sale and more such houses are sold at each moment in time.

An empirically more relevant question is how the normed transaction volume T_S/H may vary with the environment rather than how T_S itself varies. Of course, if T_S is increasing in α and q, T_S/H is increasing in the two exogenous variables too. However, T_S/H need not be increasing in H even if T_S is increasing in H. Intuitively, with the same population of households demanding for-sale housing but a greater housing supply, there should be fewer houses sold per unit of the housing stock.

Conjecture 3 *The normed transaction volume in the for-sale market T_S/H is decreasing in H.*

In a wide variety of quantitative analyses that we undertake, we find the Conjecture indeed always holds.[26]

2.3.2 Sale Price and Market Rental

We next turn to the determination of the sale price and market rental. Solving (3)-(13), we find that

$$p = \frac{\rho + q\eta_S(\theta_S)}{(2\rho + 2\delta + q\mu_S(\theta_S))\rho} \upsilon, \tag{21}$$

[26]The analyzes assume $\eta_S(\theta) = \eta_R(\theta) = a\theta^b$ for some $a > 0$ and $b \in (0,1)$. The computer codes for all quantitative analyzes to follow are available upon request.

$$r=\frac{\left(\rho+\delta+q\mu_R\left(\theta_R\right)\right)\left(\rho+q\eta_R\left(\theta_R\right)\right)+\delta q\mu_R\left(\theta_R\right)}{\left(\rho+\delta+q\mu_R\left(\theta_R\right)\right)^2}\upsilon. \tag{22}$$

Proposition 4 *The sale price p is decreasing in H and α. The market rental r is decreasing in H and α if $\rho>\delta$.*

Proof. The sale price p, as given in (21), is increasing in θ_S. The market rental r, as given in (22), is guaranteed increasing in θ_R if $\rho>\delta$. The two market tightness is decreasing in H and α, as shown in the proof of Proposition 2. ■

With a larger housing stock, there would more for-sale vacant houses on the market and a more sluggish market follows. Sale price p then falls as the market slows down.

The reason for how p should be decreasing in α too is as follows. First observe that $\theta_S>\theta_R$ in equilibrium from (A), which results from the difference in the bargaining environment facing agents in the two markets. In the for-sale market, the surplus of a match is equal to $W-\widetilde{W}$, independent of the sale price p and the value of a vacant unit V. That the sale price – a transfer between the buyer and the seller – is not part of the surplus is not surprising. More subtle is that the value of a vacant unit V is also not a factor, which is the result of the effect of the seller forgoing a vacant unit and the effect of the buyer becoming an owner of one cancelling out. In the rental market, the surplus of a match is $R-V$ in addition to $U-\widetilde{U}$, the latter roughly equal in value to $W-\widetilde{W}$ in that the gains to both the rental household and the owner-occupier by becoming matched are equal to the PV of the flow payoff υ. Apparently, there tends to be a larger surplus in the rental market than in the for-sale market, other things equal. In the rental market, when a match takes place, the owner of the previous units occupied by the rental household suffers a capital loss $V-\widetilde{R}$. This loss can be thought of as an externality as it is not taken into account in the bargaining between the rental household and his new landlord. Had it been considered by the agents, the loss would have canceled out the gain $R-V$. All this then tends to cause V_R to exceed V_S for the same market tightness. The no-arbitrage condition thus requires a tighter for-sale market to equalize V_S and V_R. In this case, given θ_R and θ_S, transferring one household from the less tight rental market to the tighter for-sale market raises the overall market tightness and therefore lowers the overall measure of vacant houses $N_{V_R}+N_{V_S}$ from the previous level. To restore the equality in (E) requires a decline in market tightness in one or both markets. But since, by the no-arbitrage condition (A), the tightness in the two markets must move in the same direction, θ_S and θ_R decline in tandem, where the decline in the former causes p to go down.

In the meantime, in a slower market, there will not just be more houses for sale, but also more houses for rent as well. Then, perhaps the market rental should also go down too in response to either a decline in H or α. Surprisingly, it does not seem possible to ascertain such a tendency without further assumptions. The present setup assumes that the renter continues to stay in the old house and pay the same rent r

after the match between the household and house is broken while the household is searching for a new house. In a slower market, on average, a new match is found sooner. This raises the renter's surplus from entering into the lease in the first place, knowing that on average he/she is only obliged to continue paying the same rent for a shorter period of time after the newly found house no longer matches his/her need. It cannot be ruled out in general that this effect dominates the direct negative effect of a slower market on equilibrium rental unless $\rho > \delta$. In some extensive quantitative analyses, however, we have not uncovered a single instance where indeed r becomes increasing in H or α even for ridiculously large δ. In all, it seems safe to conclude that an increase in the two variables should also lower r, in addition to lowering p for any practical values of δ.

We next move to check how p and r should vary with q. First, notice that by (21), holding constant θ_S and therefore the instantaneous meeting probabilities $\eta_S(\theta_S)$ and $\mu_S(\theta_S)$, the sale price p is increasing in q – other things equal, a higher matching probability means that mismatched houses, when they can be sold faster on average, are more liquid assets commanding a higher sale price. In the meantime though, a higher matching probability would be followed by a decline in θ_S to meet the restrictions in (E), which tends to lengthen the time mismatched houses can be sold, turning them into less liquid and therefore less valuable assets. Our quantitative analyses indicate that the direct positive effect dominates for small q but the indirect negative effect dominates for large q; i.e., p is initially increasing but eventually decreasing in q. Intuitively, when q is large to begin with, there tends to be a relatively small initial θ_S. With η_S a concave function, for the relatively small initial θ_S, a given decline in θ_S results in a large decline in $\eta_S(\theta_S)$. In this way, the instantaneous sale probability $q\eta_S(\theta_S)$ tends to be decreasing in q for large q.

Not unexpectedly, a similar initially positive and eventually negative effect of q is also found on market rental r in our quantitative analyses.

Conjecture 5 *The sale price p and the market rental r are both initially increasing but eventually decreasing in q.*

2.3.3 Rent-to-Price Ratio

Dividing (22) by (21) and making use of (A),

$$\frac{r}{p} = \left(\frac{\rho + q\eta_R(\theta_R)}{\rho + q\eta_S(\theta_S)} \frac{\eta_S(\theta_S)}{\eta_R(\theta_R)} + \frac{\delta q \mu_S(\theta_S)}{(\rho + \delta + q\mu_R(\theta_R))(\rho + q\eta_S(\theta_S))} \frac{\theta_S}{\theta_R} \right) \rho. \tag{23}$$

Other things equal, owners should require a rental yield above the discount rate to willingly offer his/her property for rent instead of for sale given that a rental unit would earn no rental income every now and then when the unit becomes vacant. The above equation indeed says that $r/p > \rho$ is guaranteed to hold if $\eta_S(\theta_S) = \eta_R(\theta_R)$ as in where vacant houses are sold and rented out at the same instantaneous probability.

By Proposition 4, increases in H or α cause p to decline for sure and should also lead to a smaller r. How the rent-to-price ratio should behave then depends on how the elasticities of p compare against the elasticities of r with respect to H and α. Our conjecture is that the first set of elasticities should be greater than the second set. A larger H or α causes p and r to vary through the changes in the respective market tightness in the two markets. In the rental market, as we argue above, an increase in θ_R has both a usual positive and an indirect negative effects on r, whereas in the for-sale market, there is just the usual positive direct effect of θ_S on p. If indeed p decreases proportionately more than r does in response to a given increase in H or α, to follow is a larger r/p. The conjecture holds up always in our quantitative analyses.

The ratio r/p, given in (23), is guaranteed to be decreasing in q for small ρ and/or δ, holding constant the two market tightness, meaning that an increase in the matching probability, other things equal, tends to exert stronger positive effects on the sale price than on the market rental. This is reasonable in that an improvement in asset liquidity should mainly help raise the sale price, whereas the market rental should not directly benefit from the improvement.

The rent-to-price ratio can also vary in response to the given increase in q to the extent that the increase in q impacts on the two market tightness, on which p and r depend. By Conjecture 5, an increase in q should cause both p and r to go up initially but to fall eventually as the two market tightness decline. Along the increasing phase, p should increase proportionately more than r does, just as p rises faster than r in response to a given decline in H or α. In this case then, r/p should fall. Along the decreasing phase, p should decrease proportionately more than r does, to which a larger r/p should follow. In our quantitative analyzes, however, we find that r/p is everywhere decreasing in q. Apparently, any tendency for r/p to increase with q arising from the changes in market tightness is more than offset by the direct negative effect of q on r/p.

Conjecture 6 *The rent-to-price ratio r/p is increasing in H and α but decreasing in q.*

In sum, r/p varies in the way it does as described in the Conjecture is due to the tendency that the sale price should be more sensitive to changes in market tightness arising from changes in supply H and demand composition α and that improvements in market liquidity should help raise the sale price more than the market rental.

2.3.4 Correlation between the Transaction Volume and Rent-to-Price ratio

Proposition 2 and Conjecture 6 together imply that:

Corollary 7 *Variations in α induce a positive correlation between T_S/H and r/p, i.e.,*

$$corr\,(T_S/H, r/p) > 0. \tag{24}$$

Variations in q induce a negative correlation between T_S/H and r/p, i.e.,

$$corr\left(T_S/H, r/p\right) < 0. \tag{25}$$

Conjectures 3 and 6 together imply that:

Corollary 8 *Variations in H induce a negative correlation between T_S/H and r/p,*

$$corr\left(T_S/H, r/p\right) < 0. \tag{26}$$

By construction, the housing units in the model are observationally identical and as such the model has no cross-sectional implications. But then, if we imagine, instead of there being one unified housing market, that there are numerous segmented markets, each of which is as described in the model above. The theoretical results can then be interpreted as cross-sectional implications. The last two Corollaries suggest that across the segmented markets, transaction volumes and the rental yields can exhibit negative correlations or positive correlations. Our empirical analyses in the following are in the main motivated to resolve this ambiguity of the theoretical model.[27]

3 Our Econometric Framework

In this section, we describe our econometric framework for testing the predictions of the theoretical model with a cross-sectional data set. In particular, we test whether a change in the fraction of owner-occupants α, or in the probability of good match q, or in the housing stock H would induce the rent-to-price ratio (r/p) and the transaction ratio in the sale market (T_S/H) to move in the *same or opposite directions*. Two issues need addressing to achieve these goals. First, in the model, housing units are homogeneous, while in practice they are heterogeneous. Second, we must also obtain a measure of the probability of good match q, which would vary across different sub-markets (which are REDs in the current context). One possible method is to adopt a hedonic-type regression in order to control for the heterogeneity among housing units. Unfortunately, q is *not directly observable*. In addition, as our sample size is relatively small (132 RED), listing all the control variables on the right hand side of the regression equation would exhaust our degree of freedom, and adversely affect subsequent analysis. Therefore, a hedonic-type regression approach is *not* practical.

[27] Notice also that our model is an on-the-house-search model, meaning that mismatched households would stay in the house, even when it generates zero utility flow to the households, until they find new ones. Since utility flows to households are not directly observable, the total number of "house seekers" may not be directly observable neither. Hence, important variables such as market tightness may not be observable. In this sense, the empirical focus of this paper, namely, the turnover rate and the rent-to-price ratios may very well be the only variables that can be observed and measured without controversy in the model.

Another common approach is to use principal component (PC) analysis, which "combines" different variables into a few "factors" or PCs. We can then use those PCs as the right hand side variables to control for differences across RED. However, a prerequisite of adopting the PC analysis is that those variables are significantly correlated, and that a few variables can summarize most of the total variations (Jolliffe, 2005). Unfortunately, this is not the case in our sample. It takes six PCs to account for 80% of the total variations, as shown in the appendix, and hence adopting the PC analysis would not be helpful in our context. Furthermore, neither the hedonic nor PC approach provides an empirical proxy for q. An alternative approach is required.

The starting point of our approach is the observation that *variations in the fraction of owner-occupants, in the probability of good match and in housing attributes can be interpreted as changes in particular popularity measures.* For instance, a RED that is more popular (i.e. has some good features) would attract more owner-occupants. Since owners tend to stay longer than renters, we would observe less turnover on average. On the other hand, a RED with a high matching probability means that, other things being equal, the housing units are more liquid. This attracts investors and increases the popularity of the RED. Similarly, as the number of households targeting each submarket is normalized to unity, H can be interpreted as the relative demand for housing in a certain submarket. Thus, the smaller H is, the stronger the relative demand for residence in a particular submarket. Obviously, the more desirable the structural and neighborhood attributes are, the greater the demand for the RED, and hence a smaller H and a higher popularity. Accordingly, we can summarize the major implications of the search and matching model (from (24) to (26)) as follows:

$$\frac{\partial\left(r/p\right)}{\partial(Popularity)} \gtrless 0,\ \frac{\partial\left(T_S/H\right)}{\partial(Popularity)} \gtrless 0. \tag{27}$$

As the theoretical analysis does not help us to dictate the relationship, we switch to the empirical analysis. Thus, our next step is to build a *composite popularity index* based on housing affordability and a bundle of structural, neighborhood, and location characteristics. A significant advantage of this approach is that we *consolidate* the information that are contained in the structural and neighborhood characteristics of a RED into one variable, hence we can increase the degree of freedom in our regression. Thus, the "popularity" index can be interpreted as a "summary statistics" of many housing characteristics as well as a "proxy" for the matching probabilities which we do not observe directly, and therefore it needs to be inclusive. Our choice of the "ingredients" of the popularity index is influenced by the hedonic pricing literature, motivated by exploratory interviews with market practitioners,[28] and constrained by

[28] The two major real estate agencies in Hong Kong are Centaline Property and Midland Realty. One of the authors personally visited one branch of each. Questions like "What kind of condominiums is more popular among home buyers?", "What characteristics are most valued in home purchase decision?" are explicitly asked.

data availability. Following a large literature on index construction, our popularity index is a linear sum of equally weighted selected indicators.[29] We provide more details about the construction of the index in a later section and in the Appendix. It is clearly an imperfect proxy of the conceptual "popularity" in the model. Even if we were able to include all of the characteristics of a RED in the index, it is unlikely it could address all the concerns surrounding relative demand and the likelihood of a successful match, which result from the interactions of consumers with the housing market (Griliches 1971).

If we take the existence of a popularity proxy for granted, a natural formulation reflecting the simultaneous relationship between the rent-to-price ratio (r/p) and the transaction rate (T_S/H) on one side, and how both variables may relate to the Popularity index on the other is:

$$\begin{aligned} \frac{r}{p} &= \pi_0 + \pi_1 (Popularity) + \pi_2 \left(\frac{T_S}{H}\right) + X\gamma + \varepsilon_1, \\ \frac{T_S}{H} &= \beta_0 + \beta_1 (Popularity) + \beta_2 \left(\frac{r}{p}\right) + Z\phi + \varepsilon_2, \end{aligned} \tag{28}$$

where X and Z are row vectors of other controls, whose detailed contents will be discussed later. π_i, β_i, $i = 0, 1, 2$ are all scalars, while γ and ϕ are column vectors. ε_1 and ε_2 are residual terms. In line with the model implications, we should expect $\pi_1 < 0$, $\beta_1 > 0$, π_2, $\beta_2 < 0$. In words, it means that we expect more frequent trading in popular REDs. And the rental yield must be lower to compensate for higher popularity. Clearly, this simple simultaneous equations system (28) can be rearranged as

$$\begin{aligned} \varsigma \left(\frac{r}{p}\right) &= \pi_0' + \pi_1' (Popularity) + (X\gamma + \pi_2 Z\phi) + (\varepsilon_1 + \pi_2 \varepsilon_2), \\ \varsigma \left(\frac{T_S}{H}\right) &= \beta_0' + \beta_1' (Popularity) + (\beta_2 X\gamma + Z\phi) + (\beta_2 \varepsilon_1 + \varepsilon_2), \end{aligned} \tag{29}$$

where $\varsigma = (1 - \pi_2\beta_2)$, $\pi_0' = \pi_0 + \pi_2\beta_0$, $\pi_1' = \pi_1 + \pi_2\beta_1$, $\beta_0' = \beta_0 + \pi_0\beta_2$, $\beta_1' = \beta_1 + \pi_1\beta_2$. Several issues emerge. Clearly, (T_S/H) is correlated with ε_1, and (r/p) is correlated with ε_2. Thus, there are "*cross-equation restrictions*" implicitly imposed in (28),[30] and if we estimate each equation in (28) separately, the ordinary least squares (OLS) estimates are inconsistent. Similarly, even though ε_1 and ε_2 in (28) can be assumed to be independent error terms, $(\varepsilon_1 + \pi_2\varepsilon_2)$ and $(\beta_2\varepsilon_1 + \varepsilon_2)$ in (29) are clearly not

[29] Notice that while some of our indicators are continuous variables, some are dummy variables. And without any prior knowledge of the detailed preference of the households, we consider the choice of using an equal-weighted linear sum of different indicators to be both sensible and consistent with the practice of the literature on index construction.

[30] For more discussion of "cross-equation restrictions," see Sims (1980), among others.

independent. Again, it is a kind of "*cross-equation restriction*." Thus, ignoring such a restriction and estimating the two equations in (29) separately with OLS may not deliver efficient estimates, as the covariance between $(\varepsilon_1 + \pi_2\varepsilon_2)$ and $(\beta_2\varepsilon_1 + \varepsilon_2)$ are not properly taken into considerations. In this case, structural estimators should be preferred. More specifically, we use 3SLS, which is a full information estimator, and estimate the two equations as a system.[31] The procedure is as follows:

$$\begin{aligned} & \text{Stage 1} \\ \left(\frac{r}{p}\right) &= \theta_0 + \theta_1 \left(Popularity\right) + X\theta + Z\delta + \omega_1, \\ \left(\frac{T_S}{H}\right) &= \xi_0 + \xi_1 \left(Popularity\right) + X\xi + Z\kappa + \omega_2, \end{aligned}$$

where θ_0, θ_1, ξ_0, ξ_1 are scalars, and θ, δ, ξ, κ are vectors. We obtain the OLS projections of the endogenous variables, $\widehat{\left(\frac{r}{p}\right)}$ and $\widehat{\left(\frac{T_S}{H}\right)}$.

$$\begin{aligned} & \text{Stage 2} \\ \left(\frac{r}{p}\right) &= \pi_0 + \pi_1 \left(Popularity\right) + \pi_2 \widehat{\left(\frac{T_S}{H}\right)} + X\gamma + u_1, \qquad (30) \\ \left(\frac{T_S}{H}\right) &= \beta_0 + \beta_1 \left(Popularity\right) + \beta_2 \widehat{\left(\frac{r}{p}\right)} + Z\phi + u_2, \qquad (31) \end{aligned}$$

where γ, ϕ are vectors. We retain the residuals to form a consistent estimate of the covariance matrix of the disturbances.

Stage 3
Perform Generalized Least Squares (GLS) estimation on (30), (31)
with the covariance matrix in Stage 2 as the optimal weighting matrix.

For the future reference, (30) would be labelled as the rent-to-price equation, and (31) as the turnover equation.

4 Data and Popularity Index construction

In this paper, we use data from three different sources. First, we retrieve micro-transaction records from EPRC, a commercial Hong Kong real estate database. Second, we obtain demographic and socioeconomic data from the 2011 Hong Kong Census. We also use online resources, including real estate brokers' and government websites, geographic information system, and popular local internet discussion forums,

[31] Among others, see Wooldridge (2010) for more discussion of 3SLS.

which provide us with complementary information on structural, neighborhood, and locational characteristics.

The EPRC data set has been employed by several previous studies.[32] We therefore first discuss the official census. To gain a better understanding of the socioeconomic environment, the Hong Kong government conducts a census of all households in Hong Kong every decade, and a by-census for a selected sub-sample between the main censuses. Previous censuses and by-censuses published at most the district level data for public reference.[33] In the 2011 Census, the Census and Statistics Department (CSD) of the Hong Kong government provides publicly accessible statistics of "major housing estates," (i.e. REDs) "with at least 3,000 residents or 1,000 domestic households" at the time of survey.[34] Information is readily available for a total of 487 REDs (or "estates"), of which 163 are private residential REDs. We exclude all public housing, as the allocation of these units is not decided by the market and the corresponding rents and prices are highly subsidized. This collection of private residential RED is far from being ideal, but it does provide information on demographics, income and housing related expenditure disaggregated at the RED level. Hong Kong has no zip code system, but as most of the Hong Kong population lives in high-rise buildings, the RED level analysis in Hong Kong can be comparable to those zip code-level analyses of the US. We include variables for the total population, the young population aged from 25 to 44 years, the percentage of residents' with a college education, the household monthly median income, the median monthly mortgage payment, and residents working in real estate consulting industry.

The 2011 Census was conducted from June 30 to August 2, 2011. Correspondingly, we extract rental contracts, as well as "purchase and sale agreements" (i.e. contracts of for-sale housing units) from EPRC that occurred in the 163 private REDs in the *subsequent period*. Typically, a transaction is completed in two months; hence we restrict the time horizon to range from November 2011 to October 2012.[35] For each RED, we are able to observe the actual transaction volume. With some data cleaning efforts, we collect 3,616 transactions in the rental market and 29,671 transactions in the for-sale counterpart.[36] Notice that in any given period of time, a housing unit is *either* transacted in the sale market *or* the rental market, *but not both*. Therefore, to find the rent-to-price ratio while holding the quality of the housing unit constant, we

[32] Among others, see Leung et al. (2002).

[33] There are eighteen administrative districts in Hong Kong. For a list of these districts, please visit the District Council homepage http://www.districtcouncils.gov.hk/index.html .

[34] The official webpage is http://www.census2011.gov.hk/en/major-housing-estates.html.

[35] Notice that the data set is constructed in a manner that the potential simultaneous bias is mitigated. According to Leung, Leong and Chan (2002), most housing transactions in Hong Kong are finished within two months. Therefore, when we collected housing transactions between November 2011 to October 2012, even the initial sale agreements were signed after the 2011 Census was finished, and hence we can take the variables from the household survey as "pre-determined" variables. For more justifications on using pre-determined as exogenous variables, see Greene (2008), Sims (2010).

[36] See the Appendix for the details of our construction of the data set.

need to match each transaction from the rental market with one in the sale market with "similar characteristics."[37] Following Huang and Leung (2014), we require the unit for-sale and the unit for rental both to be transacted in the sampling periods, and to have the same estate code, floor level, living room orientation, the gross square footage to differ by not more than ten feet, the net gross ratio to differ by less than one percent, and to have a building age disparity of less than three years. This data-matching methodology ensures that each rental unit is "matched" with a for-sale unit that is similar in observable attributes, and inevitably implies that *some REDs will not qualify*. For instance, during our sampling period, some estates have only sale transactions but no lease transactions. Some have new leases signed but no sale transaction, or they have both but the units transacted in the for-sale and rental markets are so different that we cannot "match" them. From this, a total of 132 REDs remain in our sample.[38] As it is very likely to have more than one transaction in the rental market in each of these REDs, we would therefore obtain several rent-to-price ratios. For the estimation purpose, we take the mean value of these rent-to-price ratios in each RED.

To further enrich our empirical analysis, we compute the RED level rent growth rate and housing price appreciation rate *before* our sampling period begins.[39] We collect six years of rental records for each RED, and transform the nominal rent into real values.[40] To control for the possible quality difference, we *match* the rental units transacted with those transacted in previous years, based on the same matching criteria applied to the for-sale units and their rental counterpart. We therefore obtain five consecutive yearly growth rates of rent in a rolling fashion. The historical house price appreciation rate is constructed in the same manner, but with only two years of data.[41]

[37] According to the literature review by Todd (2006), matching estimators have a long history in the literature. See McMillen (2012), Rivkin et al. (2005), among others, for some recent applications.

[38] Clearly, this data-matching methodology is like a sample-screening process and the final results could be affected. We provide more discussion on this in a later section.

[39] Notice that our sampling period is somehow constrained by the period the 2011 Census was conducted, which is itself an administrative decision. Using the data before the sampling period would enrich our analysis and help us to capture the potentially different "initial conditions" of the REDs.

[40] We have consulted market practitioners, and we are told that in Hong Kong, most rental contracts last for one to two years. Hence, six years of rental contract collection should be sufficient for our purpose to control for "rental market conditions before the sampling period".

[41] Notice that the growth rate of rent exhibits cyclical behavior. Reverse chronologically, the average rent growth rate across our sampled REDs is 19.83% in 2010, -11.6% in 2009, 10.58% in 2008, 7.967% in 2007, and -1.267% in 2006. Such pattern suggests a longer horizon would be more reasonable.

In terms of the data conversion, Year 2005 is used as the benchmark. The CPI data comes from IFS (International Financial Statistics), which is maintained by the IMF.

A drawback of the inclusion of this additional variable is that it further reduces the number of estates to 130, as some RED might not have very active rental markets.

Notice also that what we report in the following is based on the geometric mean in the past five

Next we examine the construction of the popularity index. According to Diamantopoulos and Winklhofer (2001), constructing an index demands a proper set of formative indicators, which must include all facets of the determinants involved; any omission can lead to an incomplete index. It is obvious what domains should be captured in our context, as we need to account for both physical attributes and affordability. After considering a combination of hedonic pricing literature, interviews with real estate brokers, and data availability, we include the following variables: the affordability, the RED-level average net gross ratio, swimming pool and club house availability, developer brand, distance to a MTR station,[42] distance to CBD, distance to the waterfront, and the school zone. Table 1a provides the summary statistics. We use the transaction price as a proxy of the affordability. EPRC provides certain housing characteristics, including the building age, net gross ratio, and housing estate facilities, such as swimming pool, club house, etc. We take the mean of the net gross ratios of all transacted units in a RED as the average net gross ratio of the RED, and the mean age of all buildings in a RED as the average age of the RED. Other "ingredients" for the popularity index are derived from various online resources. For instance, we determine the distances in meters from an estate to the nearest subway stations and to the central business district (CBD) with data from the government geographic information system.[43] We also measure the distance from an estate to its nearby waterfront. Hong Kong has a long and sinuous coastline, so proximity to the waterfront means possible sea view from home, and accessibility to a nearby waterfront park or promenade.

(Table 1a about here)

The access to quality schools is also included in our popularity index.[44] Most Hong Kong students at the end of Secondary Grade Six take the Hong Kong Diploma of Secondary Education (HKDSE) examination, which is the benchmark assessment for local higher education admission. Competition is high for the enrollment in government-funded degree-granting institutions in Hong Kong, we therefore focus

years. Replacing it with the arithmetic mean would make little difference to the final results.

[42]MTR is the train company which operates the train, the subway and the light-train system in Hong Kong.

[43]There is an online system. Its website is http://www1.map.gov.hk/gih3/view/index.jsp.

[44]The justifications are clear. First, the hedonic pricing literature, which is surveyed by Malpezzi (2008), among others, indicates that access to quality schools is an important determinant of house price. Studies such as Black (1999), Dhar and Ross (2012), Figlio and Lucas (2004), Hanushek and Yilmaz (2007, 2013), Nguyen-Hoang and Yinger (2011) also confirm the relationship between school quality and property value.

The recent literature seems to confirm that quality of education matters. Among others, see Hanushek et al. (2008), Hanushek and Woessmann (2012).

on the quality of secondary schools in our study.[45] There is no official ranking of high schools in Hong Kong, but parents do exchange information and rank schools on internet discussion forum, such as "Book of School." This website lists the top 100 out of 490 local secondary schools, based on parents' ratings on the campus, the teachers, students' academic performance, and non-academic attainment in music and sports.[46] We group the top 100 secondary schools by their physical locations, and consider their distributions in different districts.

Finally, we include information on nearby land auctions and the developer brand, as the former is related to the *quantity* dimension of housing supply and the latter to the *quality* dimension of housing supply. Clearly, land auctions in adjacent areas signify increases in the local housing supply, which tend to encourage prospective home buyers to exercise the option of waiting (i.e. to delay their purchases until, say, the new housing units come to the market). We retrieve land area, in terms of square meters auctioned for residential development, from the Lands Department of the Hong Kong government.[47] To match with our sample, we restrict ourselves to the auction records occurred between November 2011 and October 2012, and to those sites within 2 kilometers radius around any of our 132 REDs.[48] We combine the location and lot number in the land sale files, land surveyors' plan maps, and the street index and the lot address cross-reference table to ascertain the exact position of a site.[49]

We have clear reasons for including the developer brand in the popularity index. The primary residential real estate market in Hong Kong is dominated by a few developers; for example, in 2010, over seventy percent of new home sales come from

---

[45] The following figures may illustrate the keen competition for the enrollment in the 8 government-funded degree-granting institutions in Hong Kong. For instance, in the year 2014 HKDSE, the total number of participants is about 79,572 and only 20,089 of them (25.25%) are admitted by the 8 local UGC funded institutions. For more details, see http://cdcf.ugc.edu.hk/cdcf/indepthAnalysis.do

[46] Hong Kong has a rather complicated primary and secondary education system consisting of seven types of schools according to their funding source including government established schools, aided schools, caput schools, schools under direct subsidy scheme (DSS), private schools, international schools, and schools operated by English schools foundation (ESF). Other than international and ESF schools catering the need of foreigners, other schools mainly accommodate local children. A more detailed description of various types of schools in Hong Kong can be found in a public document provided by the Hong Kong University of Science and Technology at https://www.ab.ust.hk/hro/PubDoc/new_staff_guide/resources/school_types_in_hk.pdf. The top 100 secondary schools listed on Book of School at http://www.bookofschool.com/school/controller/schoolSearch?reporttype=secondaryranking&schoolcategory=2 are either government schools, or aided schools, or DSS schools. The former two kinds are public-sector schools, which are fully subsidized by the government and provide local children with free education. DSS schools are private schools receiving government subsidies.

[47] The web link for download is http://www.landsd.gov.hk/en/landsale/records.htm.

[48] See Lam and Cheung (2000) for more discussion on the 2km radius.

[49] These information can be found on http://www.landsd.gov.hk/mapping/en/landsale/2011/ls_t.htm and http://www.landreg.gov.hk/tc/public/pu-si_agree.htm.

Cheung Kong Holdings and Sun Hung Kai Properties, two of the largest property developers in Hong Kong.[50] Other years are not as extreme, but the oligopolistic structure remains. In our sample, nine developers, alone or in joint venture, participate in the developments of 119 projects.[51] All other developers account for the design and construction of only thirteen projects.[52] Severe information asymmetry could potentially affect the housing market, as unlike buyers, the developers possess the full knowledge regarding the development quality. Home owners may only discover the true quality of the housing units after moving in, or even after years of occupation. This clearly discourages market transactions. In this market environment, some firms put more efforts into reputation building than others, such as providing high-quality units more consistently, to signal their quality.[53] The major developers are therefore not equally favored by the market. Sung Hung Kai Properties appears to be the most valued, and Cheung Kong Holdings the least valued. We carried out a Factiva keywords search to investigate this.[54] A search for "Sung Hung Kai" and "award", retrieved 163 news articles reporting numerous accolades.[55] Sung Hung Kai

[50]See Vivian Kwok (2010 August 12), Two developers tower over market, *South China Morning Post*.

In this paper, the real estate developer information are extracted from real estate broker's websites. We also obtain the information about the RED scale, which will be further discussed later, from these websites. For instance, the website of Centaline Property is http://hk.centanet.com/eng/ehome.htm, and the link for Midland Realty is http://www.midland.com.hk/en/.

[51]These nine property developers are Cheung Kong Holdings, Chinachem Group, Hang Lung Properties, Henderson Land Development, MTR Properties, Nan Fung Group, New World Development, Sino Group, and Sun Hung Kai Properties.

Notice that Cheung Kong Holdings belongs to Hong Kong's most famous multi-national conglomerate, Cheung Kong Group. The group has another company called Hutchison Whampoa Limited, which also actively participated in the real estate market. In this paper, we consider REDs developed by Hutchison Whampoa Limited as if they were developed by Cheung Kong Holdings.

[52]There are economic reasons suggested for such oligopolistic market structure, despite the fact that a significant amount of land sale is conducted through open auction. For instance, due to the population density, high-rise buildings are needed as the dominant form of residential housing in Hong Kong. The development of high-rise buildings, however, demands specialists and a large sunk cost. Thus, it is unlike that such buildings will be provided by a continuum of small developers, as some textbooks would assert. Among others, see Leung and Tang (2015) for more discussion for the institutional details of the Hong Kong market.

[53]Among others, see Mas-Colell et al. (1995) for more discussion of the signalling theory and reputation.

[54]Factiva.com is a global news database with wide coverage in Chinese publications as well. We initiate the search in the language of traditional Chinese.

WiseNews, a local news database more focused on Chinese newspapers and magazines, may fulfill the same goal. However, it does *not* allow for subject classification. Searching with a combination of keywords listed above generates too many noises. Factiva.com enables us to restrict ourselves to the more relevant "Corporate Awards" subject. As a subsidary of the Dow Jones & Reuters Company, it is the industry leader and arguably has more advanced technology. So we mainly use Factiva.com for news search.

[55]In particular, two honors stand out. For both 2012 and 2013, four REDs of Sung Hung Kai Properties have been awarded the annual "Best REDs" by the Hong Kong Professional Build-

won the most awards between March 8, 2006 and the time of the search on March 11, 2014. In contrast, a similar search for Cheung Kong Holdings from November 30, 2005 to the same ending date returned only 88 pieces of news on corporate awards, many of which are not for its contribution to the real estate development, but for the company's investment value as a multi-national conglomerate, its corporate social responsibility, or for the lifetime achievement of Li Ka Shing, the founder and leader of the company. In fact, by a keyword search for "estate" and "quality," we find in four articles documenting consumer complaints and even lawsuits concerning the poor quality of buildings designed and built by Cheung Kong Holdings in Beijing, Guangzhou, Hong Kong, and Vancouver.[56]

As economists, we believe that if there is a consensus of quality difference, it should be priced in the market. We verify this by evaluating the developer brands in a hedonic pricing framework with our underlying purchase and sale data. We introduce nine dummy variables for the nine major developers, and the control group includes all other developers. In addition, we control housing *structural characteristics* including gross square footage, net gross ratio, floor level, condominium orientation, and building age, *neighborhood characteristics* including estate scale, RED facilities, and the number of quality schools in the district, and *locational characteristics* including distance to subway stations and to the central business district, and proximity to the waterfront. We also add district dummies, in an attempt to capture some unobservable district-specific effects. Considering the possible intraclass correlation of prices within the same RED, we use the clustered standard errors for reference. The regression results are reported in Table 1b. In line with the previous discussions, Sung Hung Kai Properties is the *only developer with a significantly positive coefficient*, implying its developments sell at a premium even after controlling for different characteristics. The price discounts (i.e., negative price premium) of both Cheung Kong Holdings and Chinachem Group are about 5%. Thus, this regression suggests that the *reputation of the developers matters*, even after controlling for other factors.

[Table 1b about here.]

ing Inspection Academy, which is founded by a well-known local engineer Tsim Chai Nam in the middle of 2012. This academy aims to provide consultancy for locals with respect to quality inspections and oversights of construction, decoration, remodeling, as well as maintenance. The other two estates awarded this accolade are of Sino Group. See http://ps.hket.com/content/48406/ and http://www.hkcd.com.hk/content/2013-05/31/content_3201810.htm. Besides, Sung Hung Kai Properties has been awarded "Prestigious Corporate Brand Awards" each year ever since 2007 when the program was initiated. This brand recognition award is jointly presented by Ming Pao, one of the most credible local newspapers, and Department of Marketing of The Chinese University of Hong Kong. See http://pcbaward.com/cfm/other.cfm?html=13result for a complete list of all award winning brands for 2013, and http://www.shkp.com/en-US/Pages/awards-and-recognition/ for the accolades wined by Sung Hung Kai in each year.

[56] The articles are in traditional Chinese as well.

Based on the discussion above, our popularity index combines nine indicators, which are sorted in ascending order of desirability. As a reference point, a condominium is deemed to be the more desirable if it has a higher net gross ratio, is in a RED with both swimming pool and club house, is closer to a subway station, the CBD, and the waterfront, has good schools in the district, and is sold at a low price. Some formative indicators, namely, the proximity to the waterfront and subway, swimming pool and club house availability, are dummy variables, being scored as either zero or one. Each of the six continuous variables, affordability, net gross ratio, building age, estate scale, distance to the CBD, and the number of top one hundred secondary schools, are divided into ten equal groups. We then score the groups from 0.1 to 1 from the least to the most desirable, with increments of 0.1. With regard to the developer brand, we assign a value of one to Sung Hung Kai properties, one-third to condominiums of Cheung Kong Holdings and the Chinachem Group, and two-thirds to all the other units to represent market's valuations of these developers. This scoring system ensures that all the formative indicators range from zero to one, and no indicator carries more weight than another. For each RED, we take the mean of the nine scores to be our final RED (or estate)-level popularity index, with a mean of 6.249, a standard deviation of 1.277, a sample minimum of 2.948, and a sample maximum of 9.018. A complete list of ranked REDs in our sample is presented in the Table 1c. Figure 3 relates the popularity index to the housing price in the past five years at the RED level, with each RED represented by one dot. Figure 3a shows no notable relation between the popularity index and mean price growth rate of REDs once three outliers are removed, but Figure 3b clearly demonstrates a negative correlation between the index and the associated price growth volatility. The more popular an estate is, the lower the price growth volatility it displays.

[Table 1c, Figure 3a, Figure 3b about here.]

We attempt to make our popularity index all-inclusive, but it does have some shortcomings. It only considers the physical characteristics of a condominium, which are more objectively measured. Nevertheless, without integrating the market forces of supply and demand, our constructed "popularity" can only serve as an imperfect measure of the "popularity" in the theoretical model. To illustrate this we can compare two REDs, The Belcher's and Kingswood Villas. The former is more appealing than the latter in almost every aspect. The Belcher's was developed by Sung Hung Kai Properties, and is situated just $2,500$ meters away from the CBD, 310 meters away from the beautiful waterfront of Victoria Harbor. In contrast, Kingswood Villas is a Cheung Kong Holdings development, and is situated over $25,000$ meters away from the CBD, and nowhere near the seaside. Kingswood Villas is six years older than The Belcher's, and has only seven high-quality secondary schools in the vicinity, whereas The Belcher's has nine. Under our construction, other things being equal, The Belcher's should be ranked higher.[57] However, other things are not equal. During the sampling period, the mean selling price is more than 13,000 Hong Kong Dollars (HKD) per square feet for The Belcher's, but under 3,500 HKD per square feet for Kingswood Villas.[58] The Belcher's residents have a median monthly household income of 93,580 HKD, while the residents of Kingswood Villas merely have 26,000 HKD, according to the census.[59] Kingswood Villas is obviously more attractive to ordinary people in Hong Kong.

5 Empirical Findings

The previous sections provide detailed discussions on the data set used, the construction of the popularity index and the econometric framework. This section reports *how* the theoretical model proposed is verified empirically, and *what* the empirical findings are. Existing literature has suggested four types of empirical determinants of the rent-to-price ratio, which are the macroeconomic variables, regional market conditions, property attributes, and investor characteristics.[60] This paper complements the literature by exploring the linkage between rental yields on the one hand, and property attributes and the characteristics of the owner-occupiers, which can also be interpreted as the "investors" of the housing market, on the other hand. *To achieve this goal*, we estimate the empirical model (28) with a dataset compiled by a cross-sectional data set of household survey and a panel of housing transactions. We have discussed in details how the rent-to-price ratio, the turnover rate and the popularity index are constructed. We now provide more explanations about how the vectors X and Z in (28) are constructed. The choice ingredients in vector X can be easily justified. They include the average rent growth rate and house price appreciation in the previous years at the RED level. As property attributes do not change significantly over time, especially for condominium, those previous year rent and price data enable us to have a good control of the "initial conditions."[61] The vector X set also

[57] Belcher's is ranked the sixth, and Kingswood Villas is ranked the 37th out of 132 REDs in our sample.

[58] The official exchange rate is 7.8 HKD for one US dollar. And one square meter is roughly equal to ten square feet.

[59] Alternatively, we could use hedonic estimates as weights. But then we will not be able to dictate the "affordability" because price is the dependent variable in the hedonic regression.

[60] While earlier studies focus predominantly on the effects of macroeconomic variables, such as interest rates, expected inflation, and stock market risk and return (Ambrose and Nourse 1993; Evans 1990; Jud and Winkler 1995; Peng, 2013), more recent research extends this to regional market conditions, most importantly rent growth rates and the expected returns of real estate (An and Deng 2009; Campbell et al. 2009; Plazzi et al. 2010). Other relevant local market factors include vacancy rate (Sivitanidou and Sivitanides 1999), investor sentiment (Clayton et al. 2009), and credit availability (Chervachidze and Wheaton 2013).

[61] In the literature, some authors use the historical returns of the RED as proxies for the corresponding expected values.

contains information of the owner-occupiers, such as the median household monthly income, mortgage payment-to-income ratio, percent of residents with college degrees, and the percent of residents working in real estate consultation industry, as well as some physical conditions of the RED, such as the RED scale, nearby land auctions representing local market conditions, which have been argued as important determinants of the rent-to-price ratio in the discussions on different media. Here, we simply put them in an unifying framework and assess these discussions more rigorously.

Since asset quantities and trading have "received far less attention" in the literature (Lo and Wang, 2009), our choice of control variable vector Z in the empirical model (28) is based on simple economic reasoning. First, expected returns and earnings (rents) potential are obvious candidates. Partly as an investment vehicle, return-related variables must affect the trading activity of residential properties. Second, a real-estate-related occupation may also provide the advantage of being able to invest in property.[62] REDs housing more residents working in the real estate profession may therefore display more frequent trading. Third, the role of household income is more complicated in ex ante terms so that research is needed. Increases in income may influence housing demand, resulting in a higher sales volume, but a higher income means a higher opportunity cost of moving, which will dampen trading volume. Being agnostic of which effect would dominate, we include the variable in the regression and allow the data to inform us. Fourth, because early works on search models emphasize the existence of thick market effect (TME) (Diamond, 1982a, 1982b; Ngai and Tenreyro 2014), RED size is incorporated as a simple proxy for TME to allow for such possibility. The RED scale is extracted from real estate broker's websites.[63] Fifth, land auction is controlled as well for the justification in the following. Land auctions in adjacent areas suggest more housing supply in the near future. Hence, potential buyers may demand a deeper discount in transaction price (which usually lead to a longer bargaining process and less transactions in the short run), or simply delay their purchases. Finally, the young population proportion and building age are taken into account too, but the reasons for such inclusion will be elaborated in the next paragraph.

Notice that the variables in vector X and Z overlap, but not completely. Clayton et al. (2009) use mortgage payment as an indicator of investor sentiment and find that it is negatively associated with property capitalization rate. Following their approach, we control mortgage income ratio in the rent-to-price regression. This measure, however, is an average of many households who are at different phases of mortgage cycle. Some mortgages were signed a long time ago. There is no reason to believe mortgage income ratio will impact current trading activity, and thus is excluded from the turnover regression. As previously mentioned, housing is an investment vehicle. But it is also a consumption good, of which hedonic attributes

[62] For instance, see Munneke et al. (2012) and the reference therein.

[63] The website of Centaline Property is http://hk.centanet.com/eng/ehome.htm, and the link for Midland Realty is http://www.midland.com.hk/en/.

are important pricing components. For example, Shapiro (2006) demonstrates that human capital, as an indicator of local amenity, has a causal effect on rents and house prices. Motivated by his work, we incorporate the human capital environment of a community in the rental yield regression using percent of college degree holders in a RED as a proxy. Again we do not expect more college degree holders will necessarily be more or less active in trading properties, hence this variable does not show up in the turnover regression. To put it differently, mortgage income ratio and percent of residents with college degrees are our instruments to identify the turnover equation.

Similarly, percent of young population and building age appear in vector Z but not X, hence help us to identify the rent-to-price equation. Young population is presumably more mobile. Higher mobility in the 25 to 44 age-group could lead to higher turnover rates in REDs with relatively higher concentrations of young households. Our discussions with the market practitioners confirm the observation that new REDs tend to be more actively traded. Thus, the age of the buildings is included in the turnover equation. RED size and household income enter both equations in natural logs to capture some nonlinearity.

5.1 Main Results

Our theoretical model has shown that both rent-to-price ratio and turnover rate are endogenous and they are simultaneously determined in the equilibrium. In that case, ordinary least squares (OLS) estimation of system (28) suffers from simultaneous equations bias (SEB). As Greene (2008, chapter 13) once remarked, "Although we can say with certainty that d_j (ordinary least squares estimator) is inconsistent (in the context of simultaneous-equations estimation), we cannot state how serious this problem is." In other words, we need to assess the magnitude of the SEB in the current context. To proceed, we show the OLS estimates in Column (1) and their corresponding standardized coefficients in Column (4) of Table 2. While the turnover rate in the rent-to-price equation is confirmed to be endogenous by a formal Hausman test, the endogeneity (or, exogeneity) of the rent-to-price ratio in the turnover equation is a borderline case.[64] Concerned with the power of the test, it is not clear whether the rental yield is truly exogenous. Notice also that while the theory is ambiguous on the sign of the correlations between the turnover rate and the rent-to-price ratio (see from (24) to (26)), OLS estimates deliver positive coefficients in both turnover ratio equation and rent-to-price ratio equation.

[64] The Hausman test is implemented equation by equation in four steps: (1) run a reduced-form regression of the potential endogenous variable on all exogenous variables (both X and Z) using the OLS estimator, (2) predict the residuals from the above regression, (3) estimate the structural equation with residuals from last step as an extra control, (4) test if the coefficient of the residual variable is zero. The t-statistic is *2.9* for the significance test of turnover's residual and *1.4* for that of rent-to-price ratio's residual. The critical value for *10%* relative bias toleration is about *1.6*.

[Table 2 about here.]

Given these results, we consider it appropriate to proceed with a consistent, full information estimator, which is the three-stage least squares (3SLS). But before that, we need to verify that our system of equations is identifiable. The order condition, which requires the number of exogenous variables excluded from one equation to be at least as large as the number of endogenous variables included in the equation, is clearly satisfied in the current context. And our system passes the rank condition test (Baum 2007) as well. A final specification test, the Hansen-Sargan test of over-identifying restrictions (Baum 2006),[65] confirms our instruments are all valid.

Since we have validated our approach, we now present more detailed results. Column (3) of Table 2 reports the 3SLS estimates *with small sample adjustment* on the covariance matrix of the regression residuals. Column (6), a list of corresponding standardized coefficients, gives the economic magnitude of our results. Consistent with the prediction of our theoretical model, investors of the less popular REDs are compensated with higher rental yields, probably in the form of a lower purchasing price. With an increase in the popularity index by one standard deviation, the rent-to-price ratio would decrease by *0.158* standard deviations. This is statistically significant at the 10% significance level. We also find a *negative relationship* between the rent-to-price ratio and the turnover rate. Other variables held constant, if the turnover rate increases by one standard deviation, the rent-to-price ratio will decline by *0.475* standard deviations. The estimate is statistically different from zero at the 5% significance level. However, we fail to find any significant effects on turnover rate of rental yield and popularity index in the lower panel of Table 2. Such disappointing result may be due to imperfections of our popularity index, our small sample size, and our still limited understanding or imperfect modeling of the house trading process.[66] Alternatively, the insignificant empirical result may reflect the theoretical result, which is ambiguous. More research is needed on this topic.

On top of providing an empirical verification of the theoretical model, our econometric model also identifies the empirical determinants of the RED-level rent-to-price ratio and the turnover rate. The coefficients of Ln(Income) in both equations are negative and highly significant (both economically and statistically). The upper panel of Table 2 shows that the rent-to-price ratio will fall by *0.794%* for every unit increase in Ln(Income) in the RED. An interpretation of this result is that there exist some amenities which attract people with higher income and at the same time unobservable to econometricians (for instance, the "harmony" and the degree of "quietness" of the

[65] The Hansen-Sargan statistic is *2.008*, which is distributed as Chi-squared with two degrees of freedom.

[66] For instance, housing transaction taxation (HTT) is assumed away in the model. In practice, HTT is nonlinear and might be important. Among others, see Leung et al. (2015) for a recent study.

environment). Hence, a RED with a higher level of "econometrician-unobservable amenities" would have a higher level of (average) income. As higher income residents are more reluctant to move (i.e. the opportunity cost of moving effect dominates the demand effect), REDs with higher income residents are likely to have more home-owners, which tends to drive up the house price more than the rent, resulting in a lower rent-to-price ratio. The turnover rate of a RED would be *3.693%* lower with a unit increase in the median household income. Real-estate work experience does not seem to trigger more trading. Yet despite that, the occupational advantage does have a significant positive impact on the valuation, as reflected by the positive and significant coefficient of the "Work in Real Estate Consulting" variable. This is consistent with the finding in the literature that real estate agents may have informational advantages over their clients (Levitt and Syverson 2008). While the previous year's housing price appreciation rate has no significant influence on the rent-to-price ratio, it has a significantly negative effect on the turnover rate. Transaction volume typically co-moves with property price.[67] A higher price appreciation rate in last year is probably associated with a higher turnover. If the new owners keep their units in the portfolio, turnover rate in the current period would decrease. Quantitatively, if the housing price growth rate of a particular RED was *10% higher than the mean* across all estates last year, then the current normalized transaction would decline by 1%. In contrast, the average rent growth rate in past five years, which can be interpreted as a measure of the potential for the housing units to generate income flow, significantly affects the rent-to-price ratio (at 5% level) but not the turnover rate. RED scale, the proposed measure of "thick market effect" (TME), is not significant in either regression, holding other variables constant. While it is still premature to reject the existence of TME with our relatively small sample size and relatively short sampling period, it is fair to say that its empirical relevance in the housing market remains to be proven. To conclude the paragraph, we find that nearby land auctions, which proxies for future substitutes in the neighborhood have a dampening effect on current sales volume, although the impact may only be marginally significant, with a p-value of *0.189*. Potential home owners may indeed keep their option to buy later in the face of nearby land auction.

The last set of results from the 3SLS is about the four instruments. In line with Clayton et al. (2009), we also find a significantly negative association between the mortgage burden and rental yield, holding other things constant. Thus, investors are willing to take out a heavier mortgage, even with a lower rental yield at the moment, perhaps expecting a larger capital gain in the future. However, since we do not have direct measure of the expected house price changes from the household survey, we can only leave the full explanation of this empirical finding to future research. We also find that communities with a higher level of human capital, measured by the per-

[67] The literature on the house price-transaction volume correlation is too large to be reviewed here. Among others, see Leung et al. (2002) for a study of such correlation in the Hong Kong housing market.

centage of college-educated residents, tend to have higher rent-to-price ratio. Finally, properties that are build more recently and REDs with a larger young population, who are considered more mobile, are more actively traded. All these four estimates are significant at least at the 10% significance level.

5.2 Robustness

In the previous section, we report our findings on the co-determination of the rent-to-price ratio and the turnover rate, as well as the empirical determinants of the two variables. In this section, we would address some statistical concerns and show that our results are indeed robust. First, we would address the issue of econometric method. As we have explained before, our theoretical model shows that the rent-to-price ratio and the turnover rate are jointly determined. *To faithfully verify this theoretical conclusion, we need an econometric method for a system of simultaneous equations.* In this regard, 3SLS is a natural candidate as it estimates our equations, (30) and (31), jointly. Another consistent but less efficient estimator for simultaneous equations models is two-stage least squares (2SLS). Methodologically, it is essentially the first two steps of 3SLS. However, treating the disturbances of equations as independent, it is a limited information estimator and hence less efficient. 2SLS forgoes some efficiency for the benefit of a less stringent data requirement. Therefore, it can serve a useful robustness check. As shown in Column (2) and (5) of Table 4, despite being less significant, the results are very similar to those in Column (3) and (6).

The second issue is the small sample size of this paper. We bootstrap and cross-validate the data.[68] The resampling operation applied here involves bootstrapping the 3SLS regression coefficients in a thousand replications. Visualizations of the resulting 1000 bootstrapped coefficients are provided by their density distributions in Figure 4 and the appendix. The vertical lines in the figures stand for the coefficient estimates from the actual data, i.e. those in Column (3) of Table 2. Since they are almost always in the middle of the distribution, there should be no substantial bias in our estimates. In Figure 4a, more than 94% of the bootstrapped coefficients for turnover fall in the range between -1.5 and zero. In both Figure 4b and 4d, the vast majority of the estimated coefficients for popularity index have the same negative signs as in Column (3) of Table 2. The estimated range of the coefficient of rental yield in the turnover equation is however too wide to be informative, as in the case of the actual sample. Coefficient distributions of other covariates from the bootstrapping exercises can be found in the appendix. Based on these findings, it seems that the small sample size is not issue.

[68] Both the bootstrap method and cross-validation are widely used in the literature. Among others, see Diebold (2016), Efron and Tibshirani (1993), Moore (2016).

[Figure 4 about here]

It is well known that the existence of outliers could potentially affect the empirical results, especially in a relatively small sample. To address this issue, we employ the leave-one-out cross validation (LOOCV) to detect influential outliers. The idea is simple. We repeatedly re-run the same regression model with all but one observations. We iterate this process throughout our sample. Visual representations of the resulting 130 vectors of coefficients are in Figure 5 and the appendix. Although leaving certain observations out does lead to some deviations from the full-sample estimates, there is no qualitative change in results. Thus, outlier does not seem to be an issue either.

[Figure 5 about here]

We also receive some feedback about our popularity index and hence we feel obligated to provide more discussion here. This research attempts to build a novel popularity index for housing. According to the literature (Diamantopoulos and Winklhofer 2001; Jarvis et al. 2003), measurement models are of two types, refective and formative. What the former does is to look for some latent variable which causes the observable effect indicators such as economic growth and poverty reduction. Multiple effect indicators could be traced back to a common latent variable. In contrast, the objective of a formative model is to summarize a set of causal indicators into a composite outcome index. Socioeconomic status, combining income, education, and occupation, is a typical example. Conceptually, a formative model fit our purpose better. So we form the popularity index as a linear combination of nine causal indicators with equal weights. Clearly, such a simple approach may raise concerns about how it might affect the robustness of our findings. In order to alleviate such concerns, in the spirit of LOOCV we remove one sub-indicator at a time, reformulate the index based on the remaining eight sub-indicators, and re-run the 3SLS regression. This process is repeated nine times for every sub-indicator and results in several *extra coefficient estimates for every control variable* of the regression. In Figure 6 and the appendix, these point estimates are plotted around the coefficients (the horizontal line) from regressions with default popularity index. Again the deviations are very mild, suggesting our results are robust to different weighting schemes on the sub-indicators. As an alternative index construction method, we also try the principal component analysis on our data in the appendix. However, we need to retain as many as six principal components to capture less than 80% of variances in our nine indicators, which does not seem to be a promising approach given our small sample size and 3SLS procedure. Last but not the least, we also incorporate some district level information in an extended specification and the results, which are reported in the appendix, do not significantly change.

[Figure 6 about here]

6 Concluding Remarks

Research in real estate has been growing since the Great Recession. Search models endogenize vacancy and trading volume and are therefore investigated more heavily in the theoretical literature.[69] On the other hand, empirical investigations on the rental yield and normalized transaction, and the interactions among them in the residential property market, are surprisingly scarce.[70] Previous studies focus on how the macroeconomic environment, regional market conditions, property attributes, and investor characteristics would affect the rent-to-price ratio. However, few studies examine how the rental yields are related to the characteristics of the household living in those housing units. Empirical works typically treat the rent-to-price ratio as the housing market counterpart of the earning-to-price ratio in the stock market, but few consider *how the ratio is determined in equilibrium*, and how it can be related to *economic fundamentals*. This paper attempts to complement the existing literature by providing a search-theoretic model and applying it to a unique RED-level data set in Asia, where housing units are (relatively) frequently traded.

Our theoretical model shows that both the rent-to-price ratio and the turnover rate can be related to each other, and to the level of popularity. Building on previous literature, our empirical work attempts to construct a popularity index as a proxy for the underlying good-match probability or the desire for owner-occupants on a RED that econometricians may not observe directly. Second, as housing units are either transacted in the sale market or rental market *but not both*, we adopt a matching estimator to create sale-rental pairs, where both units have the same "quality." Third, we adopt a 3SLS approach to solve the simultaneity issue. Our empirical results do suggest that the OLS estimates are very likely to be biased, which justifies our 3SLS estimation approach. We confirm a significantly negative effect of the turnover rate on the rent-to-price ratio. If the trading activities are more intensive, measured by a one-standard-deviation increase in the turnover rate, the associated rental yield would be 0.475 standard deviations lower. The two-stage least squares analysis reports similar results. These findings provide supportive evidence of a *liquidity effect in the property market*, which has been widely reported in the financial market research.

[69] Among others, see Wheaton (1990), Yavas, (1992a, 1992b), Krainer (2001), Albrecht et al. (2007), Genesove and Han (2012), Diaz and Jerez (2013), Anglin (2006), Anglin and Gao (2011), Lin and Vandell (2007), Lin and Liu (2008), and Tse (2011).

[70] Exceptions include Kwok and Tse (2006), which test the price premium for condominiums easier to sell in a submarket of Hong Kong, and Liu and Qian (2012), which confirm an illiquidity compensation in property returns based on a sample of 24 office markets in the U.S.

Our constructed popularity index is also important in explaining the rental yield. A RED with a popularity index one standard deviation higher has its rent-to-price ratio 0.158 standard deviations lower. This effect is also in line with our theoretical model. Our *on-the-house-search model*, which can be considered as the housing market counterpart of the on-the-job-search model, the popularity index, and our empirical estimates of brand premium for different developers, may independently be of interest for market practitioners and academics.

In addition, we identify the empirical determinants of the rental yield or the turnover rate. We find that income has a dampening effect on both the rental yield and the turnover rate. In our sample, the opportunity cost of moving appears to dominate the wealth effect. We also find evidence of a "dichotomous structure" in the housing market: the demographic structure, and lagged price appreciation of a RED tend to affect its turnover rate, while popularity, human capital environment, mortgage burden, and long run rent growth have more influence on the rental yield. We employed several statistical methods to establish the robustness of our results. We believe our results have both academic and practical value.

This study provides a search-theoretic model which endogenizes both the rent-to-price ratio and the transaction volume. The rent-to-price equation is empirically confirmed and we make some additional empirical observations. Nevertheless, there is room for further improvement. While we can identify the empirical determinants of the cross-sectional differences in turnover rates among different REDs, we fail to find any evidence that the popularity index or the rent-to-price ratio play a role in explaining the turnover rate difference. Future research, therefore, can improve on the popularity index. In fact, there may exist two different indices, one is more relevant for the cross-sectional rental yield difference and the other for the turnover rate difference. In theoretical terms, a more sophisticated model of the housing market could also be constructed.[71]

[71] For instance, see Leung and Tse (2016), Luo (2014).

References

[1] Albrecht, J., A. Anderson, E. Smith, and S. Vroman, 2007, Opportunistic matching in the housing market, International Economic Review 42, 641-664.

[2] Albrecht, J., P. Gautier and S. Vroman, 2006, Equilibrium directed search with multiple applications, Review of Economic Studies 73, 869-891.

[3] Ambrose, B.W., and H.O. Nourse, 1993, Factors influencing capitalization rates, Journal of Real Estate Research 8, 221-237.

[4] Amihud, Y., H. Mendelson and L.H. Pedersen, 2013, Market Liquidity: Asset Pricing, Risk, and Crises, Cambridge: Cambridge University Press.

[5] An, X., and Y. Deng, 2009, A structural model for capitalization rate, San Diego State University, mimeo.

[6] Anglin, P., 2006, Value and liquidity under changing market conditions, Journal of Housing Economics 15, 293-304.

[7] Anglin, P., and Y. Gao, 2013, Integrating illiquid assets into the portfolio decision process, Real Estate Economics 39, 277-311.

[8] Banerjee, S., V.A. Gatchev, and P.A. Spindt, 2007, Stock market liquidity and firm dividend policy, Journal of Financial and Quantitative Analysis 42, 369-397.

[9] Baum, C.F., Schaffer, M.E., Stillman, S., and V. Wiggins, 2006, Overid: Stata module to calculate tests of overidentifying restrictions after ivreg, ivreg2, ivprobit, ivtobit, reg3. http://ideas.repec.org/c/boc/bocode/s396802.html

[10] Baum, C.F., 2007, Checkreg3: Stata module to check identification status of simultaneous equations system. http://ideas.repec.org/c/boc/bocode/s456877.htm

[11] Bekaert, G., C.R. Harvey, and C. Lundblad, 2007, Liquidity and expected returns: Lessons from emerging markets, Review of Financial Studies 20, 1783-1831.

[12] Blanchard, O.J., and P.A. Diamond, 1990, The cyclical behavior of the gross flows of U.S. workers, Brooking Papers on Economic Activity 2, 85-155.

[13] Brunnermeier, M. and M. Oehmke, 2013, Bubbles, financial crises, and systemic risk, Princeton University, mimeo.

[14] Burdett, K. and D. Mortensen, 1998, Equilibrium wage differentials and employer size, International Economic Review 39, 257-274.

[15] Campbell, S.D., M.A. Davis, J. Gallin, and R.F. Martin, 2009, What moves housing markets: A variance decomposition of the rent-price ratio, Journal of Urban Economics 66, 90-102.

[16] Chambers, M.; C. Garriga and D. Schlagenhauf, 2009, Accounting for changes in the homeownership rate, International Economic Review 50, 677-726.

[17] Chervachidze, S., and W. Wheaton, 2013, What determined the great cap rate compression of 2000-2007, and the dramatic reversal during the 2008-2009 financial crisis? Journal of Real Estate Finance and Economics 46, 208-231.

[18] Clayton, J., D.C. Ling, and A. Naranjo, 2009, Commercial real estate valuation: Fundamentals versus investor sentiment, Journal of Real Estate Finance and Economics 38, 5-37.

[19] Dhar, P., and S.L. Ross, 2012, School district quality and property values: Examining differences along school district boundaries, Journal of Urban Economics 71, 18-25.

[20] Diamantopoulos, A., and H.M. Winklhofer, 2001, Index construction with formative indicators: An alternative to scale development, Journal of Marketing Research 38, 269-277.

[21] Diamond, P., 1982a, Wage determination and efficiency in search equilibrium, Review of Economic Studies 49, 217-27.

[22] Diamond, P., 1982b, Aggregate demand management in search equilibrium, Journal of Political Economy 90, 881-94.

[23] Diaz, A., and B. Jerez, 2013, House prices, sales, and time on the market: A search-theoretic framework, International Economic Review 54, 837-872.

[24] Diebold, F., 2016, Dynamic econometrics with big data, Society of Financial Econometrics (SoFiE), invited lecture.

[25] Efron, B., and R.J. Tibshirani, 1993, An Introduction to the Bootstrap, New York: Chapman and Hall.

[26] Evans, R.D., 1990, A transfer function analysis of real estate capitalization rates, Journal of Real Estate Research 5, 371-379.

[27] Faig, M., and B. Jerez, 2006, Inflation, prices, and information in competitive search, Advances in Macroeconomics 6, 1-34, ISSN (Online) 1534-6013, DOI: 10.2202/1534-6013.1313.

[28] Figlio, D.N., and M.E. Lucas, 2004, What's in a grade? School report cards and the housing market, American Economic Review 94, 591-604.

[29] Financial Times, 2015, Fears of bubble rise as property investment volumes hit record, available at http://www.ft.com/cms/s/0/3439fc8c-99b1-11e4-93c1-00144feabdc0.html#axzz3SjshUzs2

[30] Fisher, J., D. Gatzlaff, D. Geltner and D. Haurin, 2003, Controlling for impact of variable liquidity in commerical real estate price indices, Real Estate Economics 31, 269-303.

[31] Fox, R., and P. Tulip, 2014, Is housing overvalued? Reserve Bank of Australia, mimeo.

[32] Genesove, D., and L. Han, 2012, Search and matching in the housing market, Journal of Urban Economics 72, 31-45.

[33] Gjerstad, S., and V. Smith, 2014, Rethinking Housing Bubbles: The Role of Household and Bank Balance Sheets in Modeling Economic Cycles, Cambridge: Cambridge University Press.

[34] Glaeser, E., and J. Gyourko, 2007. Arbitrage in housing markets, NBER Working Papers 13704.

[35] Global Property Guide, 2010, How to avoid buying into a bubble? available at http://www.globalpropertyguide.com/real-estate-school/How-to-avoid-buying-into-a-bubble

[36] Green, R., 2015, Remarks on the globalization of real estate and its consequences, presented at the 2015 Global Real Estate Submit.

[37] Green, R., and S. Malpezzi, 2003, A Primer on U.S. Housing Markets and Housing Policy, Washington, D.C.: Urban Institute Press.

[38] Greene, W. H., 2008, Econometric Analysis, 6th ed., Boston: Prentice Hall.

[39] Griliches, Z., 1971, Hedonic price indexes revisited, in Z. Griliches, ed., Price Indexes and Quality Change: Studies in New Methods of Measurement, Cambridge, MA: Harvard Univeristy Press.

[40] Guerrieri, V., 2008, Inefficient unemployment dynamics under asymmetric information, Journal of Political Economy 116, 667-708.

[41] Halket, J., and M. Pignatti, 2013, homeownership and the scarcity of rentals, paper presented at HULM-Atlanta Fed conference.

[42] Hanushek, E., Lavy, V. and K. Hitomi, 2008, Do students care about school quality? determinants of dropout behavior in developing countries, Journal of Human Capital 2, 69-105.

[43] Hanushek, E., Rivkin, S. and L. Taylor, 1996, Aggregation and the estimated effects of school resources. Review of Economics and Statistics 78, 611-27.

[44] Hanushek, E., and L. Woessmann, 2012, Do better schools lead to more growth? Cognitive skills, economic outcomes, and causation, Journal of Economic Growth 17, 267-321.

[45] Hanushek, E., and K. Yilmaz, 2007, The complementarity of Tiebout and Alonso, Journal of Housing Economics 16, 243-261.

[46] Hanushek, E., and K. Yilmaz, 2013, Schools and location: Tiebout, Alonso, and governmental finance policy, Journal of Public Economic Theory 15, 829-855.

[47] Hawkins, W., 2013, Competitive search, efficiency, and multiworker firms, International Economic Review 54, 219-251.

[48] Huang, D. J., and C.K.Y. Leung, 2014, Liquidity premium in the housing market: A matching estimator approach, paper presented in AREUEA meeting.

[49] Jarvis, C. B., S.B. MacKenzie, and P.M. Podsakoff, 2003, A critical review of construct indicators and measurement model misspecification in marketing and consumer research, Journal of Consumer Research 30, 199-218.

[50] Jolliffe, I., 2005, Principal Component Analysis, New York: Springer-Verlag.

[51] Jovanovic, B., 1989, Observable implications of models with multiple equilibria, Econometrica 57, 1431-37.

[52] Jovanovic, B. and R. Rosenthal, 1988. Anonymous sequential games, Journal of Mathematical Economics, 17, 77-87.

[53] Jud, D.G., and D.T. Winkler, 1995, The capitalization rate of commercial properties and market returns, Journal of Real Estate Research 10, 509-518.

[54] Kashiwagi, M., 2014, A search-theoretic model of the rental and homeownership markets, Journal of Housing Economics 26, 33-47.

[55] Krainer, J., 2001, A theory of liquidity in residential real estate markets, Journal of Urban Economics 49, 32-53.

[56] Krause, M., D. Lopez-Salido, & Lubik, T., 2008, Do search frictions matter for inflation dynamics? European Economic Review 52, 1464-1479.

[57] Krause, M., and T. Lubik, 2010, Instability and indeterminacy in a simple search and matching model, Economic Quarterly, Federal Reserve Bank of Richmond, issue 3Q, 259-272.

[58] Kwok, H., and C.Y. Tse, 2006, Estimating liquidity effects in the housing market, University of Hong Kong, mimeo.

[59] Lam, W.H.K. and C.Y. Cheung, 2000, Pedestrian speed/flow relationships for walking facilities in Hong Kong, Journal of Transportation Engineering 126, 343-349.

[60] Leung, C.K.Y., G.C.K. Lau and Y.C.F. Leong, 2002, Testing alternative theories of the property price-trading volume correlation, Journal of Real Estate Research 23, 253-63.

[61] Leung, C.K.Y., Y.C.F. Leong and I.Y.S. Chan, 2002, TOM: Why isn't price enough? International Real Estate Review 5, 91-115.

[62] Leung, C.K.Y., Y.C.F. Leong and K.S.K. Wong, 2006, Housing price dispersion: An empirical investigation, Journal of Real Estate Finance and Economics 32, 357-385.

[63] Leung, C.K.Y., T.T.C. Leung and B.K.P. Tsang, 2015, Tax-driven bunching of housing market transactions: The case of Hong Kong, International Real Estate Review 18, 473-501.

[64] Leung, C.K.Y., and E.C.H. Tang, 2012, Comparing two financial crises: The case of Hong Kong real estate markets, in A. Bardhan, R. Edelstein and C. Kroll, ed., Global Housing Markets: Crises, Institutions and Policies, New York: John Wiley & Sons.

[65] Leung, C.K.Y., and E.C.H. Tang, 2015, Availability, affordability and volatility: The case of Hong Kong housing market, International Real Estate Review 18, 383-428.

[66] Leung, C.K.Y., and C.Y. Tse, 2016, Flipping in the housing market, forthcoming in Journal of Economic Dynamics and Control.

[67] Leung, C.K.Y., and J. Zhang, 2011, "Fire sales" in housing market: Is the house-search process similar to a theme park visit? International Real Estate Review 14, 311-329.

[68] Levitt, S., and C. Syverson, 2008, Market distortions when agents are better informed: The value of information in real estate transactions, Review of Economics and Statistics 90, 599-611.

[69] Lin, Z., and Y. Liu, 2008, Real estate returns and risk with heterogeneous investors, Real Estate Economics 36, 753-776.

[70] Lin, Z., and K. D. Vandell, 2007, Illiquidity and pricing biases in the real estate market, Real Estate Economics 35, 291-330.

[71] Lo, A. W., and J. Wang, 2009, Stock market trading volume, in Y. Ait-Sahalia and L. P. Hansen, ed., Handbook of Financial Econometrics, New York: Elsevier.

[72] Lubik, T., 2012, Aggregate labour market dynamics in Hong Kong, Pacific Economic Review 17, 257-279.

[73] Luo, G., 2014, Household credit histories: Slow jobless recovery from mortgage crises. Northwestern University, mimeo.

[74] Malpezzi, S., 2008, Hedonic pricing models: A selective and applied review, in T. O'Sullivan and K. Gibb, ed., Housing Economics and Public Policy: Essays in honor of Duncan Maclennan, Oxford: Blackwell Science Ltd.

[75] Malpezzi, S., 2014, On the rental market of the United States, University of Wisconsin, Madison, mimeo.

[76] Mas-Colell, A., M. Whinston and J. Green, 1995, Microeconomic Theory, Oxford: Oxford University Press.

[77] Mayer, C.J., 1995, A model of negotiated sales applied to real estate auctions, Journal of Urban Economics 38, 1-22.

[78] McMillen, D.P., 2012, Repeat sales as a matching estimator, Real Estate Economics 40, 745-73.

[79] Moen, E., and A. Rosen, 2004, Does poaching distort training, Review of Economic Studies 71, 1143-1162.

[80] Moore, A., 2016, Cross-validation for detecting and preventing overfitting, Carnegie Mellon University, mimeo.

[81] Mortensen, D., and C. Pissarides, 1994, Job creation and job destruction in the theory of unemployment, Review of Economic Studies 61, 397-415.

[82] Munneke, H., J. Ooi, C. F. Sirmans and G. Turnbull, 2012, Real estate agents, house prices, and liquidity, mimeo.

[83] Ngai, L. R., and S. Tenreyro, 2014, Hot and cold seasons in the housing market, American Economic Review 104, 3991-4026.

[84] Nguyen-Hoang, P., and J. Yinger, 2011, The capitalization of school quality into house values: A review, Journal of Housing Economics 20, 30-48.

[85] Novy-Marx, R., 2009, Hot and cold markets, Real Estate Economics 37, 1-22.

[86] Peng, L., 2013, How Integrated Is the Commercial Real Estate Asset Market? Evidence from Transaction Cap Rates, University of Colorado, mimeo.

[87] Petrongolo, B., and C. Pissardes, 2001, Looking into the black box: A survey of the matching function, Journal of Economic Literature 39, 390-431.

[88] Piazzesi, M., and M. Schneider, 2009, Momentum Traders in the Housing Market: Survey Evidence and a Search Model, American Economic Review Papers and Proceedings 99, 406-411.

[89] Pissarides, C., 1994, Search unemployment with on-the-job search, Review of Economic Studies 61, 457-475.

[90] Pissarides, C., 2000, Equilibrium Unemployment Theory, 2nd ed., Cambridge, MA: MIT Press.

[91] Plazzi, A., W. Torous, and R. Valkanov, 2010, Expected returns and expected growth in rents of commercial real estate, Review of Financial Studies 23, 3469-3519.

[92] Qian, W., and P. Liu, 2012, Does (and What) Illiquidity Matter for Real Estate Prices? Measure and Evidence, mimeo.

[93] Rubinstein, A., 1990, Bargaining and Markets, New York: Emerald Group Publishing Limited.

[94] Rutherford, R. C., Springer, T. and A. Yavas, 2005, Conflicts between principals and agents: Evidence from real estate brokerage, Journal of Financial Economics 76, 627-665.

[95] Sivitanidou, R., and P. Sivitanides, 1999, Office capitalization rates: Real estate and capital market influences, Journal of Real Estate Finance and Economics 18, 297-322.

[96] Rivkin, S., E. Hanushek and J.F. Kain, 2005, Teachers, schools, and academic achievement, Econometrica 73, 417-458.

[97] Sargent, T., 1979, Macroeconomic Theory, New York: Academic Press.

[98] Sato, Y., and W. Xiao, 2015, Land development, search frictions and city structure, Regional Science and Urban Economics, 50, 63-76.

[99] Shapiro, J. M., 2006, Smart cities: Quality of life, productivity, and the growth effects of human capital, Review of Economics and Statistics 88, 324–335.

[100] Sims, C.A., 1980, Macroeconomics and reality, Econometrica 48, 1-48.

[101] Sims, C.A., 2010, But economics is not an experimental science. Journal of Economic Perspectives 24, 59-68.

[102] Smith, B. A., and W.P. Tesarek, 1991, House prices and regional real estate cycles: Market adjustments in Houston, Real Estate Economics 19, 396-416.

[103] StataCorp, L., 2013, Stata Base Reference Manual, College Station, TX: Stata Press.

[104] Todd, P., 2006, Matching estimators, University of Pennsylvania, mimeo.

[105] Tse, C. Y., 2004, Search frictions, market power, and long-run growth, Journal of Economic Theory 116, 323-346.

[106] Tse, C. Y., 2011, The Spatial origin of commerce, International Economic Review 52, 349-377.

[107] Tse, C. Y., C.K.Y. Leung and W.Y.F. Chan, 2002, Unemployment and Vacancy in the Hong Kong Labour Market, Applied Economics Letters 9, 221-29.

[108] Wheaton, W.C., 1990, Vacancy, search, and prices in a housing market matching model, Journal of Political Economy 98, 1270.

[109] Williams, J.T., 1995, Pricing real assets with costly search, Review of Financial Studies 8, 55-90.

[110] Wong, E., A. Tsang, S. Kong, 2016, How does loan-to-value policy strengthen resilience of banks to property price shocks – evidence from Hong Kong, International Real Estate Review 19, 120-149.

[111] Wooldridge, J., 2010, Econometric Analysis of Cross Section and Panel Data, 2nd edition, Cambridge: MIT Press.

[112] Wu, J., J. Gyourko and Y. Deng, 2012, Evaluating conditions in major Chinese housing markets, Regional Science and Urban Economics 42, 531-543.

[113] Xiao, W., 2014, Search frictions, unemployment, and housing in cities: Theory and policies, Journal of Regional Science, 54, 422–449.

[114] Yavas, A., 1992a, Marketmakers versus matchmakers, Journal of Financial Intermediation 2, 33-58.

[115] Yavas, A., 1992b, A simple search and bargaining model of real estate markets, Real Estate Economics 20, 533-548.

Table 1a. Summary Statistics

Variables	count	mean	sd	min	max
Mean R/P of the RED (%)	132	3.772	0.476	2.629	5.631
Median R/P of the RED (%)	132	3.681	0.464	2.596	5.308
Turnover Rate in Sample Period (%)	132	7.181	3.160	0.484	18.972
Average Rent Growth Rate in Past 5 YRs (%)	130	4.294	4.022	-13.096	20.626
Last Year HP Growth Rate (%)	132	21.564	5.889	2.231	35.803
Population Resided in the RED (Persons)	132	7640.803	6978.537	2263	39361
Young Population (%)	132	40.139	7.994	23.055	62.218
Middle Age Population (%)	132	29.314	5.451	12.903	42.328
Median Age of the Residents (Years)	132	39.008	2.390	34.100	46.100
Married Population (%)	132	56.921	3.220	48.117	65.096
Mandarin Speaker (%)	132	3.933	2.620	0.500	12.900
College Education Level (%)	132	46.477	11.429	20.900	75.500
Household Monthly Median Income (2005 HKD)	132	37907.360	17274.432	13090.000	133450.000
Median Monthly Mortgage Payment (2005 HKD)	132	8875.867	4691.953	3400.000	30132.500
Mortgage Income Ratio (%)	132	20.345	2.647	14.900	28.500
Home Ownership Rate in the RED (%)	132	76.314	10.111	41.300	94.000
RED Popularity	132	6.249	1.277	2.948	9.018
Average Net Gross Ratio (%)	132	78.090	4.710	63.953	88.654
Average Building Age (Years)	132	19.214	8.701	2.933	41.121
Estate Scale	132	2899.932	2474.556	900.000	15836.000
Distance to CBD (Meters)	132	12168.281	6775.361	1152.100	25431.000
Distance to MTR (Meters)	132	380.302	151.925	52.284	500.000
Distance to Waterfront (Meters)	132	381.629	158.259	34.000	500.000
# Top 100 Secondary School in the Same District	132	5.742	2.995	0	13
Nearby Land Auction (Square Meters)	132	6140.515	15539.870	0	76167

Table 1b. Hedonic Pricing: Developers' Ranking

VARIABLES	Ln(Price)
Cheung Kong Holdings	-0.050**
	[0.038]
Chinachem Group	-0.055**
	[0.013]
Hang Lung Properties	-0.037
	[0.182]
Henderson Land Development	-0.008
	[0.710]
MTR Properties	-0.035
	[0.257]
Nan Fung Group	0.011
	[0.725]
New World Development	0.021
	[0.381]
Sino Group	0.043
	[0.314]
Sun Hung Kai Properties	0.053***
	[0.008]
Constant	12.843***
	[0.000]
Housing Structural Characteristics	YES
Neighborhood Characteristics	YES
Location Characteristics	YES
Observations	29,647
Adjusted R-squared	0.813
District FE	YES
Clustered Std Errors	YES
F	124.2

pval in brackets

*** p<0.01, ** p<0.05, * p<0.1

Table 1c. Popularity Ranking of Real Estate Development (RED)

Rank	Estate	Popularity Index	District
1	HARBOUR PLACE	9.018	Kowloon City
2	HARBOUR GREEN	8.814	Yau Tsim Mong
3	OCEAN SHORES	8.640	Sai Kung
4	HENG FA CHUEN	8.611	Eastern
5	PARK AVE	8.248	Yau Tsim Mong
6	BELCHER'S	8.211	Central and Western
7	LAGUNA CITY	8.209	Kwun Tong
8	SORRENTO	8.191	Yau Tsim Mong
9	LOHAS PARK	8.189	Sai Kung
10	GRAND PROMENADE	8.138	Eastern
11	ISLAND HARBOURVIEW	8.115	Yau Tsim Mong
12	TIERRA VERDE	8.101	Kwai Tsing
13	CITY GDN	8.085	Eastern
14	WATERFRONT	7.961	Yau Tsim Mong
15	VILLA ESPLANADA	7.905	Kwai Tsing
16	ROYAL PENINSULA	7.663	Kowloon City
17	VILLA ATHENA	7.646	Sha Tin
18	TAIKOO SHING	7.645	Eastern
19	GRAND WATERFRONT	7.588	Kowloon City
20	LAKE SILVER	7.584	Sha Tin
21	MERTON	7.553	Central and Western
22	GREENFIELD GDN	7.551	Kwai Tsing
23	SUNSHINE CITY	7.543	Sha Tin
24	TELFORD GDN	7.492	Kwun Tong
25	PARK CENTRAL	7.454	Sai Kung
26	ISLAND RESORT	7.430	Eastern
27	LAGUNA VERDE	7.427	Kowloon City
28	METRO HARBOUR VIEW	7.418	Yau Tsim Mong
29	EAST POINT CITY	7.366	Sai Kung
30	COASTAL SKYLINE	7.344	Islands
31	YOHO TOWN	7.324	Yuen Long
32	VISTA PARADISO	7.312	Sha Tin
33	RESIDENCE BEL-AIR	7.302	Southern
34	LIBERTE	7.268	Sham Shui Po
35	PARK ISLAND	7.265	Tsuen Wan
36	METRO CITY	7.263	Sai Kung
37	KINGSWOOD VILLAS	7.046	Yuen Long
38	TUNG CHUNG CRESCENT	7.021	Islands
39	METRO TOWN	7.019	Sai Kung
40	GRANDIOSE	7.001	Sai Kung
41	GRAND HORIZON	6.984	Kwai Tsing

42	MANHATTAN HILL	6.965	Sham Shui Po
43	DISCOVERY PARK	6.896	Tsuen Wan
44	SOUTH HORIZONS	6.842	Southern
45	PARCVILLE	6.830	Yuen Long
46	CENTRAL PARK	6.829	Yau Tsim Mong
47	RESIDENCE OASIS	6.819	Sai Kung
48	SEAVIEW CRESCENT	6.805	Islands
49	SEA CREST VILLA	6.775	Tsuen Wan
50	PROVIDENT CTR	6.755	Eastern
51	SCENEWAY GDN	6.724	Kwun Tong
52	LA CITE NOBLE	6.703	Sai Kung
53	PACIFICA	6.693	Sham Shui Po
54	AQUAMARINE	6.679	Sham Shui Po
55	AEGEAN COAST	6.666	Tuen Mun
56	NAN FUNG PLAZA	6.638	Sai Kung
57	BANYAN GDN	6.633	Sham Shui Po
58	BELLAGIO	6.589	Tsuen Wan
59	CENTRAL PARK TWRS	6.587	Yuen Long
60	TSING YI GDN	6.554	Kwai Tsing
61	HANFORD GDN	6.490	Tuen Mun
62	RAMBLER CREST	6.454	Kwai Tsing
63	SKY TWR	6.425	Kowloon City
64	PALAZZO	6.413	Sha Tin
65	HONG KONG GOLD COAST	6.354	Tuen Mun
66	LEI KING WAN	6.344	Eastern
67	VISION CITY	6.301	Tsuen Wan
68	ROYAL ASCOT	6.289	Sha Tin
69	MIAMI BEACH TWRS	6.274	Tuen Mun
70	PARKLAND VILLAS	6.115	Tuen Mun
71	RIVIERA GDN	6.107	Tsuen Wan
72	SERENITY PLACE	6.104	Sai Kung
73	SUMMIT TERR	6.053	Tsuen Wan
74	SERENITY PARK	6.048	Tai Po
75	CITY ONE SHATIN	5.998	Sha Tin
76	NERINE COVE	5.907	Tuen Mun
77	RHINE GDN	5.901	Tsuen Wan
78	DIS. BAY	5.890	Islands
79	WHAMPOA GDN	5.890	Kowloon City
80	SUN YUEN LONG CTR	5.863	Yuen Long
81	CHELSEA COURT	5.849	Tsuen Wan
82	RAVANA GDN	5.797	Sha Tin
83	QUEEN'S TERR	5.787	Central and Western
84	CASTELLO	5.771	Sha Tin
85	TAI HING GDNS	5.705	Tuen Mun

86	GRAND DEL SOL	5.686	Yuen Long
87	BELVEDERE GDN	5.658	Tsuen Wan
88	UPTOWN PLAZA	5.621	Tai Po
89	TAK BO GDN	5.620	Kwun Tong
90	CHELSEA HTS	5.577	Tuen Mun
91	WELL ON GDN	5.576	Sai Kung
92	SERENO VERDE	5.540	Yuen Long
93	SHA TIN CTR	5.536	Sha Tin
94	VIANNI COVE	5.502	Yuen Long
95	BAGUIO VILLA	5.404	Southern
96	CHI FU FA YUEN	5.370	Southern
97	WONDERLAND VILLAS	5.315	Kwai Tsing
98	DAWNING VIEWS	5.286	North
99	JUBILEE GDN	5.258	Sha Tin
100	SUN TUEN MUN CTR	5.256	Tuen Mun
101	SHERWOOD	5.233	Tuen Mun
102	PIERHEAD GDN	5.220	Tuen Mun
103	LIDO GDN	5.211	Tsuen Wan
104	CARIBBEAN COAST	5.191	Islands
105	MAYFAIR GDNS	5.183	Kwai Tsing
106	AVON PARK	5.154	North
107	GARDEN RIVERA	5.144	Sha Tin
108	FLORA PLAZA	5.105	North
109	AMOY GDN	5.073	Kwun Tong
110	MEI FOO SUN CHUEN	5.027	Sham Shui Po
111	FANLING CTR	5.014	North
112	WHAMPOA EST	5.006	Kowloon City
113	TSUEN KING GDN	4.964	Tsuen Wan
114	TUEN MUN TOWN PLAZA	4.863	Tuen Mun
115	LUK YEUNG SUN CHUEN	4.800	Tsuen Wan
116	BELAIR MONTE	4.780	North
117	ALLWAY GDN	4.752	Tsuen Wan
118	ABERDEEN CTR	4.721	Southern
119	BRAEMAR HILL MANS	4.490	Eastern
120	TAI PO CTR	4.447	Tai Po
121	JUBILANT PLACE	4.380	Kowloon City
122	BELAIR GDNS	4.314	Sha Tin
123	FANLING TOWN CTR	4.304	North
124	GOODVIEW GDN	4.170	Tuen Mun
125	GOLDEN LION GDN	4.100	Sha Tin
126	GOLDEN LION GDN	4.013	Sha Tin
127	PICTORIAL GDN	4.005	Sha Tin
128	NEW KWAI FONG GDN	3.932	Kwai Tsing
129	LUCKY PLAZA	3.902	Sha Tin

130	TSUEN WAN CTR	3.853	Tsuen Wan
131	HONG KONG GDN	3.373	Tsuen Wan
132	WALDORF GDN	2.948	Tuen Mun

Table 2: Main Results

	Coef.			Beta Coef.		
	OLS	2SLS	3SLS	OLS	2SLS	3SLS
	(1)	(2)	(3)	(4)	(5)	(6)
Equation: R/P						
Turnover Rate %	0.016	-0.064*	-0.072**	0.106	-0.427*	-0.475**
Popularity of the RED	-0.075*	-0.062	-0.068*	-0.174*	-0.144	-0.158*
College Education Level %	0.003	0.004	0.007*	0.070	0.093	0.161*
Ln(Income)	-0.450***	-0.724***	-0.794***	-0.378***	-0.609***	-0.668***
Mortgage Income Ratio %	-0.046***	-0.040***	-0.037**	-0.256***	-0.221***	-0.207**
Work in Real Estate Consulting %	0.028*	0.044**	0.044**	0.118*	0.182**	0.182**
Last Year HP Growth Rate %	0.001	-0.012	-0.012	0.008	-0.148	-0.148
Rent Growth Rate in Past 5 YRs	0.012	0.020*	0.020**	0.097	0.165*	0.171**
Ln(RED Scale)	0.106*	0.080	0.078	0.135*	0.102	0.100
Land Auction	-0.007	-0.014*	-0.015	-0.061	-0.120*	-0.123
Observations	130	130	130	130	130	130
RMSE	0.351	0.401	0.413	0.737	0.842	0.867
F/Chi-squared	14.42	121.42	80.24	14.42	121.42	80.24
R-squared	0.479	0.258	0.213	0.479	0.258	0.213
Equation: Turnover						
Mean R/P of the RED	1.540*	-1.525	-1.188	0.232*	-0.230	-0.179
Popularity of the RED	-0.172	-0.330	-0.296	-0.060	-0.115	-0.103
Young Population %	0.093**	0.098**	0.070**	0.234**	0.247**	0.178**
Ln(Income)	-2.141***	-3.578**	-3.693**	-0.271***	-0.454**	-0.468**
Work in Real Estate Consulting %	0.090	0.198	0.211	0.057	0.125	0.133
Last Year HP Growth Rate %	-0.089**	-0.113**	-0.102**	-0.166**	-0.210**	-0.191**
Rent Growth Rate in Past 5 YRs	0.077	0.115	0.114	0.098	0.147	0.145
Average Building Age	-0.103***	-0.059	-0.093**	-0.284***	-0.162	-0.257**
Ln(RED Scale)	0.102	0.342	0.285	0.020	0.066	0.055
Land Auction	-0.060	-0.097	-0.092	-0.076	-0.122	-0.115
Observations	130	130	130	130	130	130
RMSE	2.579	2.681	2.644	0.816	0.849	0.837
F/Chi-squared	6.10	75.59	61.56	6.10	75.59	61.56
R-squared	0.389	0.278	0.298	0.389	0.278	0.298

Note: OLS and 2SLS estimations are performed equation by equation. Estimates of standard errors are adjusted to be robust to misspecifications arising from neglecting heteroscedasticity. 3SLS estimates the two structural equations simultaneously. Small sample adjustment is made on the covariance matrix of the regression residuals. Standard errors and intercepts are suppressed to save space. The stars *, **, and *** indicate the significance level at 10%, 5%, and 1% respectively.

Figure 3a Popularity Index and Housing Price Growth

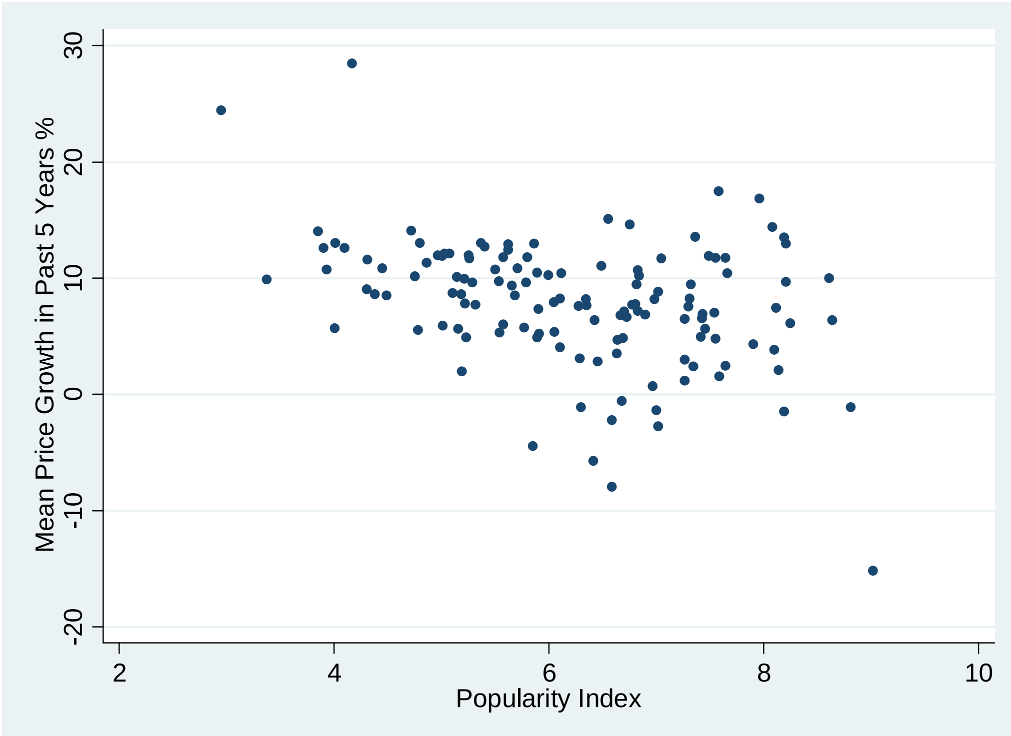


Figure 3b Popularity Index and Housing Price Volatility

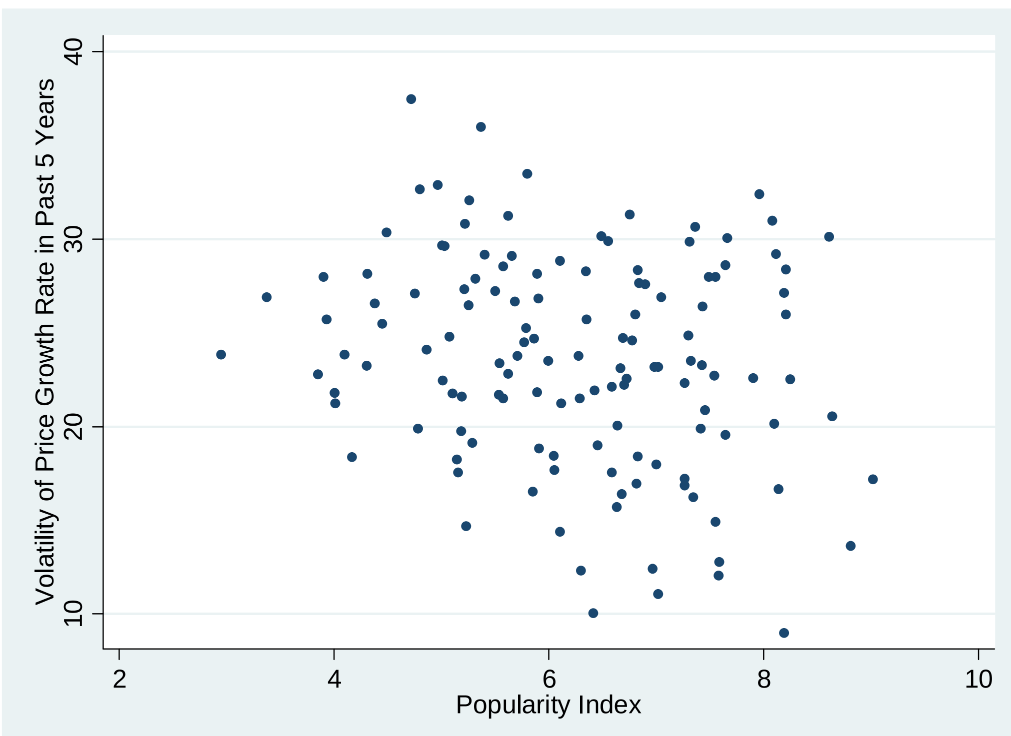
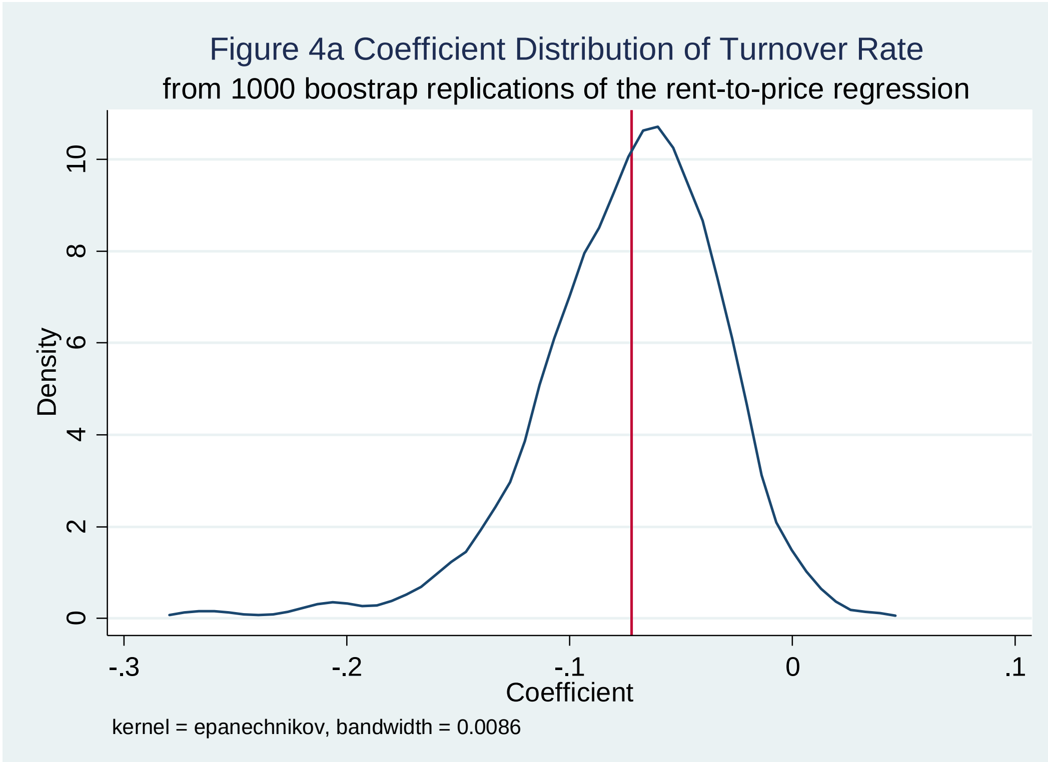
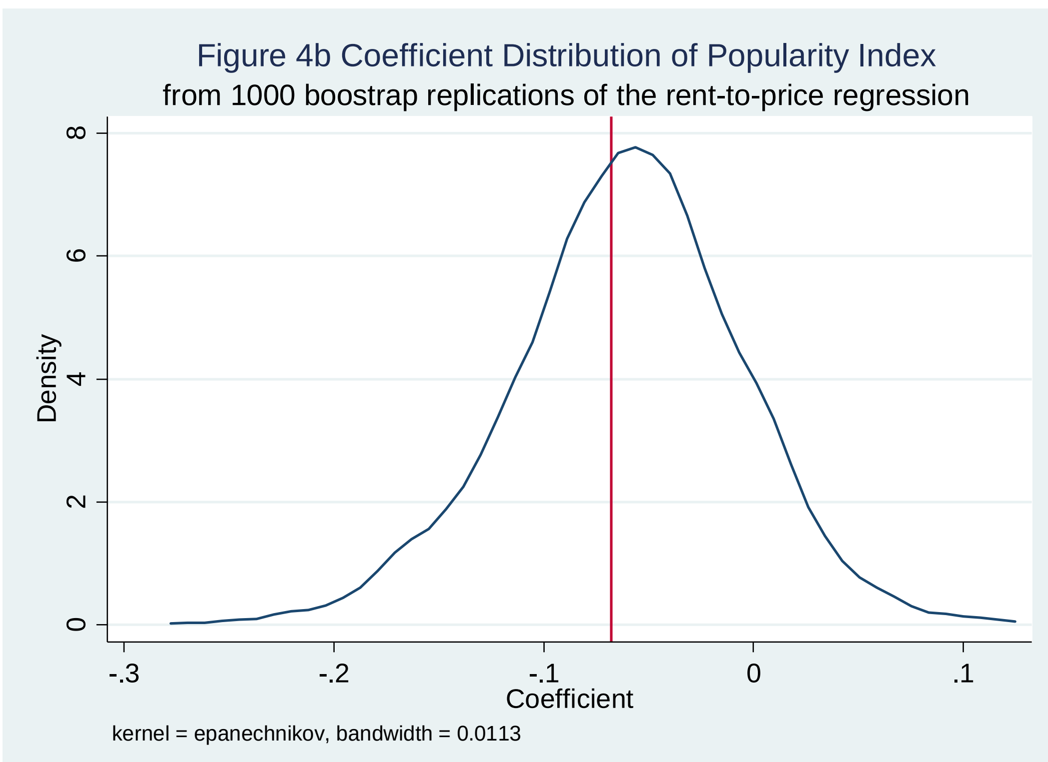


Figure 4 Bootstrap Replications





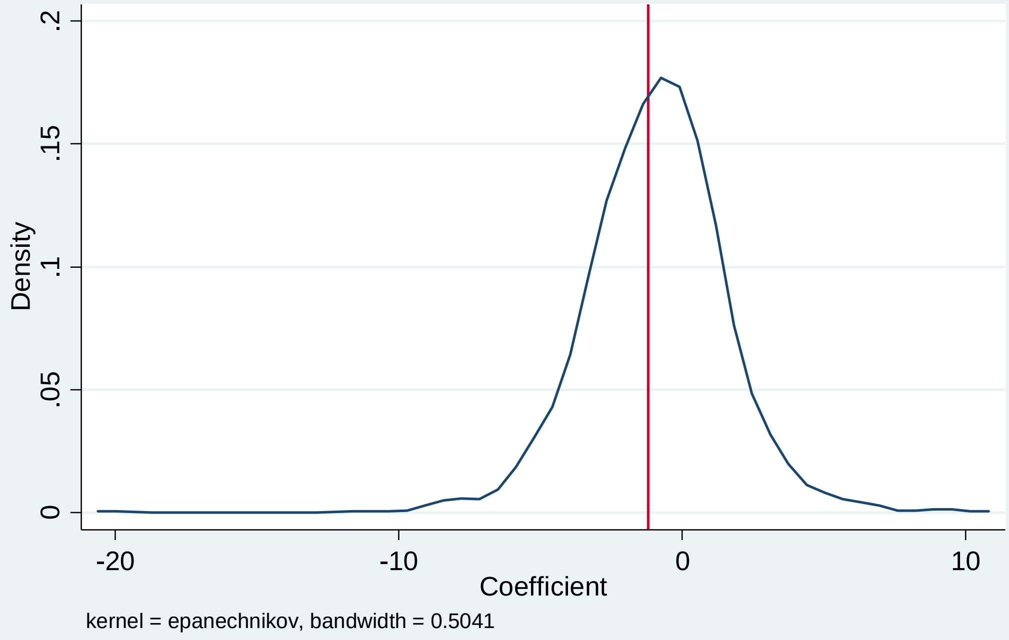


Figure 4c Coefficient Distribution of Rent-to-Price Ratio
from 1000 boostrap replications of the turnover regression

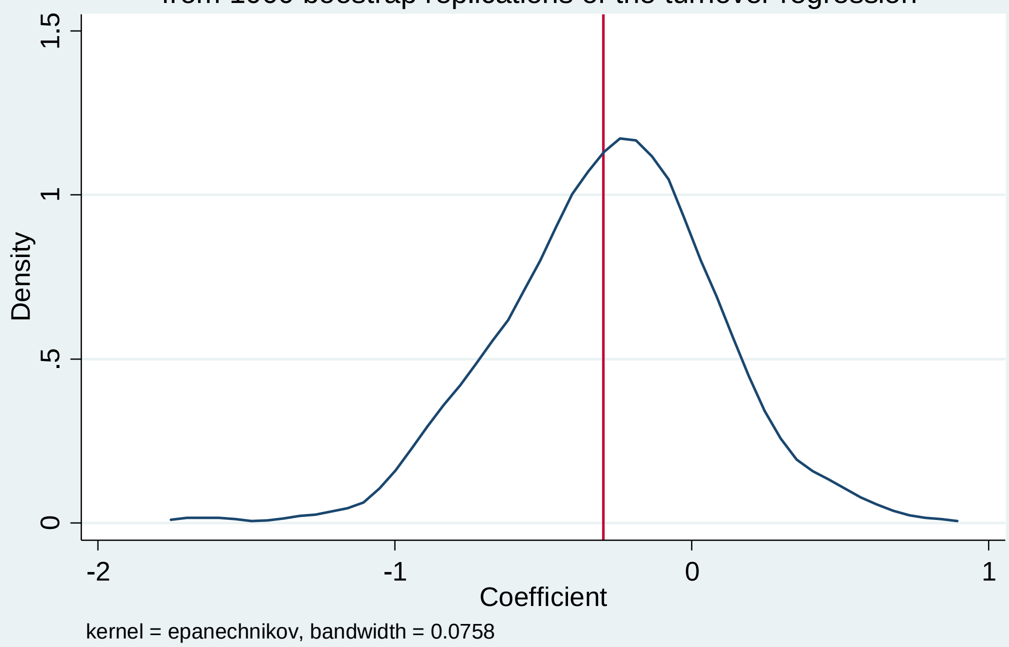
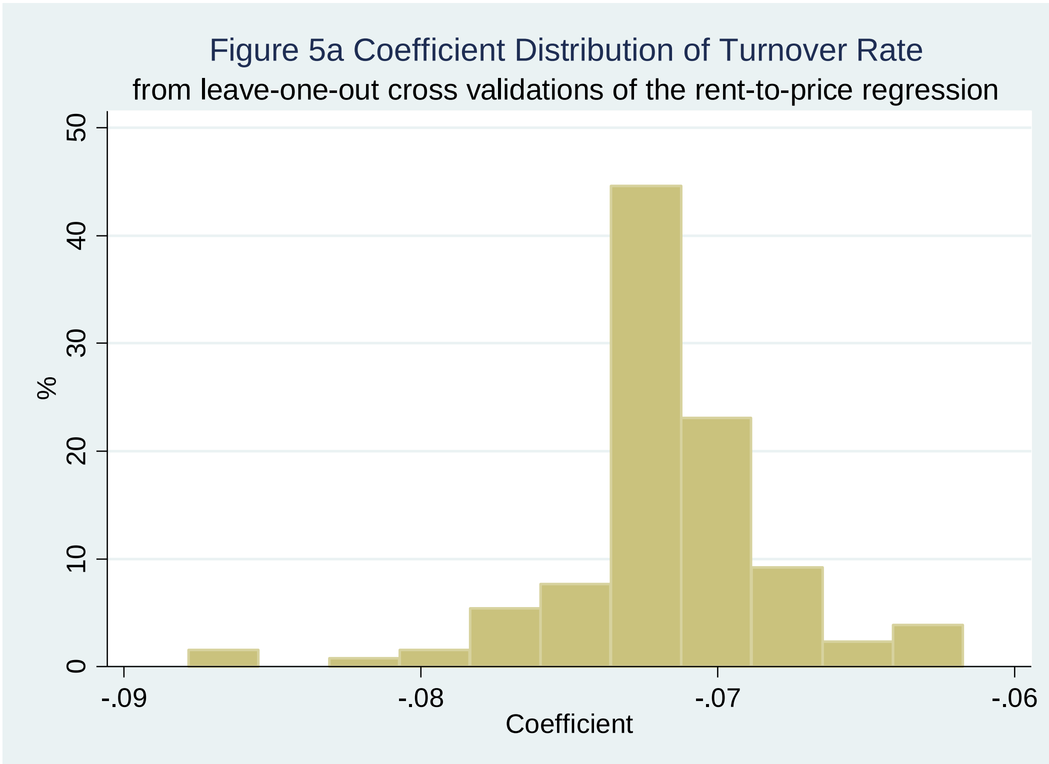
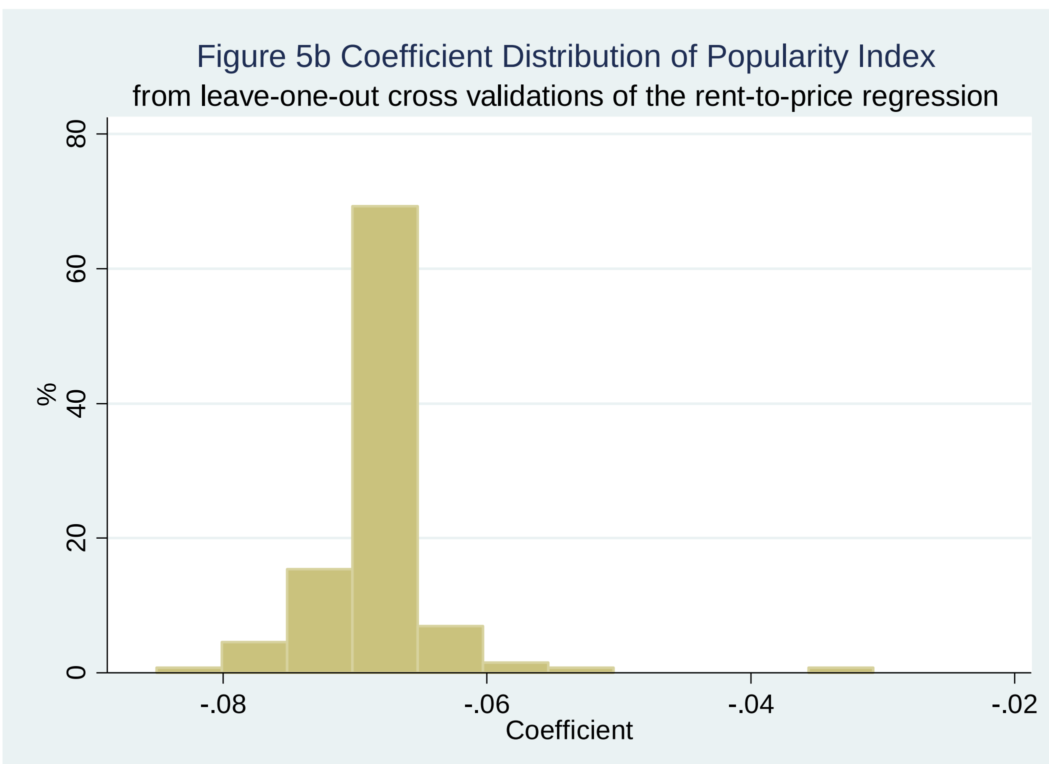


Figure 4d Coefficient Distribution of Popularity Index Ratio
from 1000 boostrap replications of the turnover regression

Figure 5 Leave-one-out Cross Validations





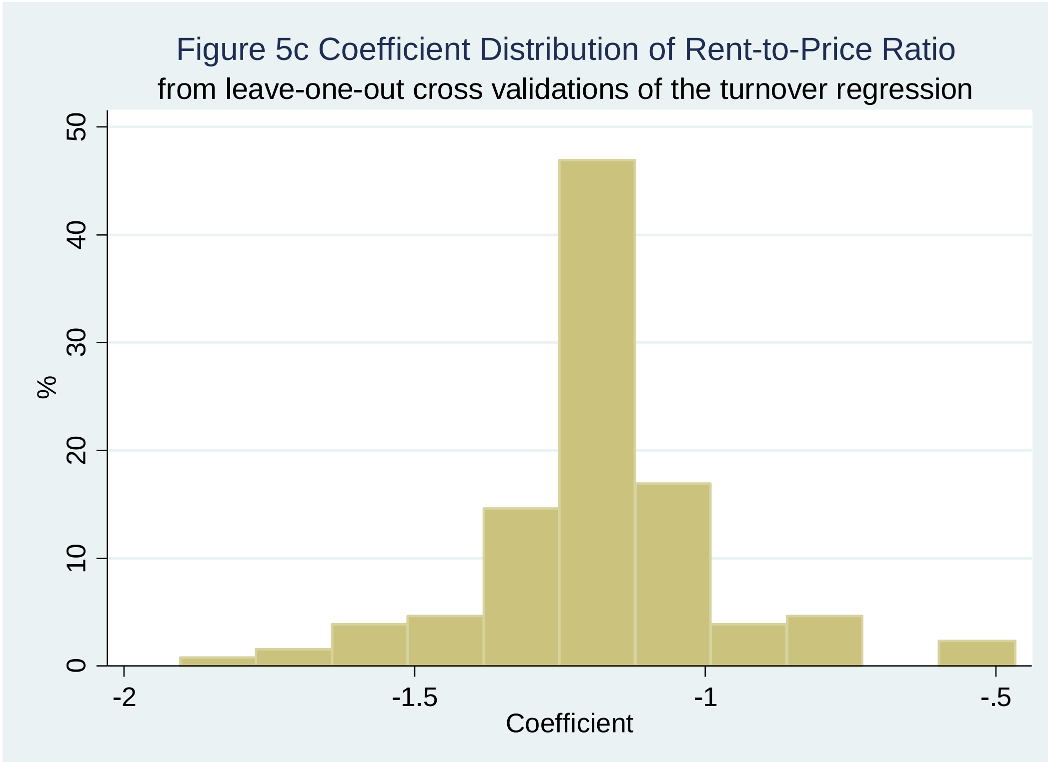


Figure 5c Coefficient Distribution of Rent-to-Price Ratio
from leave-one-out cross validations of the turnover regression

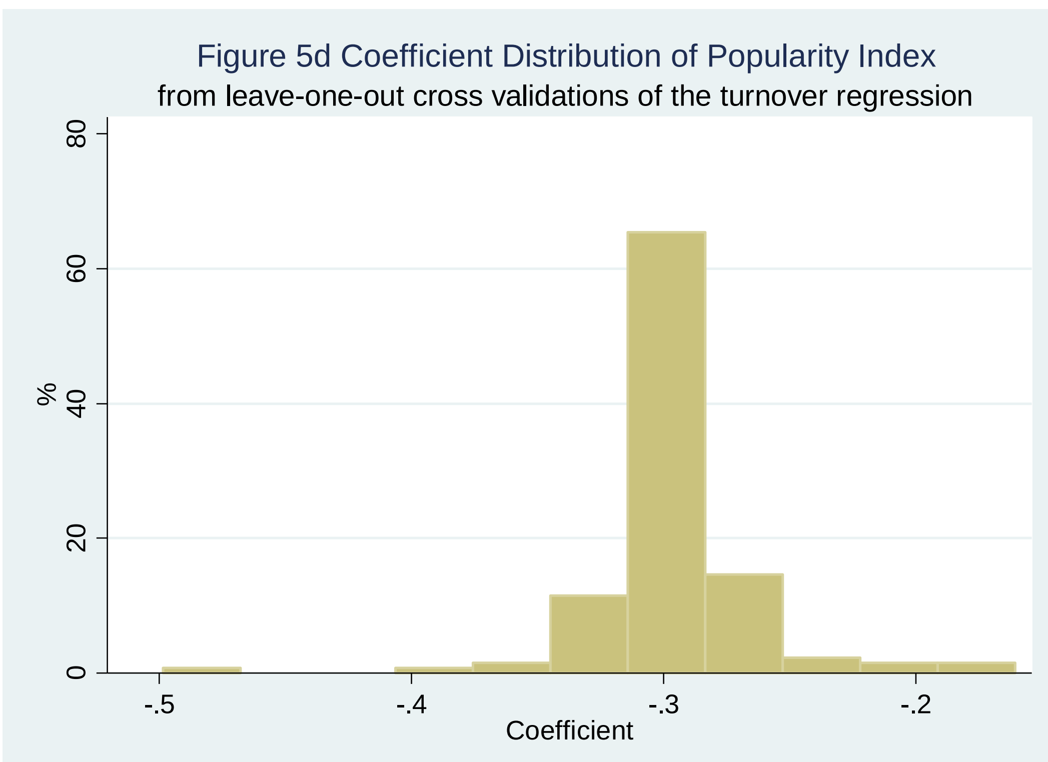
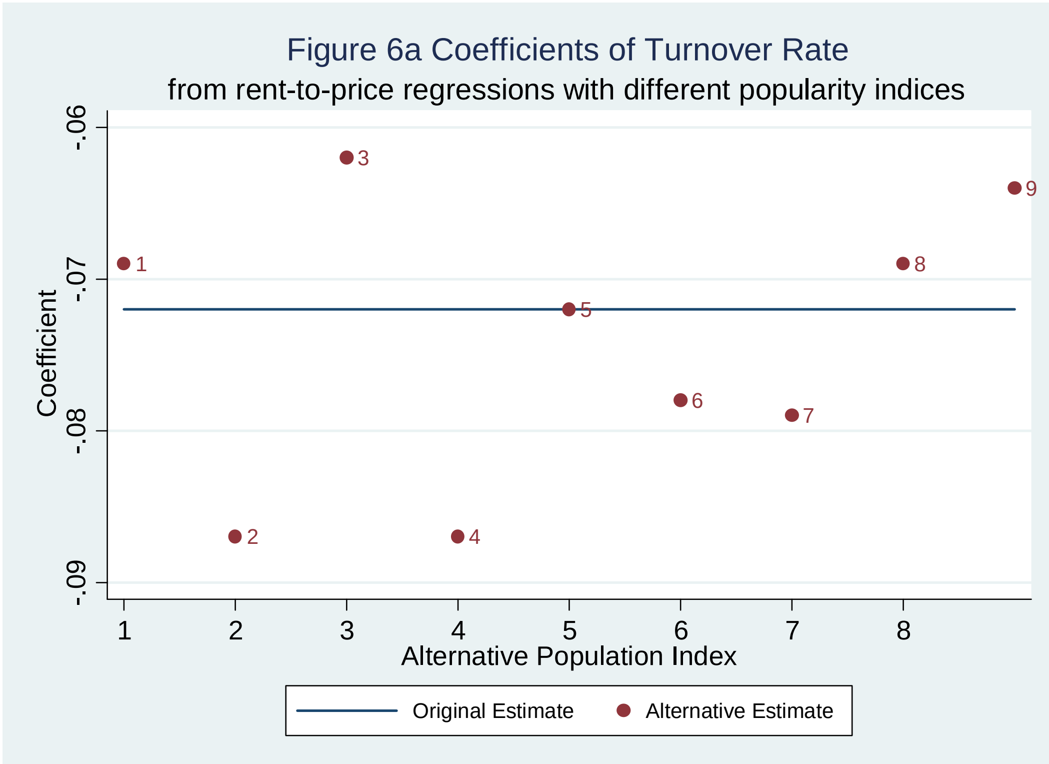


Figure 5d Coefficient Distribution of Popularity Index
from leave-one-out cross validations of the turnover regression

Figure 6 Different Popularity Indices



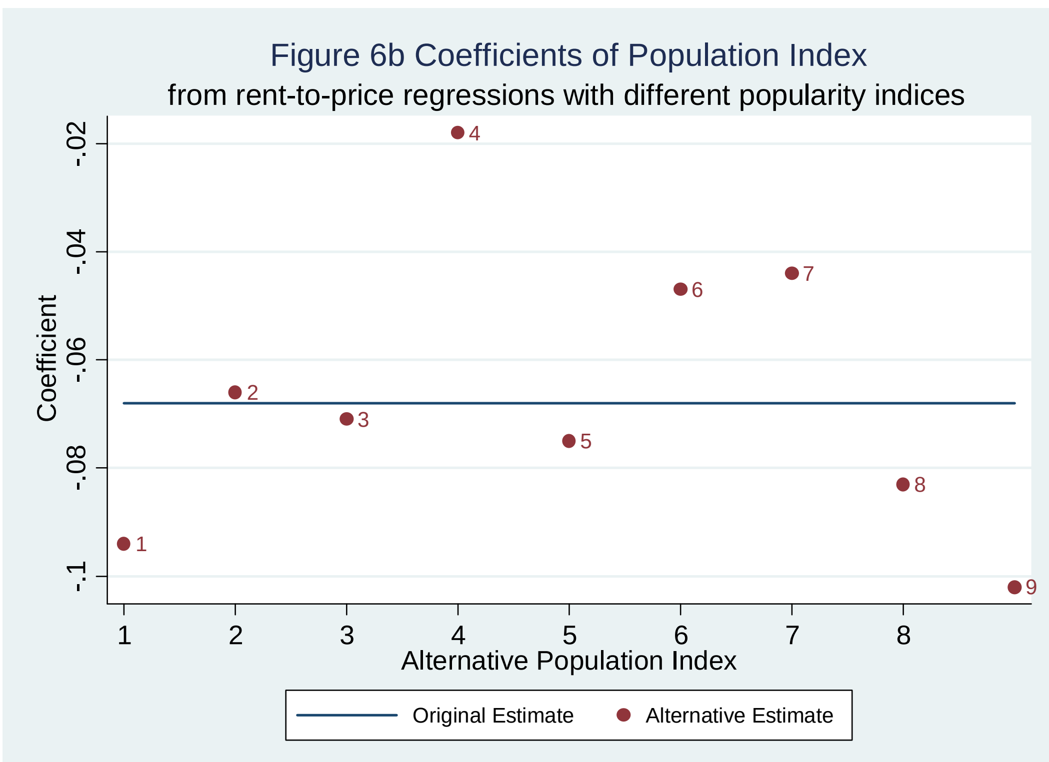


Figure 6c Coefficients of Rent-to-Price Ratio
from turnover regressions with different popularity indices

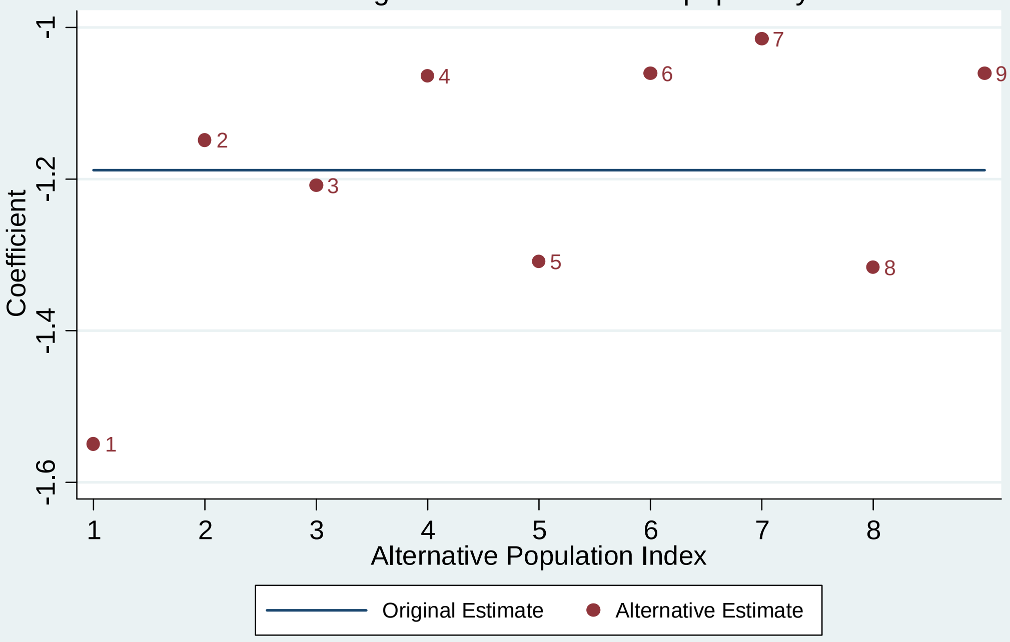
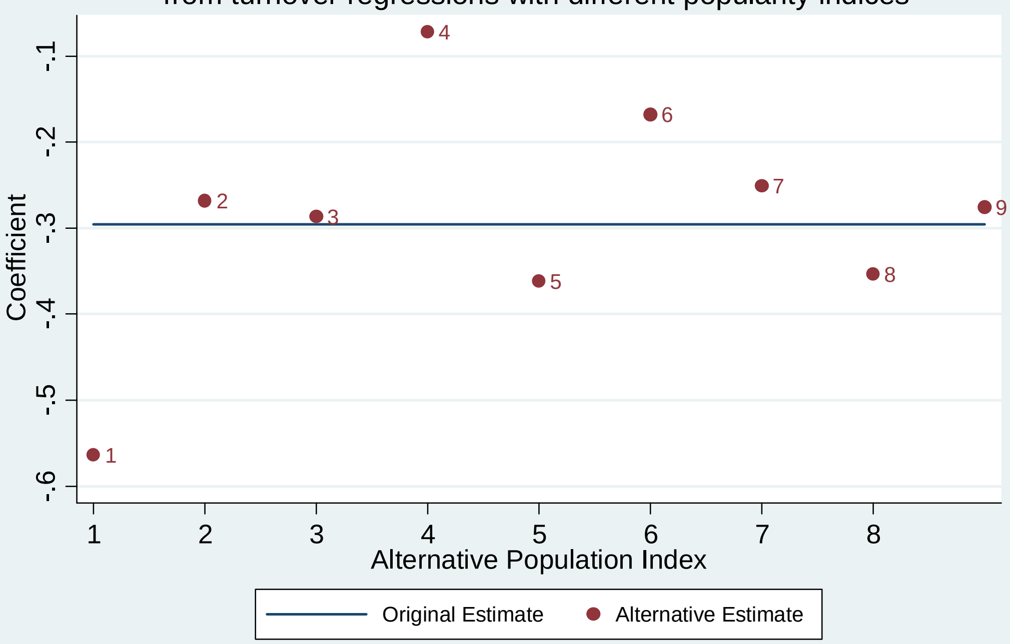


Figure 6d Coefficients of Population Index
from turnover regressions with different popularity indices



Description

Symbol	Indicator Removed
1	Swimming Pool
2	Club House
3	Average Net Gross Ratio
4	Distance to Metro
5	Developer Brand
6	Distance to CBD
7	School Zone
8	Affordability
9	Distance to Waterfront

Appendix

This appendix provides some supplementary materials to the main text.

Appendix A provides the proof of proposition 1.

Appendix B provides the results of our bootstrap replications for each of our explanatory variable and show that our results are indeed robust.

Appendix C provides the results of our Leave-one-out Cross Validations (LOOCV) for each of our explanatory variable and show that our results are indeed robust.

Appendix D compares the point estimate for each potential empirical determinant in (28) with alternatively constructed popularity index and show that our results are indeed robust.

Appendix E provides correlation tables among different variables and discusses the construction of the popularity index.

Appendix F contains the 3SLS estimation with additional district level data.

Appendix G provides a robustness check for altering the variable choices in X and Z.

Appendix H provides a discussion of the possible interactions between the labor and housing markets.

A Proof of Proposition 1

Before we formally prove the proposition, we need to first solve the steady state distribution of agents, sale price, rental (flow), value functions in both for-sale and rental market in terms of meeting rates of the buyers and home-seekers. It can be achieved by algebraic manipulation of the equations. Let us recall the symbols used in the theoretical model.

Table A1a:	List of Exogenous Variables
υ	flow utility of staying in a matched house
δ	match break-up rate
ρ	discount rate
α	fraction of households who demand for-sale housing
H	housing stock
$M_S\,(\text{Buyers}, \text{Sellers})$	matching function in for-sale market
$M_R\,(\text{Buyers}, \text{Sellers})$	matching function in rental market
q	matching probability

Table A1b: List of Endogenous Variables

	Meeting rates and market tightness
μ_S	buyer's meeting rate in for-sale market
η_S	seller's meeting rate in for-sale market
θ_S	market tightness = Buyers/Sellers in for-sale market
μ_R	house-seeker's meeting rate in rental market
η_R	landlord's meeting rate in rental market
θ_R	market tightness = Buyers/Sellers in rental market
	Price and rent
p	sale price
r	flow rental payment
	Asset values
W	matched owner-occupier
$\tilde{W}$	mismatched owner-occupier
U	matched household in rental housing
$\tilde{U}$	mismatched household in rental housing
R	unit rented to matched renter
$\tilde{R}$	unit rented to mismatched renter
V	vacant unit
V_S	vacant unit for sale
V_R	vacant unit for rent
	Measures
N_S	matched owner-occupiers
$\tilde{N}_S$	mismatched owner-occupiers
N_R	matched renter
$\tilde{N}_R$	mismatched renter
N_{V_S}	vacant unit for sale
N_{V_R}	vacant unit for rent

We can now proceed to prove the proposition. We first solve for the sale price, rental rate, and the value functions in both for-sale and rental markets, by including other equations, (3), (4), (5), (6), (7), (8), (9), (10), (11), (12). We first state the following equations, and then we will show to to derive them:

(for-sale market)

$$p = \frac{\rho + q\eta_S}{(2\rho + 2\delta + q\mu_S)\,\rho}\upsilon, \tag{32}$$

$$W=\frac{2\rho+q\mu_S}{(2\rho+2\delta+q\mu_S)\,\rho}\upsilon, \tag{33}$$

$$\widetilde{W}=\frac{q\mu_S}{(2\rho+2\delta+q\mu_S)\,\rho}\upsilon, \tag{34}$$

$$V_S=\frac{q\eta_S}{(2\rho+2\delta+q\mu_S)\,\rho}\upsilon, \tag{35}$$

(rental market)

$$r=\frac{(\rho+\delta+q\mu_R)\,(\rho+q\eta_R)+\delta q\mu_R}{(\rho+\delta+q\mu_R)^2}\upsilon, \tag{36}$$

$$R=\frac{\rho+q\eta_R}{\rho\,(\rho+\delta+q\mu_R)}\upsilon, \tag{37}$$

$$\widetilde{R}=\frac{(\rho+\delta+q\eta_R)\,\rho+(\delta+q\mu_R)\,q\eta_R}{\rho\,(\rho+\delta+q\mu_R)^2}\upsilon, \tag{38}$$

$$U=\frac{(q\mu_R-q\eta_R+\rho)\,q\mu_R-(\delta+\rho)\,q\eta_R}{\rho\,(\rho+\delta+q\mu_R)^2}\upsilon, \tag{39}$$

$$\widetilde{U}=\frac{(q\mu_R-q\eta_R)\,q\mu_R-(\delta+\rho)\,(\rho+q\eta_R)}{\rho\,(\delta+\rho+q\mu_R)^2}\upsilon, \tag{40}$$

$$V_R=\frac{q\eta_R}{\rho\,(\rho+\delta+q\mu_R)}\upsilon. \tag{41}$$

Notice that all these variables depend on the meeting rates, $\eta_S, \eta_R, \mu_S, \mu_R$. Recall by definition, $\mu_S(\theta_S)=\frac{q\eta_S(\theta_S)}{\theta_S}$, $\mu_R(\theta_R)=\frac{q\eta_R(\theta_R)}{\theta_R}$, and $\theta_S=\widetilde{N}_S/N_{V_S}$, $\theta_R=\widetilde{N}_R/N_{V_R}$. Thus, if we can solve for the 4 variables, $\widetilde{N}_S$, $\widetilde{N}_R$, N_{V_S}, and N_{V_R}, we can solve for the meeting rates $\eta_S, \eta_R, \mu_S, \mu_R$, and hence all other variables.

The derivations of equations from (32) to (41) are not difficult but would take some steps. To begin, combining (3) and (4), we get

$$W-\widetilde{W}=\frac{\upsilon+q\mu_S\,(p-V)}{\rho+\delta+q\mu_S}. \tag{42}$$

Similarly, combining (5), (6), we get

$$U-\widetilde{U}=\frac{\upsilon}{\rho+\delta+q\mu_R}. \tag{43}$$

From (7), we get

$$V_S=\left(\frac{q\eta_S}{\rho+q\eta_S}\right)p. \tag{44}$$

From (8), we get

$$V_R=\left(\frac{q\eta_R}{\rho+q\eta_R}\right)R. \tag{45}$$

Combining (9), (10), we get

$$(\rho+\delta)\left(R-\widetilde{R}\right)=q\mu_R\left(\widetilde{R}-V_R\right). \tag{46}$$

Given all these expressions, we can solve for the for-sale market equations, i.e. from (32) to (35).

From (11) and the fact that $V=V_S$ at the equilibrium, we have

$$p-V_S=(1/2)\left(W-\widetilde{W}\right). \tag{47}$$

We then substitute (47) into (42), we have

$$\left(W-\widetilde{W}\right)=\frac{2v}{2(\rho+\delta)+q\mu_S}. \tag{48}$$

In turn, we substitute (48) into (3), we get (33).
Combining (48) with (33), we get (34).
Combining (7) and (47), (48), we get (35).
Now substitute (35) and (48) into (47), we get (32).

To solve for the rental market-related equations, i.e. from (36) to (41), we combine (12) with (43), (45), and we get (37).
Combining (45) with (37), we get (41).

We then combine (37), (45) and (46), we get (38).
We then substitute (37), (38) into (9), we get (36).
We substitute (36) and (43) into (5), we get (39).
Similarly, we substitute (36) and (43) into (6), we get (40).

Since a house-owner can arbitrage between being a seller or being a landlords, we can equate (35) and (41), hence obtain condition (A), which is re-stated here for future reference:

$$V_S=V_R\Rightarrow\frac{\eta_S}{2(\rho+\delta)+q\mu_S}=\frac{\eta_R}{\rho+\delta+q\mu_R}. \tag{A}$$

Recall that the housing market equilibrium is characterized by the system of equations stated above, and as some of the equations are nonlinear in the variables in concern, it can be difficult to solve them for the equilibrium in general. Sargent (1979), among others, suggests that first identifying the "core" sub-system of the original system of equations is necessary. Thus, it is important to identify the "core" sub-system. In the current context, the "core" sub-system exists and is composed of four equations:

$$\widetilde{N}_S=\frac{\delta}{q\mu_S+\delta}\alpha, \tag{49}$$

$$\widetilde{N}_R=\frac{\delta}{q\mu_R+\delta}(1-\alpha), \tag{50}$$

$$N_{V_R} + N_{V_S} = H - 1, \tag{51}$$

$$V_S = V_R \Rightarrow \frac{\eta_S}{2(\rho+\delta) + q\mu_S} = \frac{\eta_R}{\rho + \delta + q\mu_R}. \tag{A}$$

These equations are "core" for a couple of reasons. Once we solve them, we can *recover other variables* easily. For instance, given $\widetilde{N}_S$, we can get N_S from (14), and given $\widetilde{N}_R$, we can get N_R from (15), etc. More importantly, equations (49), (50), (51), and (A) are 4 equations in 4 unknowns, $\widetilde{N}_S$, $\widetilde{N}_R$, N_{V_S}, and N_{V_R}. Observe that $\theta_S = \widetilde{N}_S / N_{V_S}$ and $\theta_R = \widetilde{N}_R / N_{V_R}$, thus, $\eta_S, \eta_R, \mu_S, \mu_R$ are all functions of $\widetilde{N}_S$, $\widetilde{N}_R$, N_{V_S}, and N_{V_R}.

Now we show how to derive (49), (50). Combine (14) with (17), we get (49). Similarly, we combine (15) with (18), we get (49).

The 4-equation sub-system can be further simplified. Combining (49), (50), with the observation that $\theta_S = \widetilde{N}_S / N_{V_S}$ and $\theta_R = \widetilde{N}_R / N_{V_R}$, with equation (51), we get (E) , which is re-stated here for convenience:

$$\frac{\delta}{q\eta_R + \delta\theta_R}(1-\alpha) + \frac{\delta}{q\eta_S + \delta\theta_S}\alpha = H - 1. \tag{E}$$

which defines a negatively-sloped (implicit) function for θ_R of θ_S. Hence, the "core" sub-system is reduced to 2 equations, (A) and (E) only. Notice that (E) contains additional information about the lower bounds of θ_S and θ_R. First, consider the case when $\theta_R \to \infty$. By (2), $\frac{\partial \eta_R(\theta_R)}{\partial \theta_R} > 0$. Hence, (E) implies that $\frac{\delta}{q\eta_S + \delta\theta_S}\alpha = H - 1$ as $\theta_R \to \infty$. Hence, the function is defined for $\theta_S > \theta_S^{\min}$, where

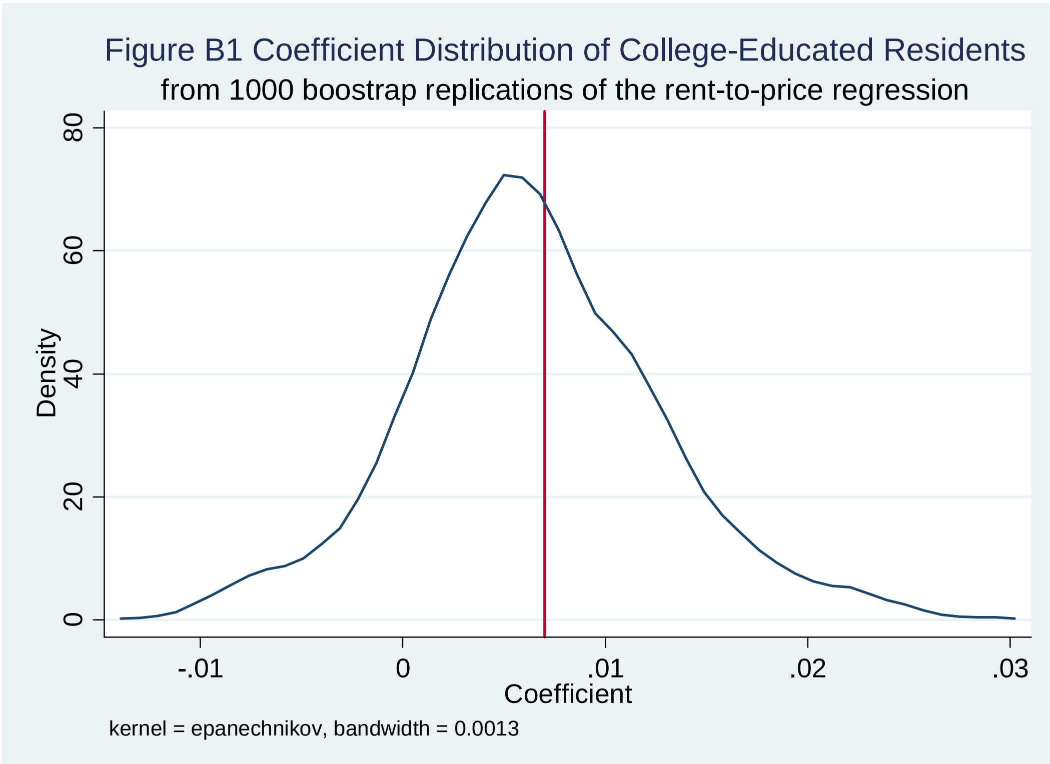
$$q\eta_S\left(\theta_S^{\min}\right) + \delta\theta_S^{\min} = \frac{\alpha\delta}{H-1},$$

at which $\theta_R \to \infty$. By the same token, consider the case when $\theta_S \to \infty$. By (1), $\frac{\partial \eta_S(\theta_S)}{\partial \theta_S} > 0$. Hence, $\theta_R \to \theta_R^{\min}$ where

$$q\eta_R\left(\theta_R^{\min}\right) + \delta\theta_R^{\min} = \frac{(1-\alpha)\delta}{H-1}.$$

That is, the function thus defined with "bounds" $\theta_S^{\min}$ on the y-axis and $\theta_R^{\min}$ on the x-axis. On the other hand, (A) defines a positively-sloped function for θ_R of θ_S that starts out at the origin and increases without bounds. A unique equilibrium exists. See Figure 1.

Appendix B: Bootstrap Replications



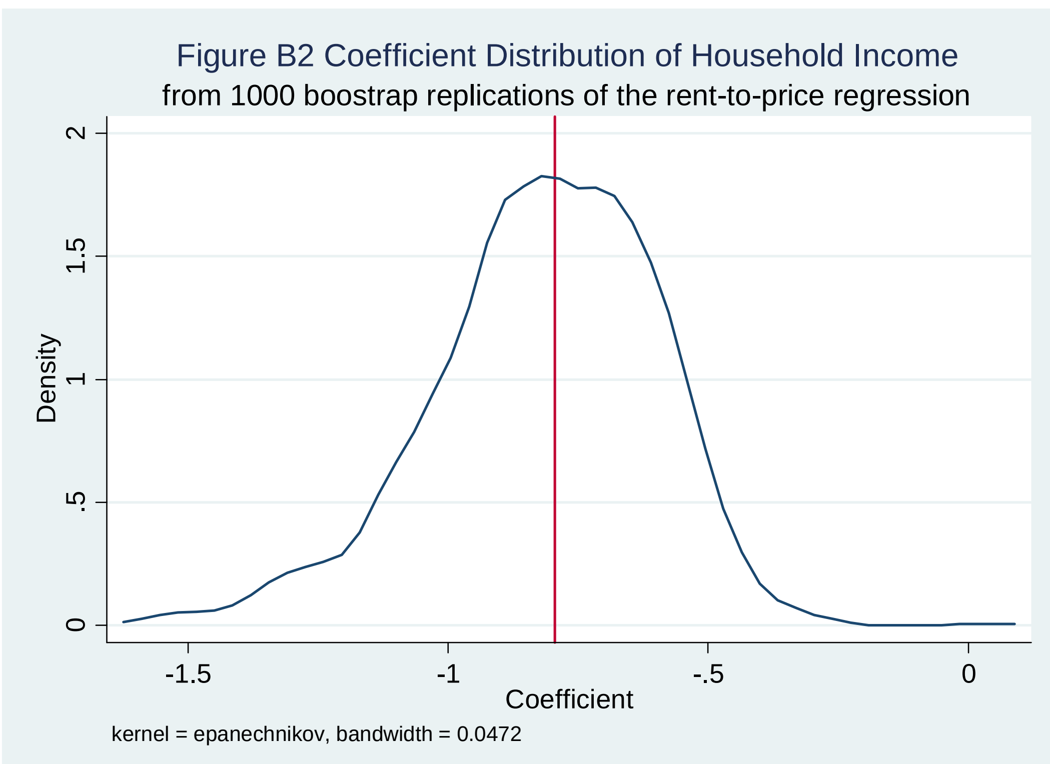
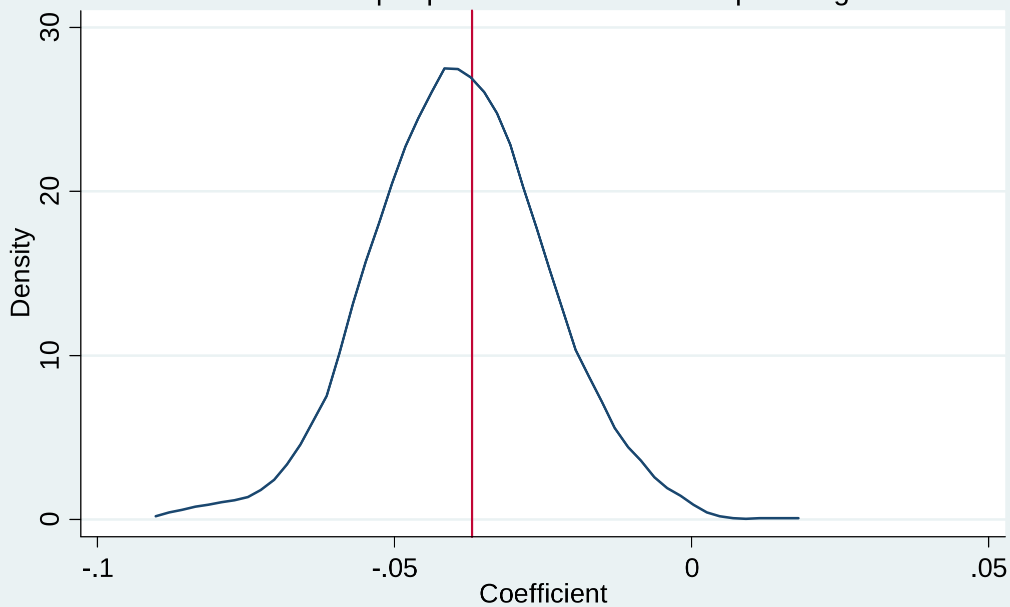


Figure B3 Coefficient Distribution of Mortgage

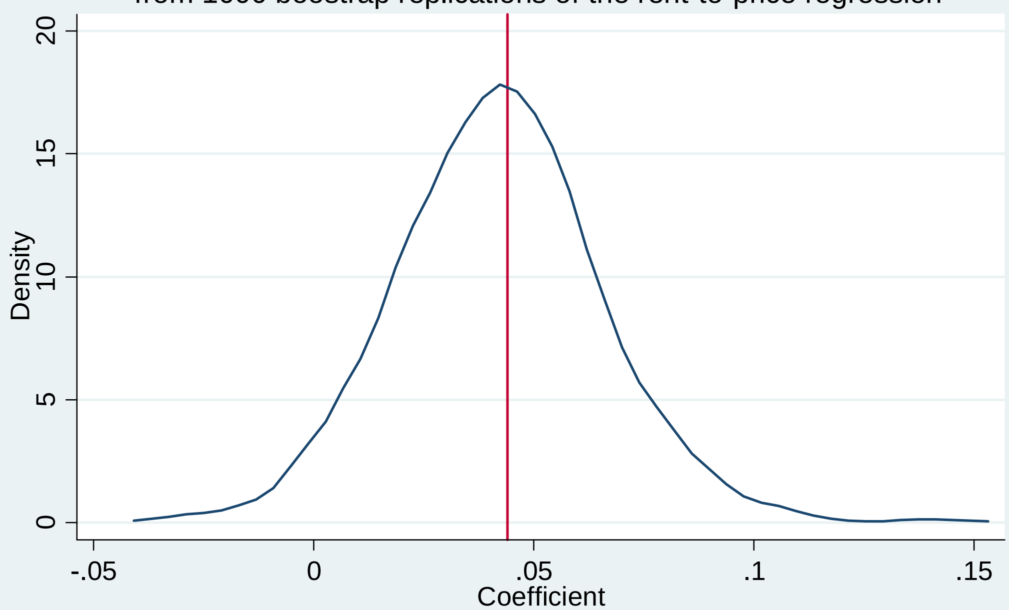
from 1000 boostrap replications of the rent-to-price regression



kernel = epanechnikov, bandwidth = 0.0032

Figure B4 Coefficient Distribution of Real Estate Work Experience

from 1000 boostrap replications of the rent-to-price regression



kernel = epanechnikov, bandwidth = 0.0051

Figure B5 Coefficient Distribution of House Price Growth Rate
from 1000 boostrap replications of the rent-to-price regression

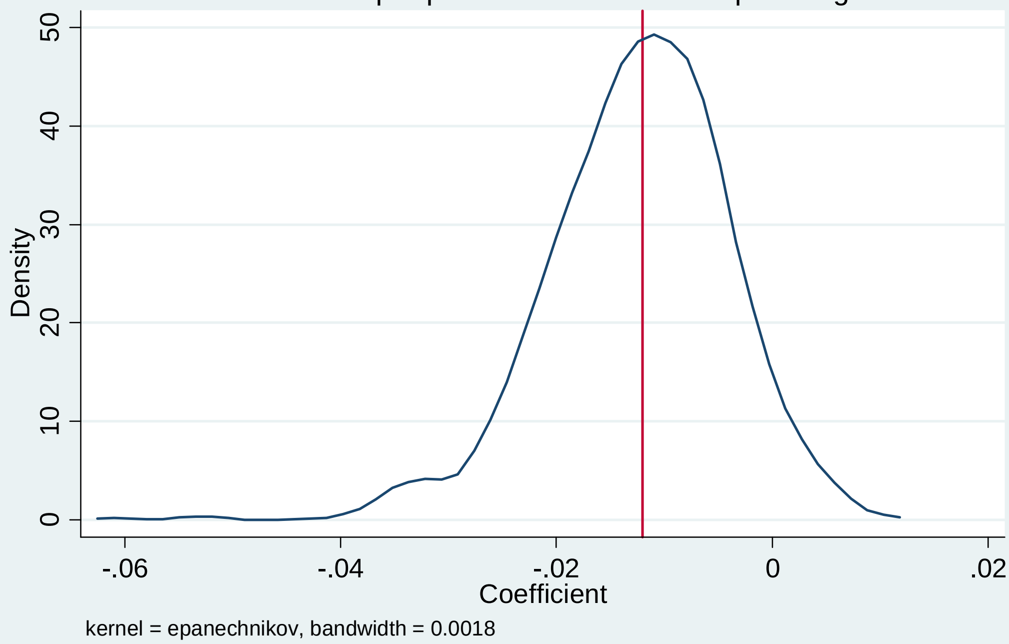
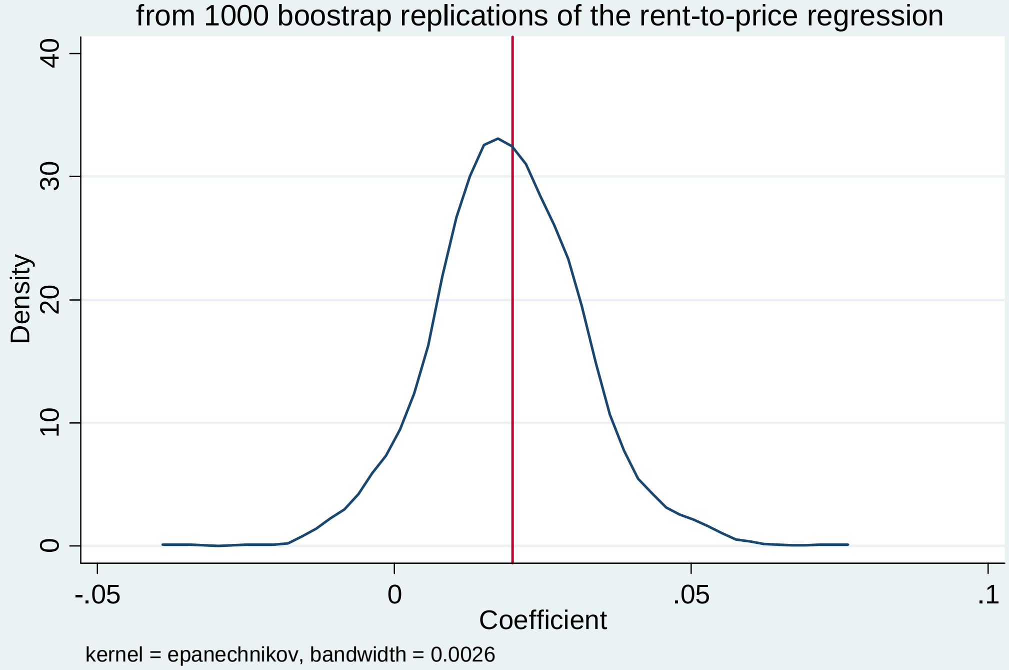


Figure B6 Coefficient Distribution of Rent Growth Rate
from 1000 boostrap replications of the rent-to-price regression



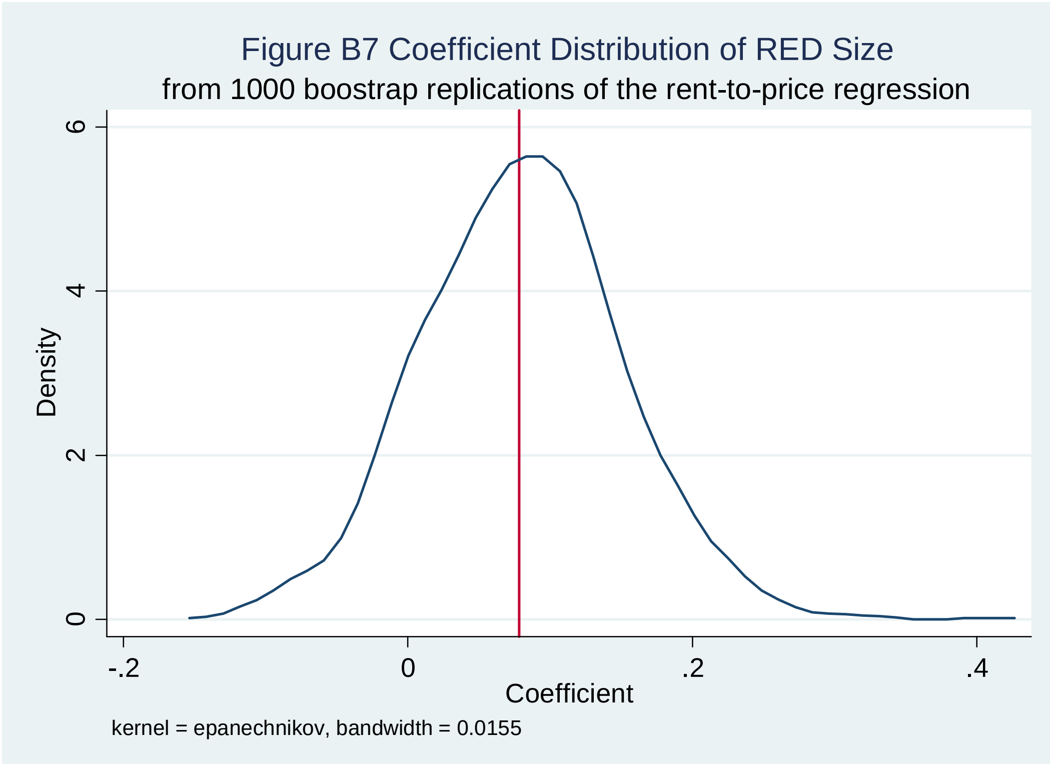


Figure B7 Coefficient Distribution of RED Size
from 1000 boostrap replications of the rent-to-price regression
Density
0
2
4
6
-.2
0
.2
.4
Coefficient
kernel = epanechnikov, bandwidth = 0.0155

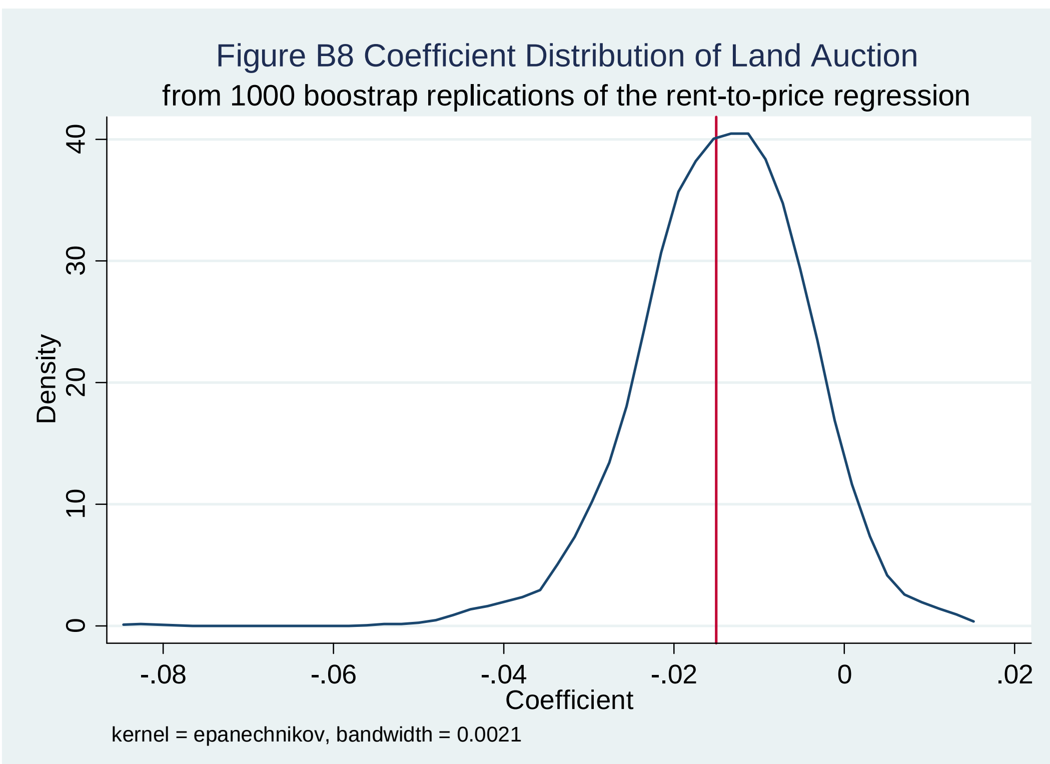


Figure B8 Coefficient Distribution of Land Auction
from 1000 boostrap replications of the rent-to-price regression
Density
0
10
20
30
40
-.08
-.06
-.04
-.02
0
.02
Coefficient
kernel = epanechnikov, bandwidth = 0.0021

Figure B9 Coefficient Distribution of Young Population
from 1000 boostrap replications of the turnover regression

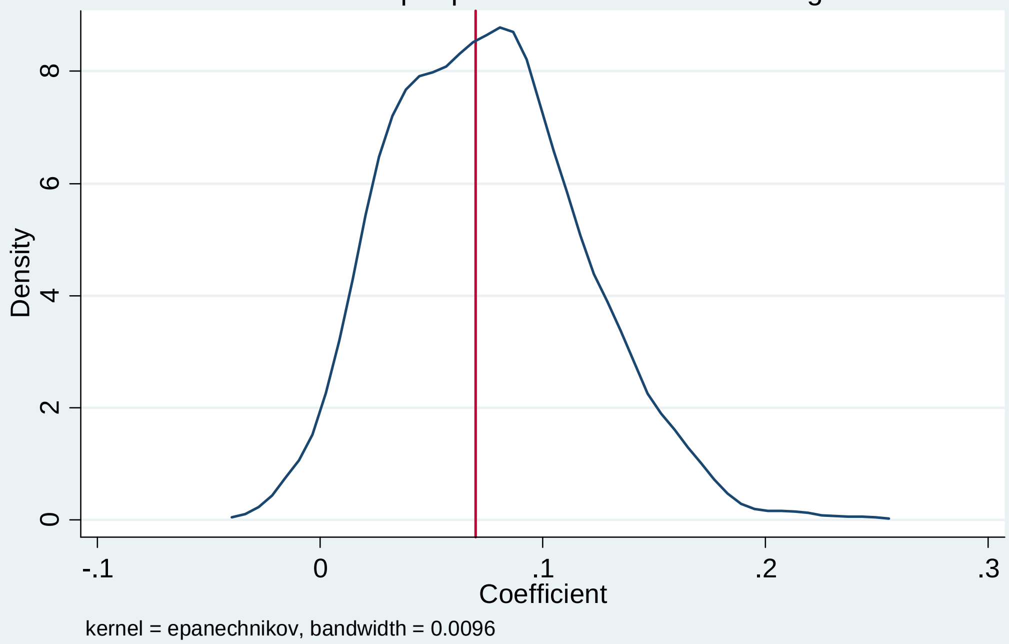


Figure B10 Coefficient Distribution of Household Income
from 1000 boostrap replications of the turnover regression

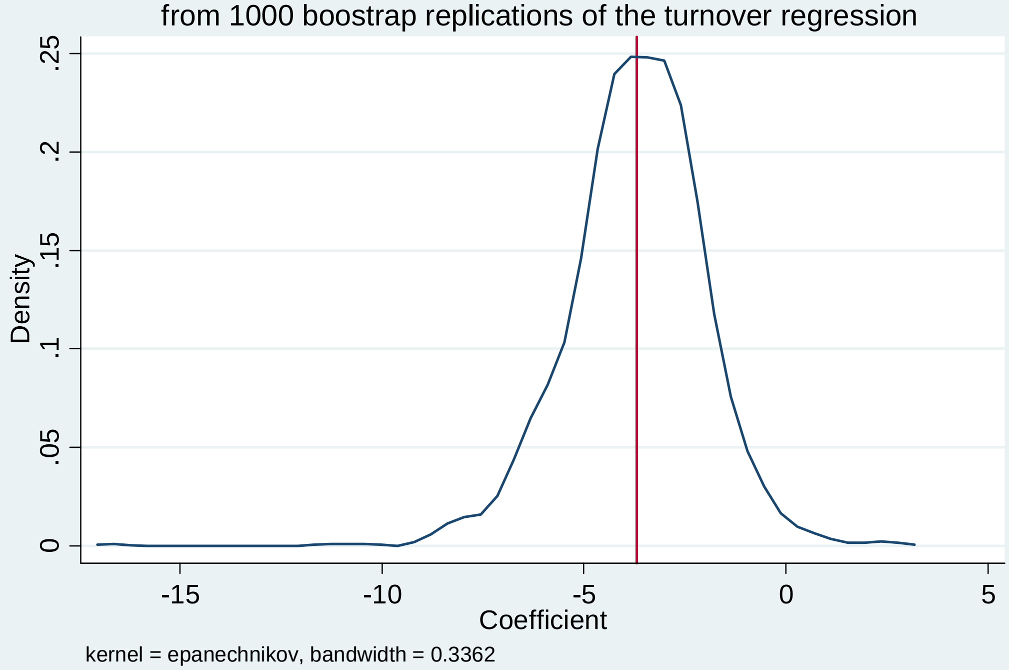
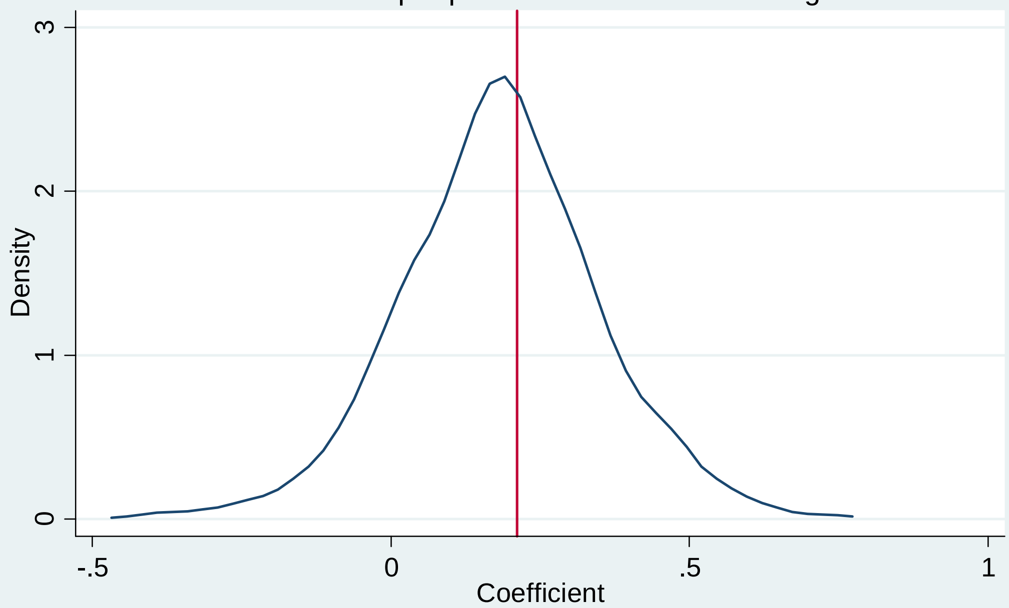
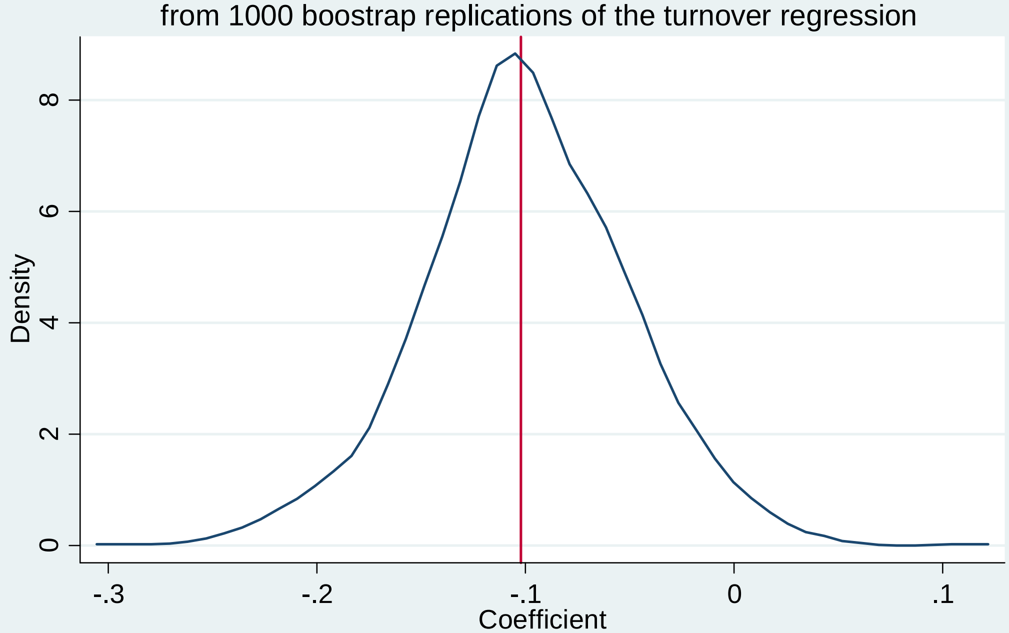


Figure B11 Coefficient Distribution of Real Estate Work Experience
from 1000 boostrap replications of the turnover regression



kernel = epanechnikov, bandwidth = 0.0347

Figure B12 Coefficient Distribution of House Price Growth Rate
from 1000 boostrap replications of the turnover regression



kernel = epanechnikov, bandwidth = 0.0107

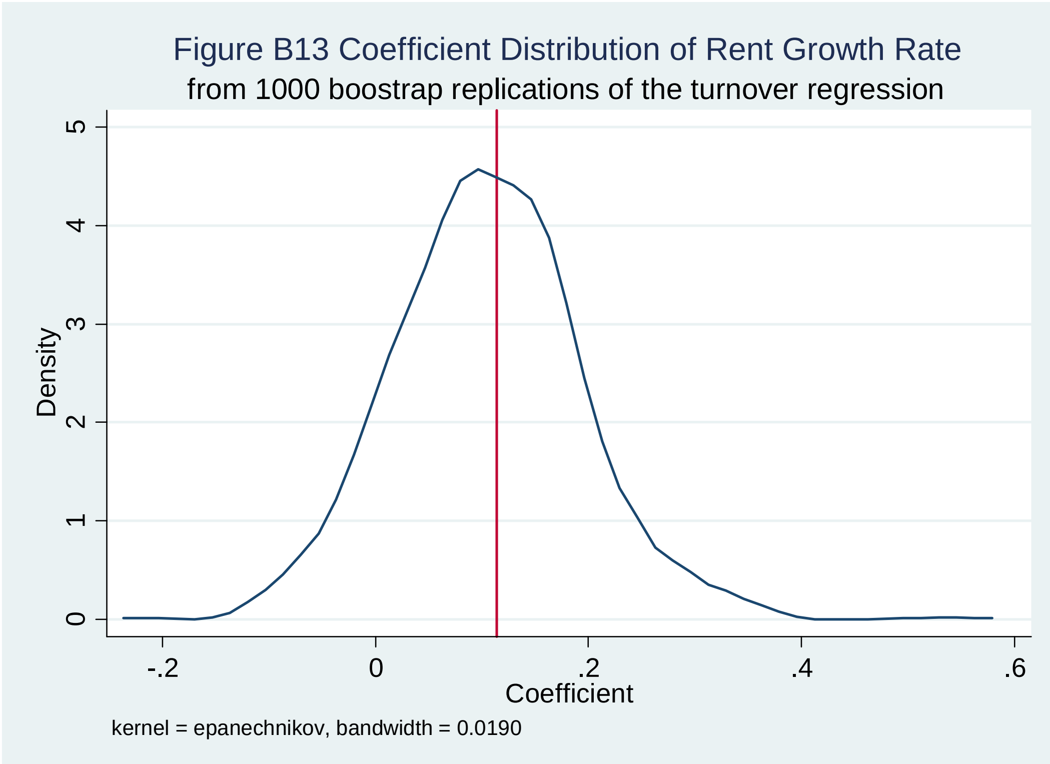


Figure B13 Coefficient Distribution of Rent Growth Rate
from 1000 boostrap replications of the turnover regression
Density
0
1
2
3
4
5
-.2
0
.2
.4
.6
Coefficient
kernel = epanechnikov, bandwidth = 0.0190

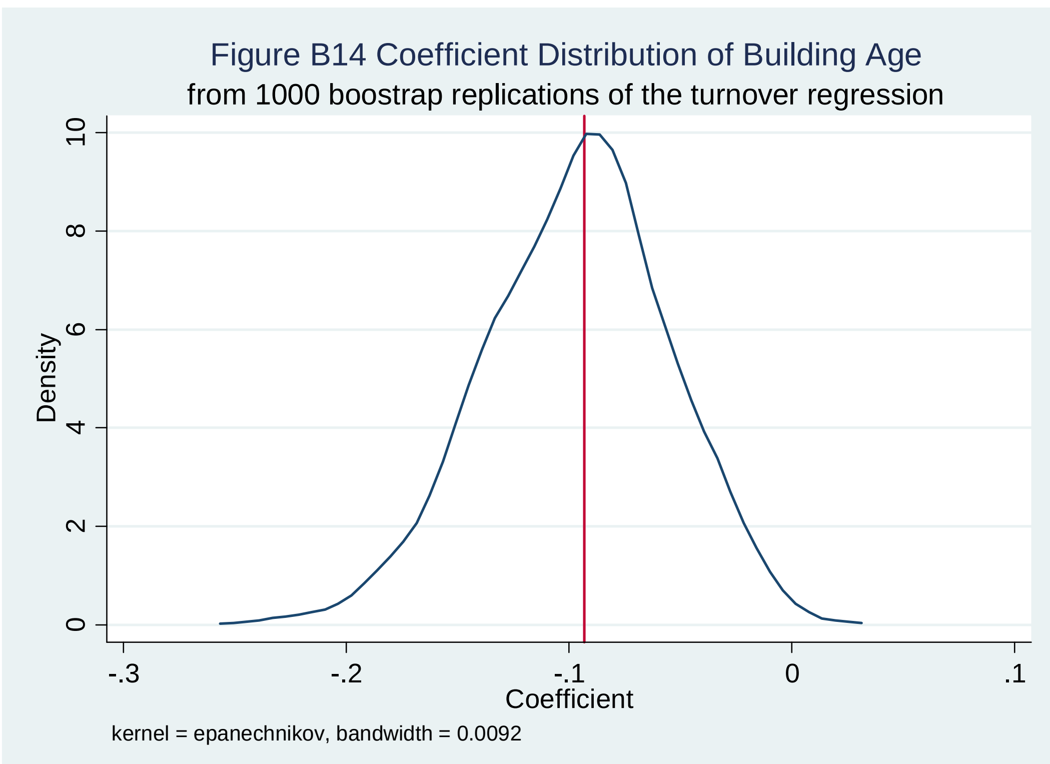


Figure B14 Coefficient Distribution of Building Age
from 1000 boostrap replications of the turnover regression
Density
0
2
4
6
8
10
-.3
-.2
-.1
0
.1
Coefficient
kernel = epanechnikov, bandwidth = 0.0092

Figure B15 Coefficient Distribution of RED Size
from 1000 boostrap replications of the turnover regression

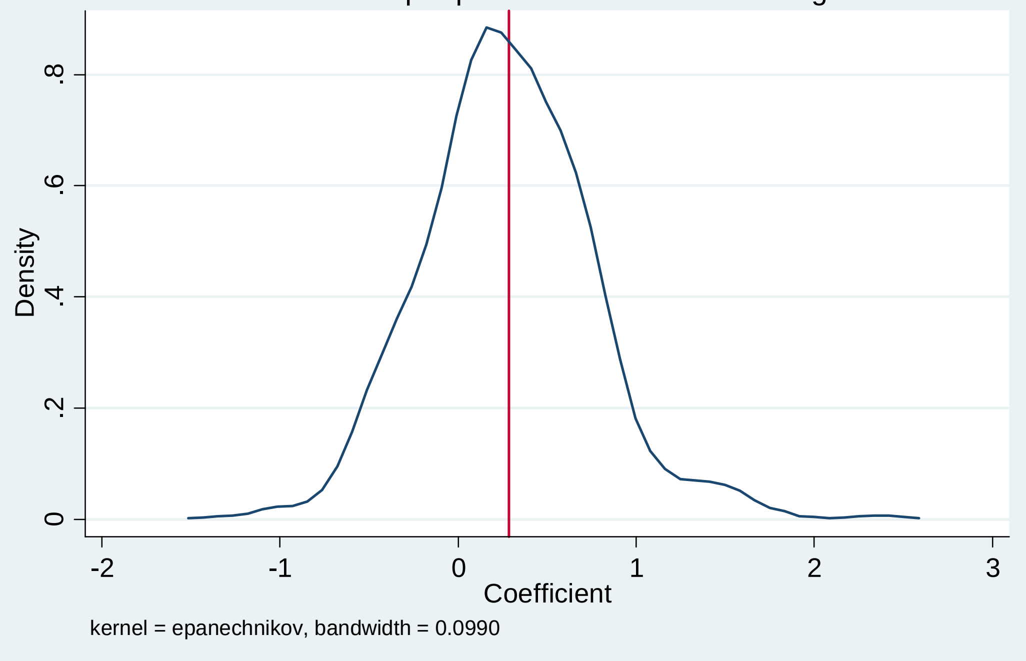
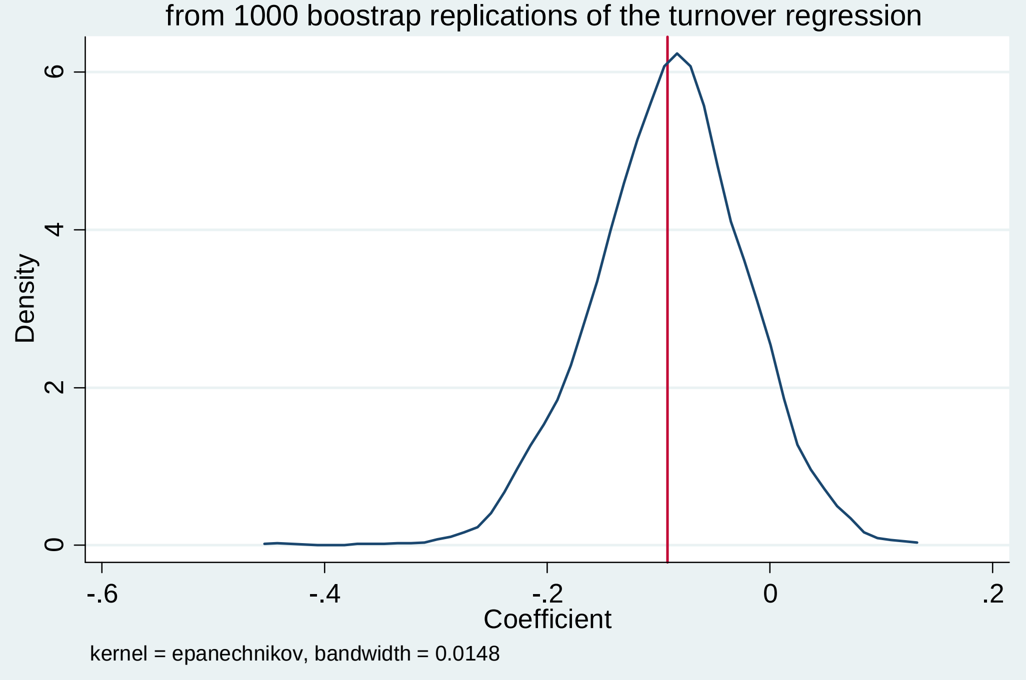
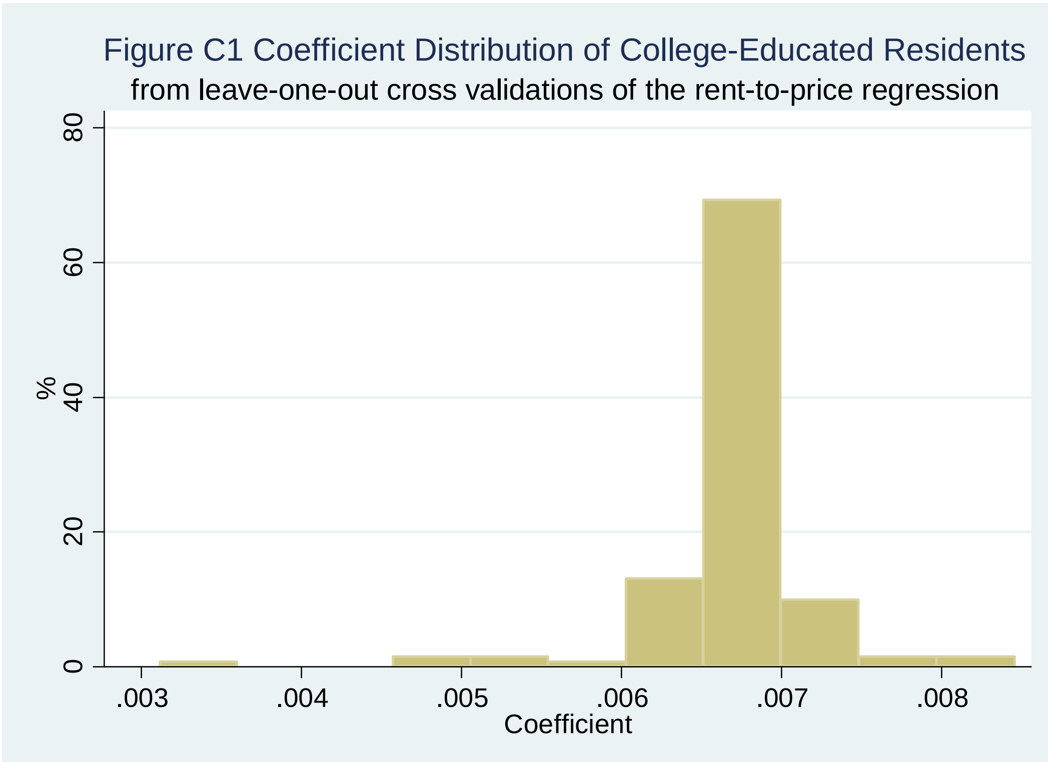
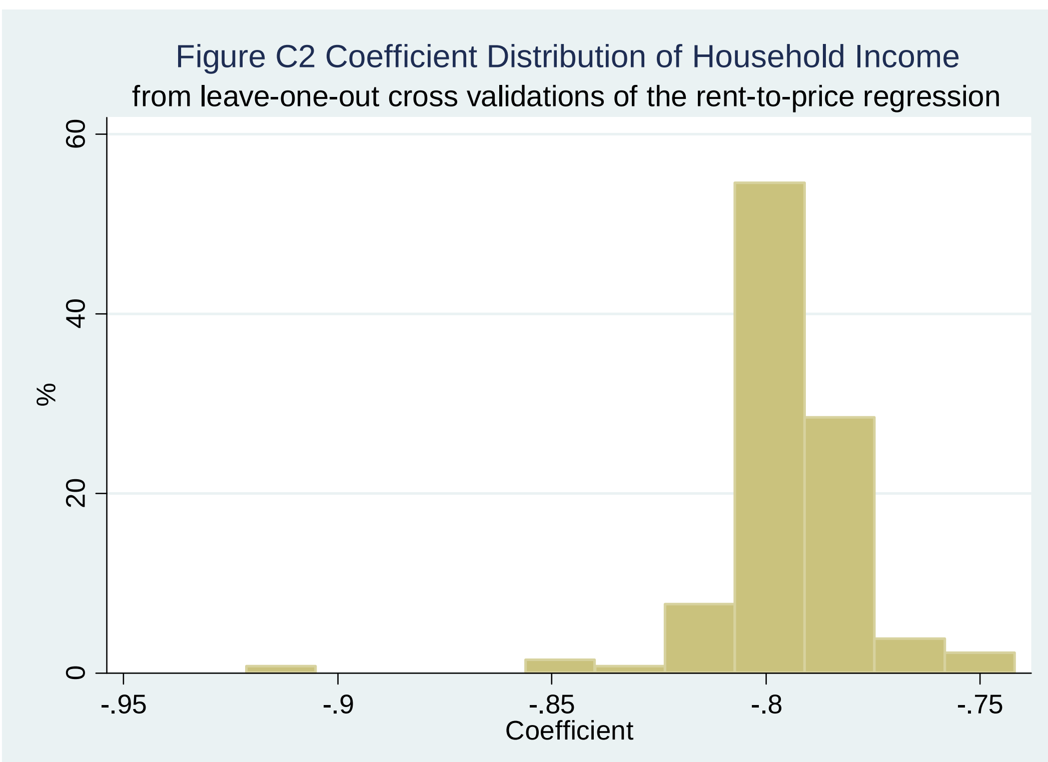


Figure B16 Coefficient Distribution of Land Auction
from 1000 boostrap replications of the turnover regression



Appendix C: Leave-one-out Cross Validations





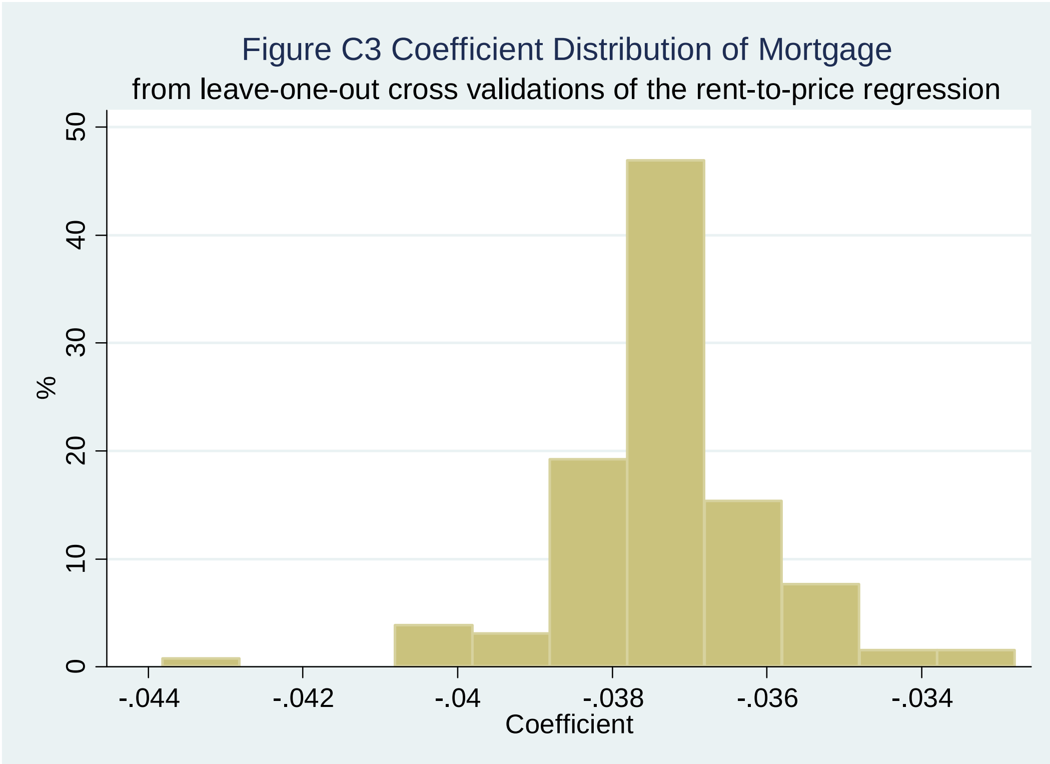


Figure C3 Coefficient Distribution of Mortgage
from leave-one-out cross validations of the rent-to-price regression

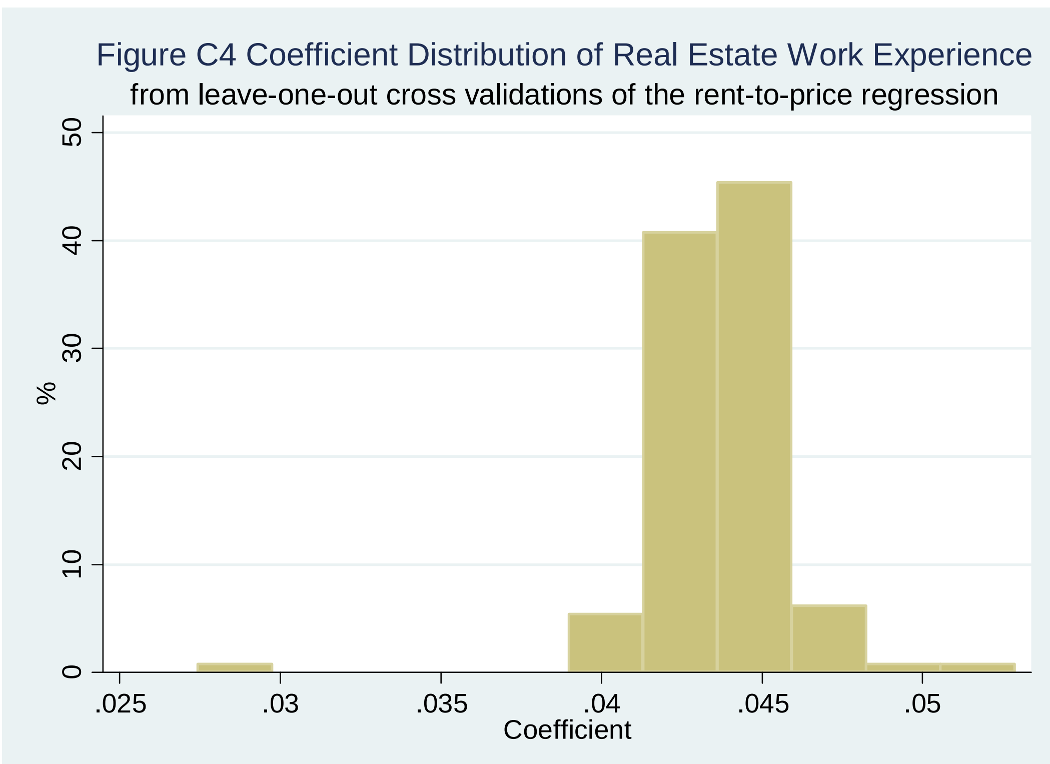


Figure C4 Coefficient Distribution of Real Estate Work Experience
from leave-one-out cross validations of the rent-to-price regression

Figure C5 Coefficient Distribution of House Price Growth Rate
from leave-one-out cross validations of the rent-to-price regression

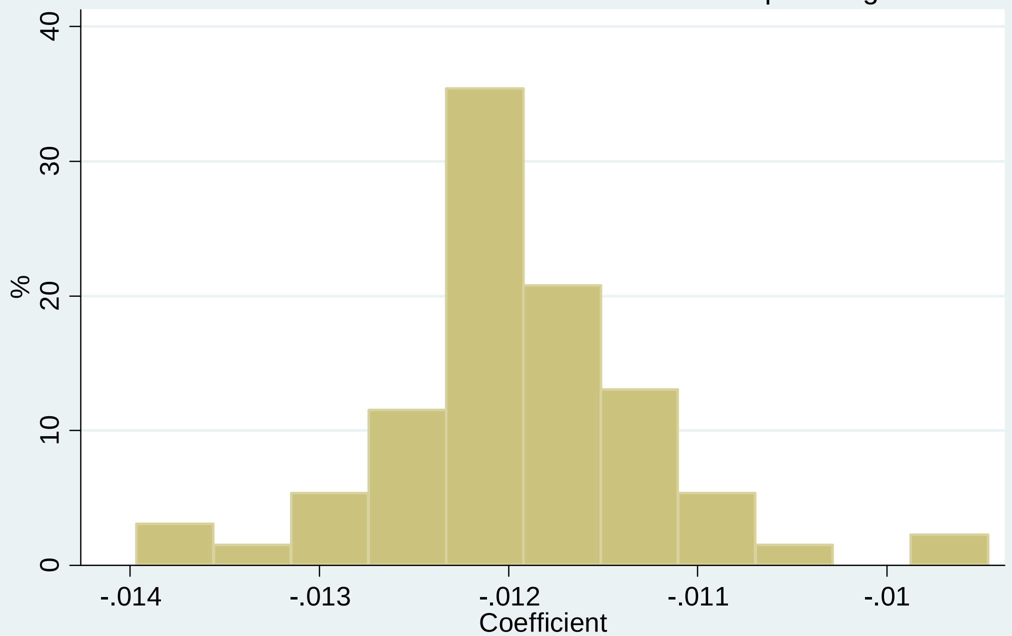
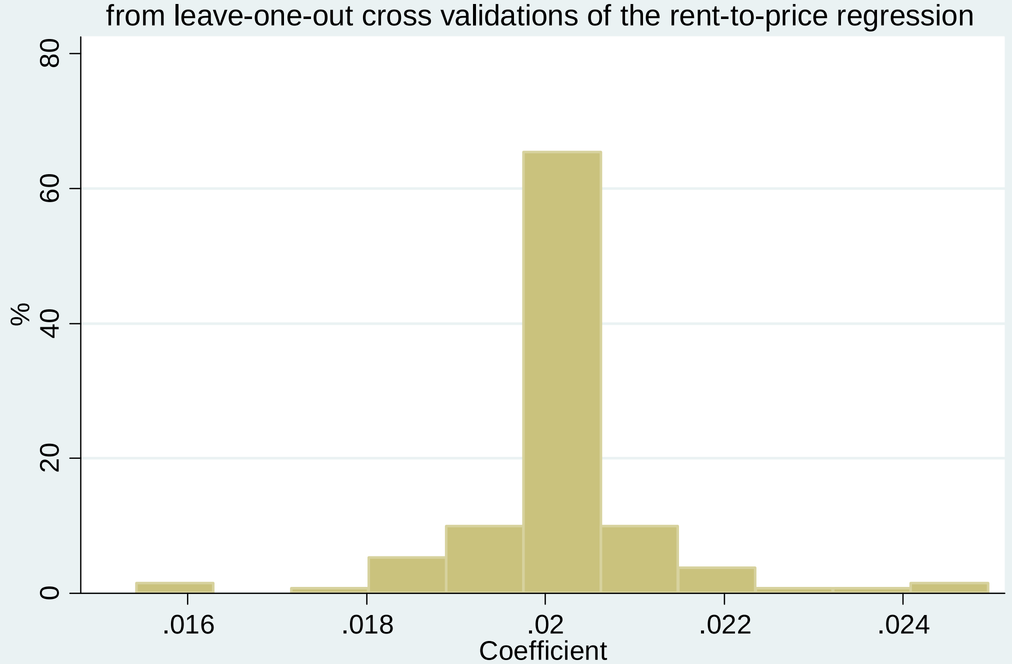


Figure C6 Coefficient Distribution of Rent Growth Rate
from leave-one-out cross validations of the rent-to-price regression



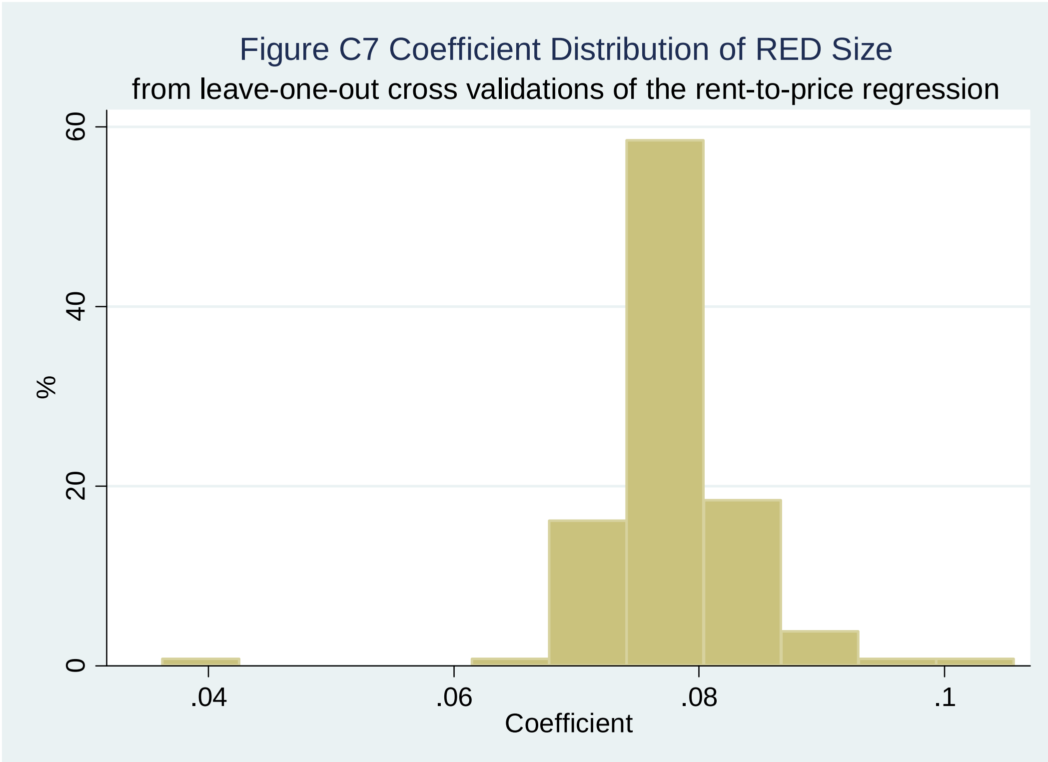


Figure C7 Coefficient Distribution of RED Size
from leave-one-out cross validations of the rent-to-price regression

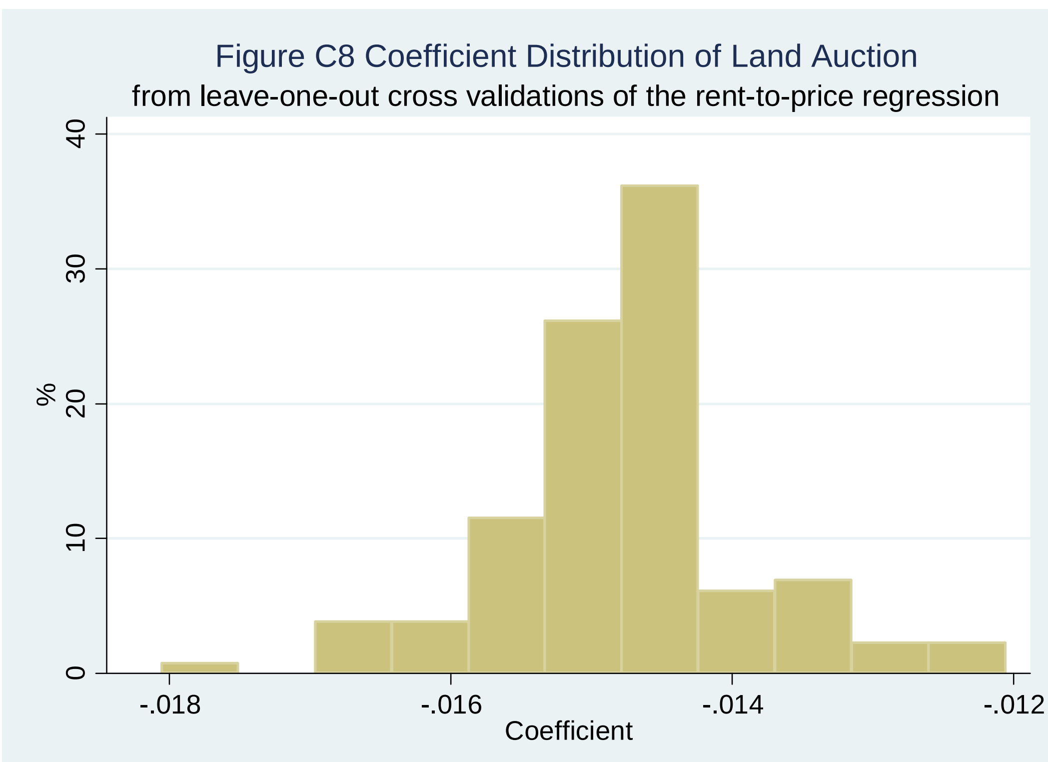


Figure C8 Coefficient Distribution of Land Auction
from leave-one-out cross validations of the rent-to-price regression

Figure C9 Coefficient Distribution of Young Population
from leave-one-out cross validations of the turnover regression

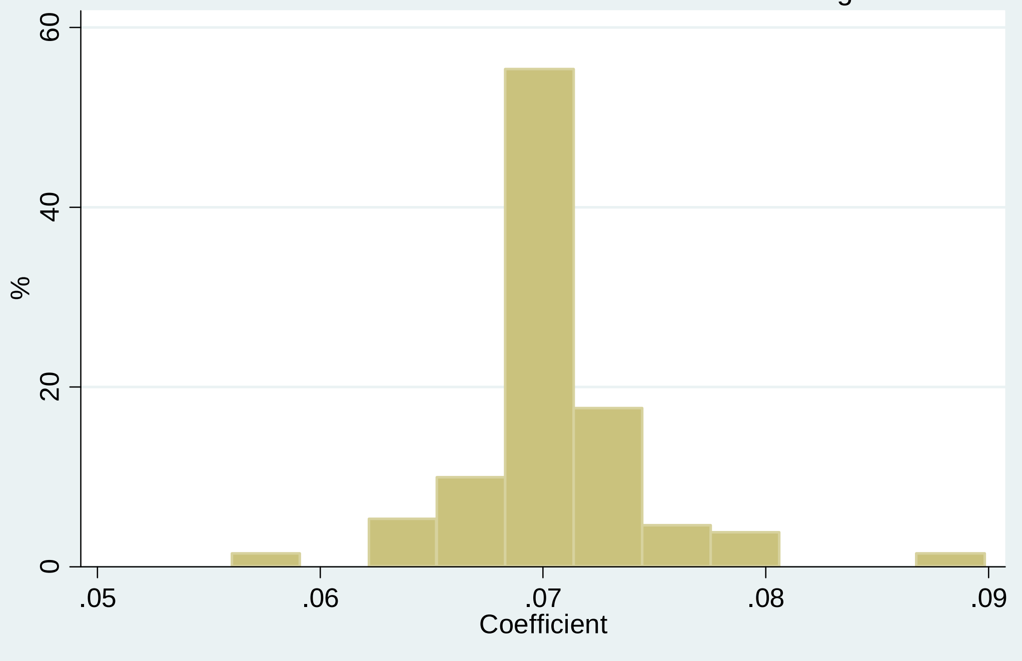


Figure C10 Coefficient Distribution of Household Income
from leave-one-out cross validations of the turnover regression

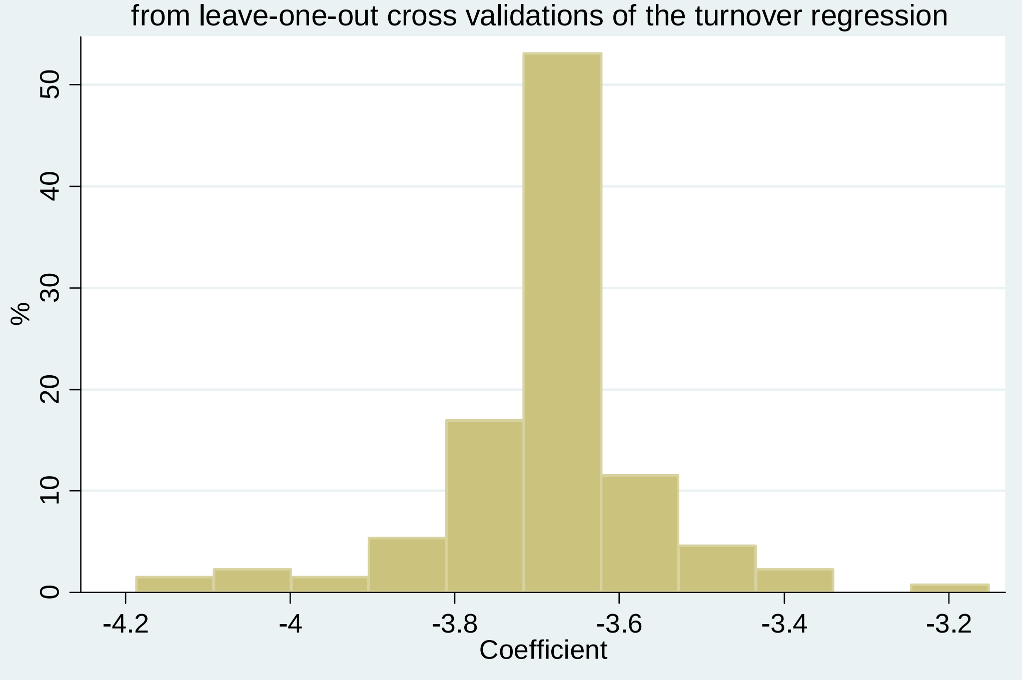


Figure C11 Coefficient Distribution of Real Estate Work Experience
from leave-one-out cross validations of the turnover regression

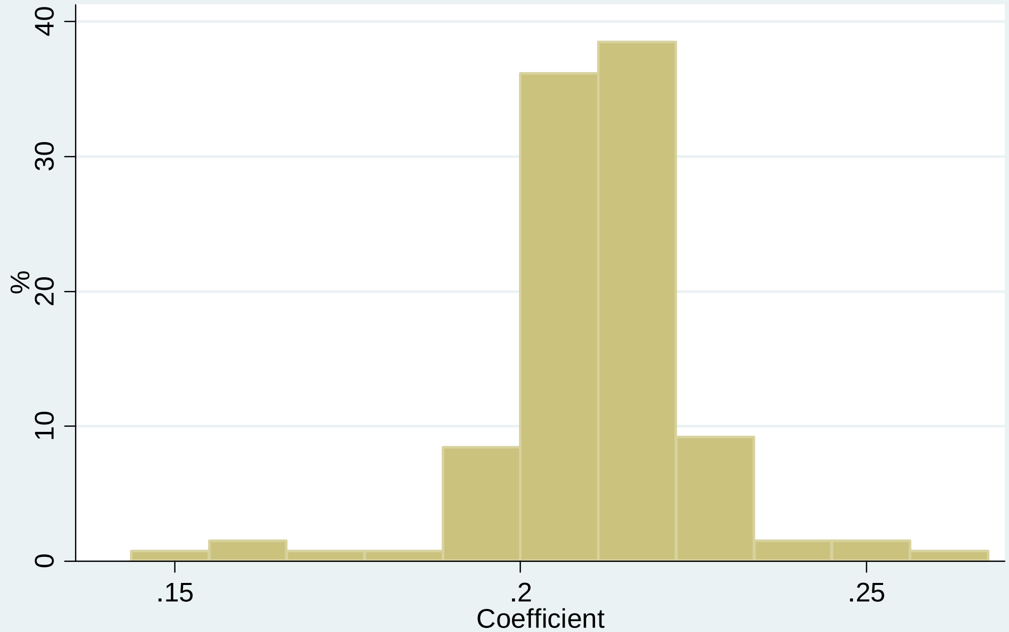


Figure C12 Coefficient Distribution of House Price Growth Rate
from leave-one-out cross validations of the turnover regression

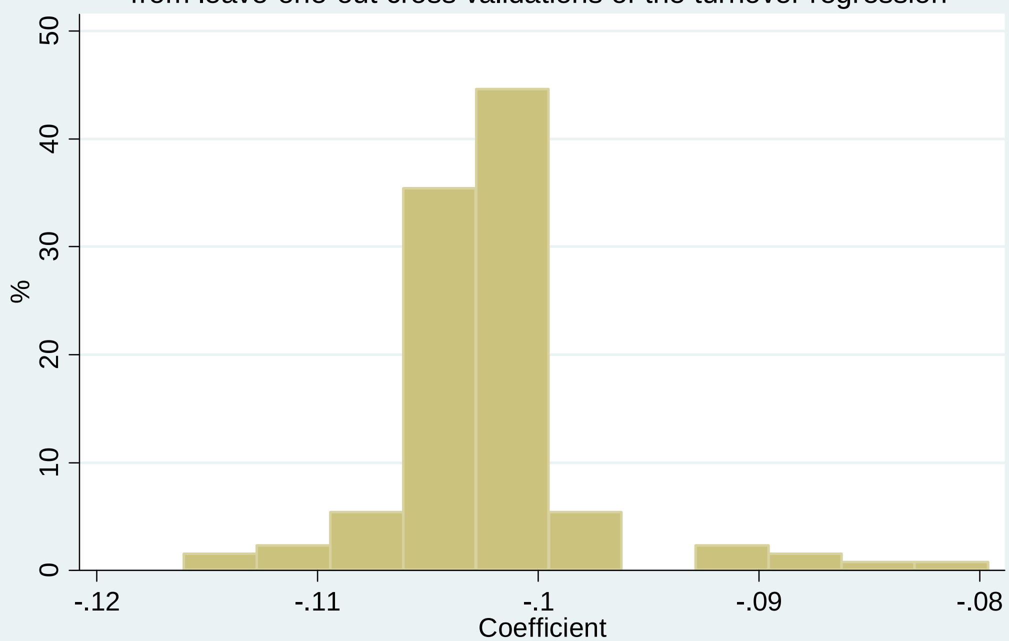


Figure C13 Coefficient Distribution of Rent Growth Rate
from leave-one-out cross validations of the turnover regression

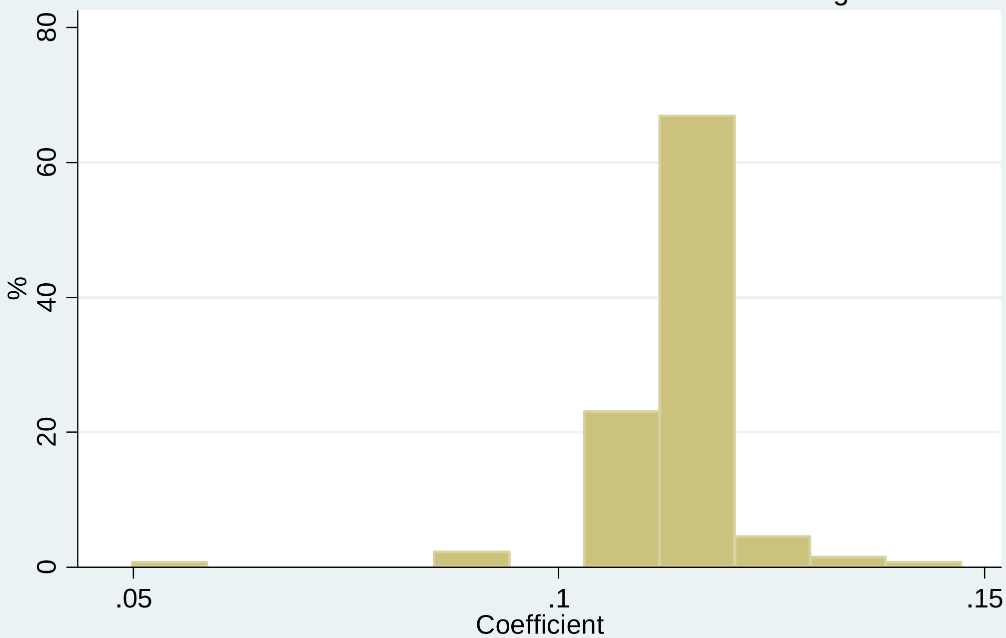
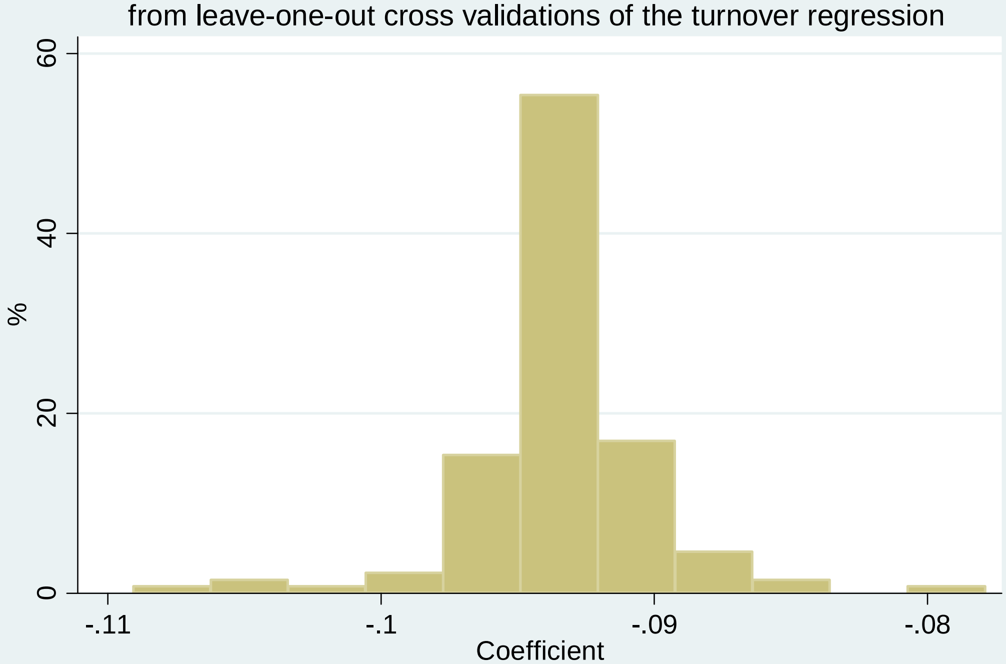


Figure C14 Coefficient Distribution of Building Age
from leave-one-out cross validations of the turnover regression



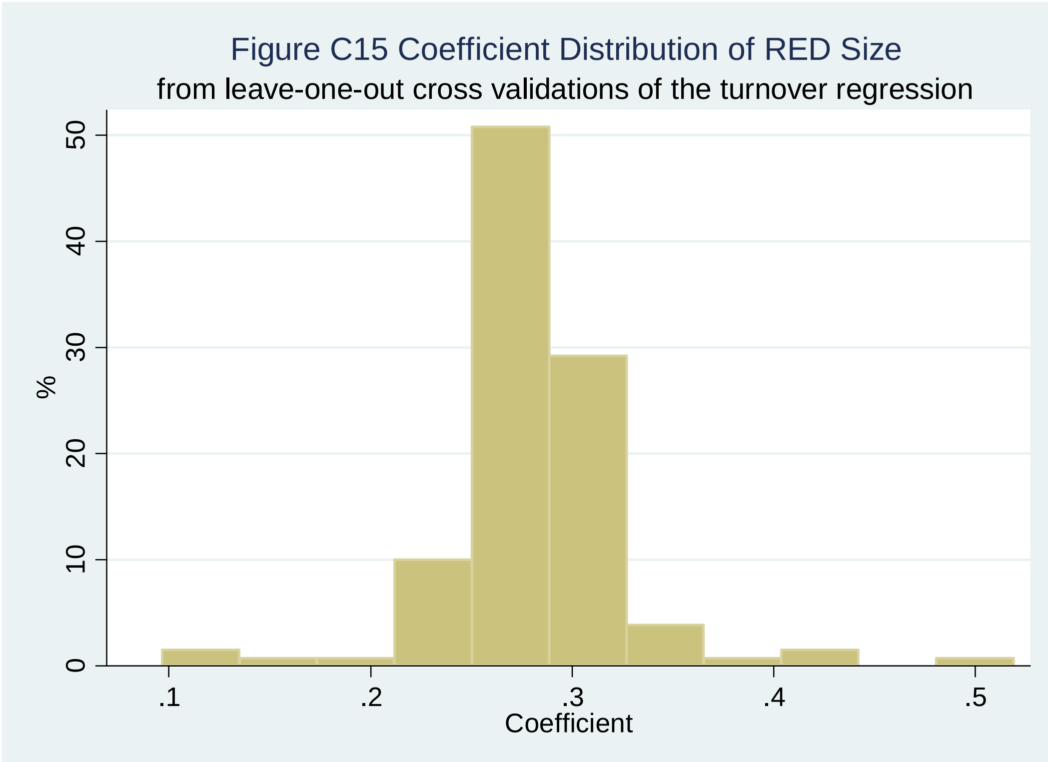


Figure C15 Coefficient Distribution of RED Size
from leave-one-out cross validations of the turnover regression
%
0
10
20
30
40
50
.1
.2
.3
.4
.5
Coefficient

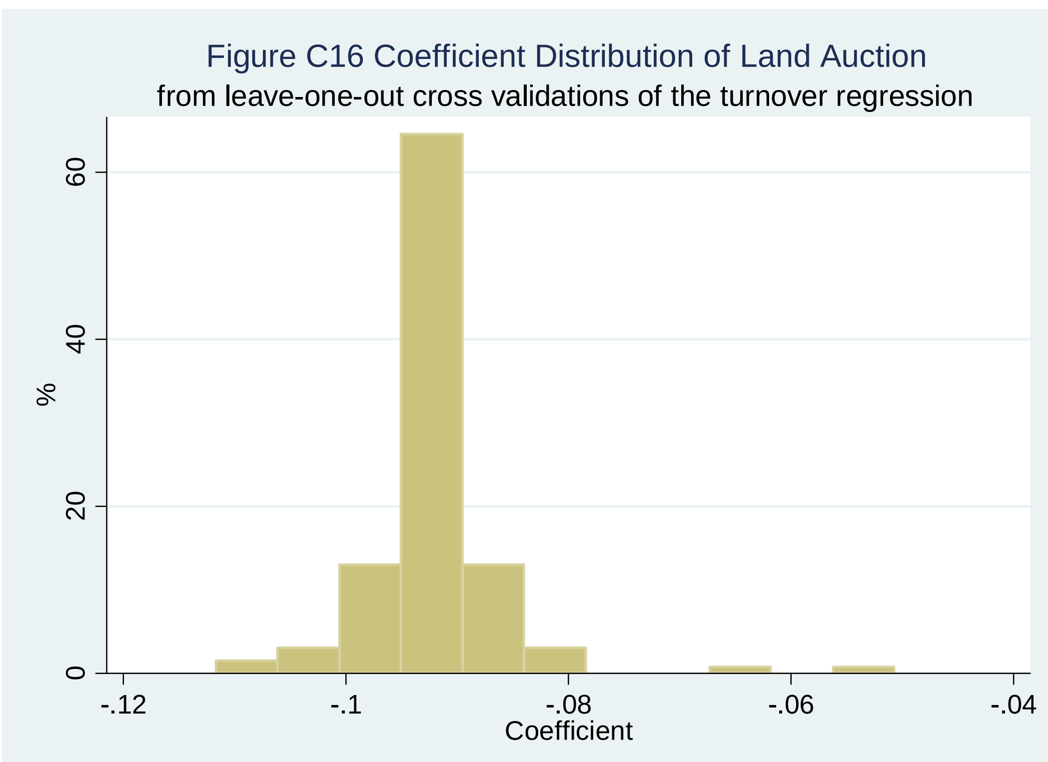


Figure C16 Coefficient Distribution of Land Auction
from leave-one-out cross validations of the turnover regression
%
0
20
40
60
-.12
-.1
-.08
-.06
-.04
Coefficient

Appendix D: Population Index

Figure D1 Coefficients of College-Educated Residents
from rent-to-price regressions with different popularity indices

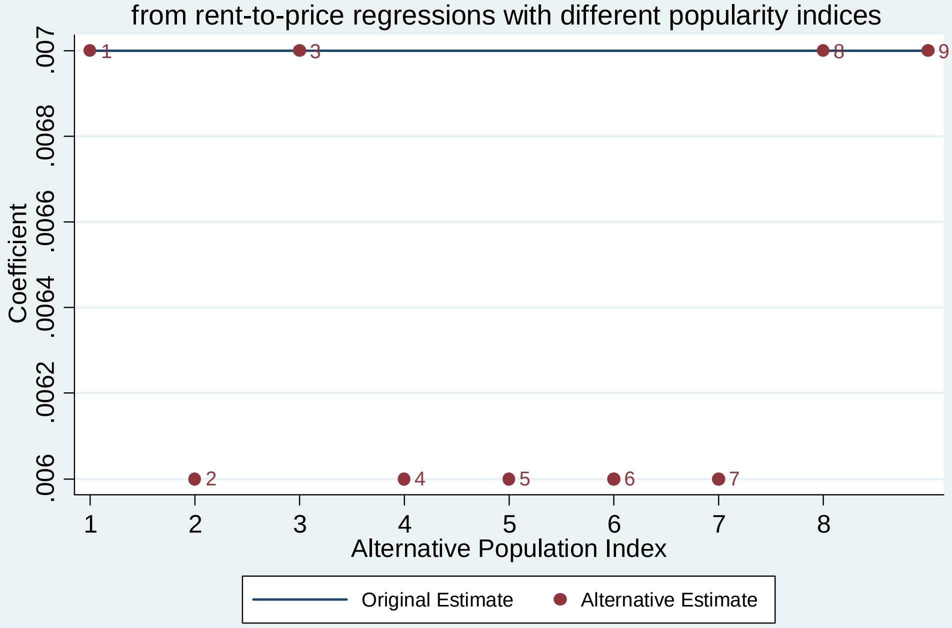


Figure D2 Coefficients of Household Income
from rent-to-price regressions with different popularity indices

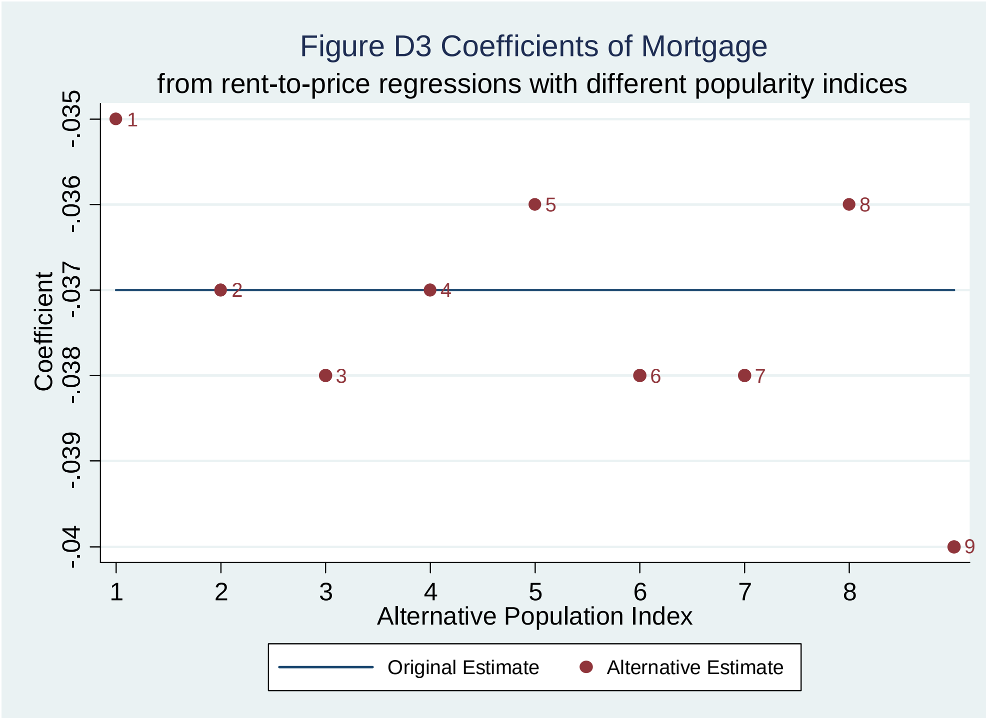


Figure D3 Coefficients of Mortgage
from rent-to-price regressions with different popularity indices
Coefficient
-.035
-.036
-.037
-.038
-.039
-.04
1
2
3
4
5
6
7
8
9
Alternative Population Index
Original Estimate
Alternative Estimate

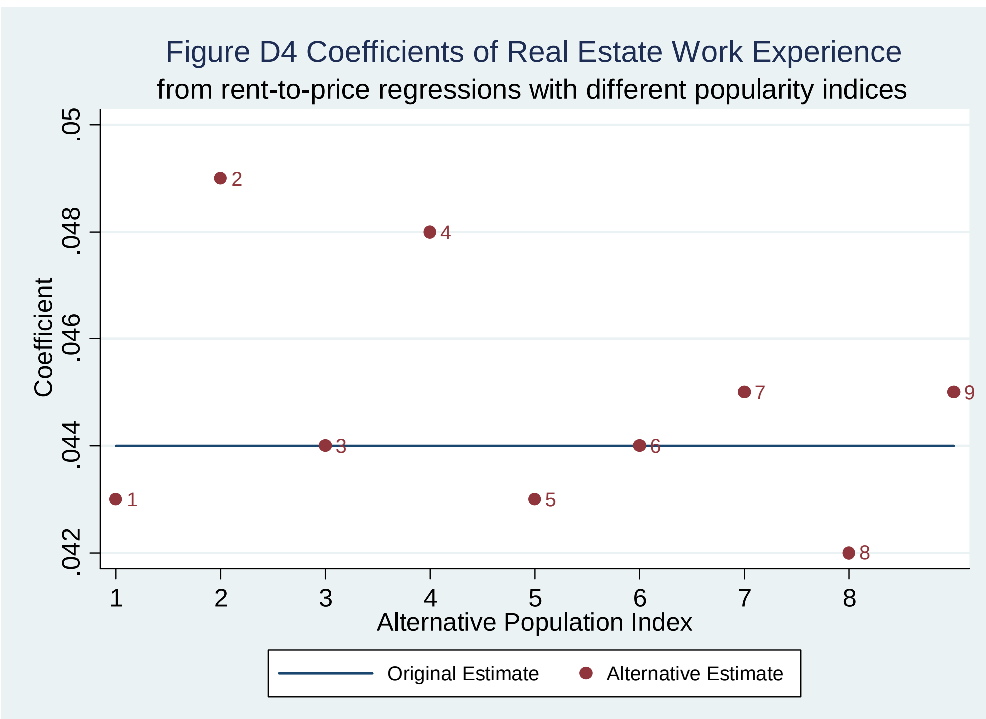


Figure D4 Coefficients of Real Estate Work Experience
from rent-to-price regressions with different popularity indices
Coefficient
.05
.048
.046
.044
.042
1
2
3
4
5
6
7
8
9
Alternative Population Index
Original Estimate
Alternative Estimate

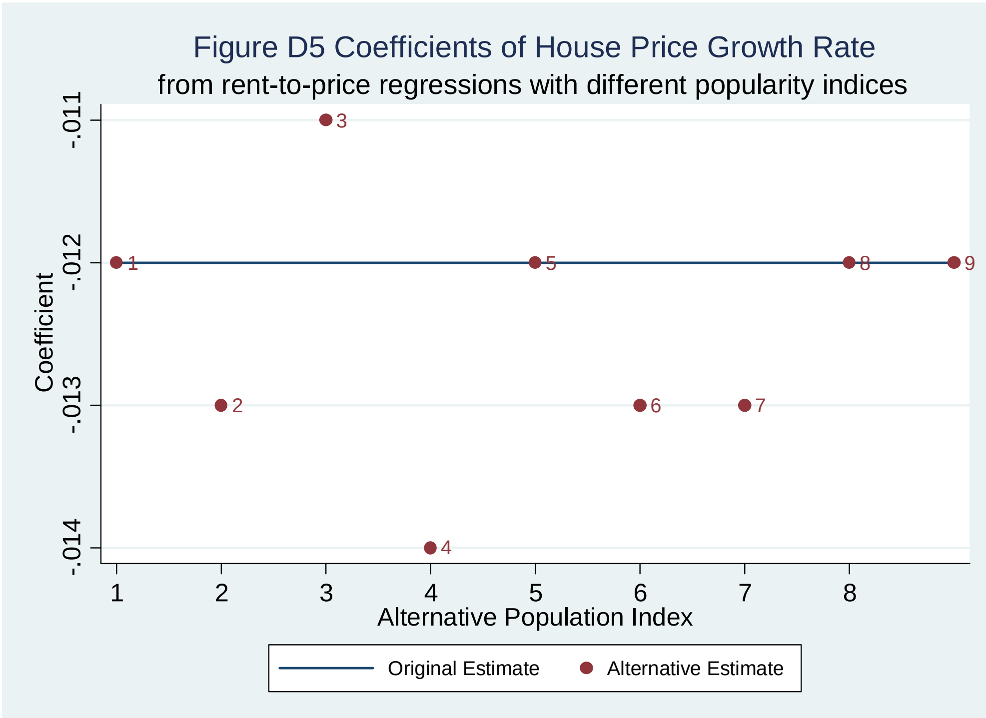


Figure D5 Coefficients of House Price Growth Rate
from rent-to-price regressions with different popularity indices
Coefficient
-.011
-.012
-.013
-.014
1
2
3
4
5
6
7
8
9
Alternative Population Index
Original Estimate
Alternative Estimate

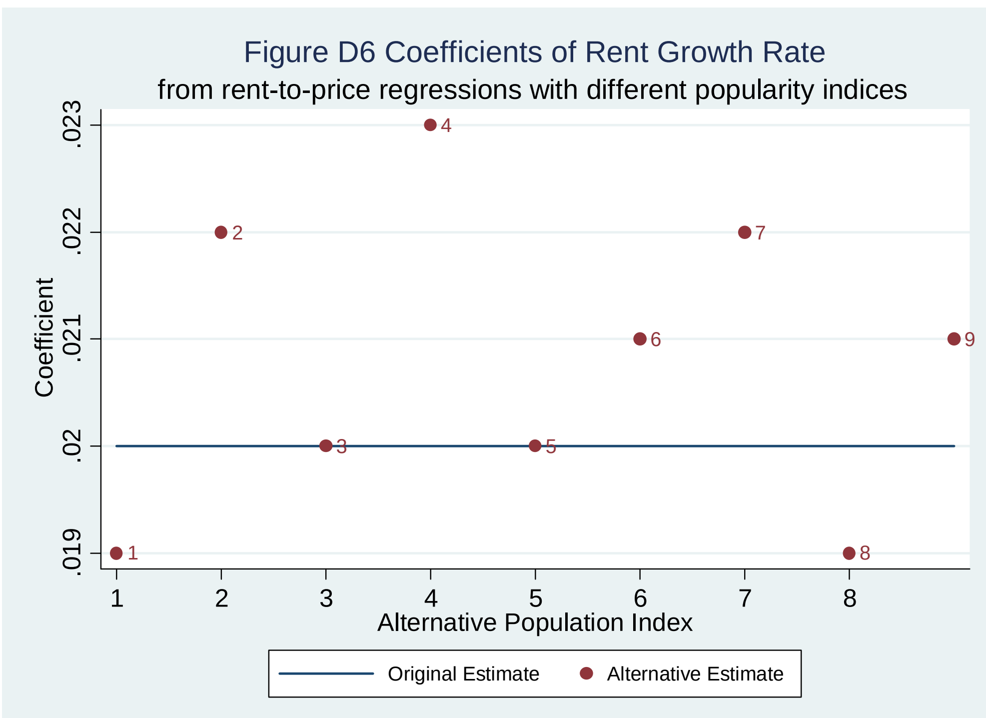


Figure D6 Coefficients of Rent Growth Rate
from rent-to-price regressions with different popularity indices
Coefficient
.023
.022
.021
.02
.019
1
2
3
4
5
6
7
8
9
Alternative Population Index
Original Estimate
Alternative Estimate

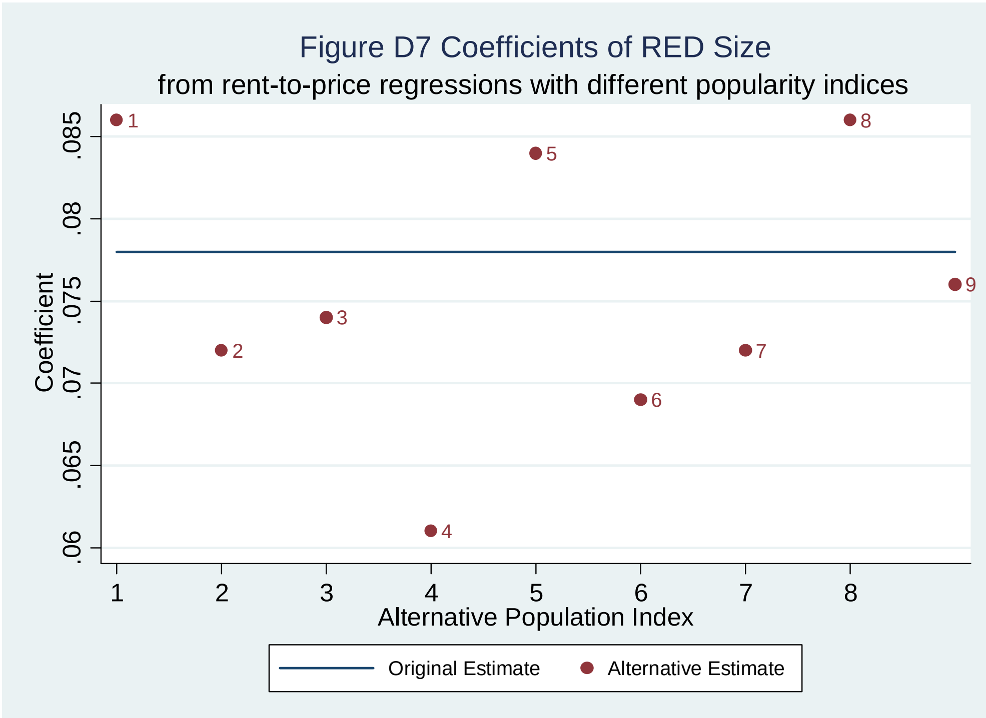


Figure D7 Coefficients of RED Size
from rent-to-price regressions with different popularity indices
Coefficient
.06
.065
.07
.075
.08
.085
1
2
3
4
5
6
7
8
9
Alternative Population Index
Original Estimate
Alternative Estimate

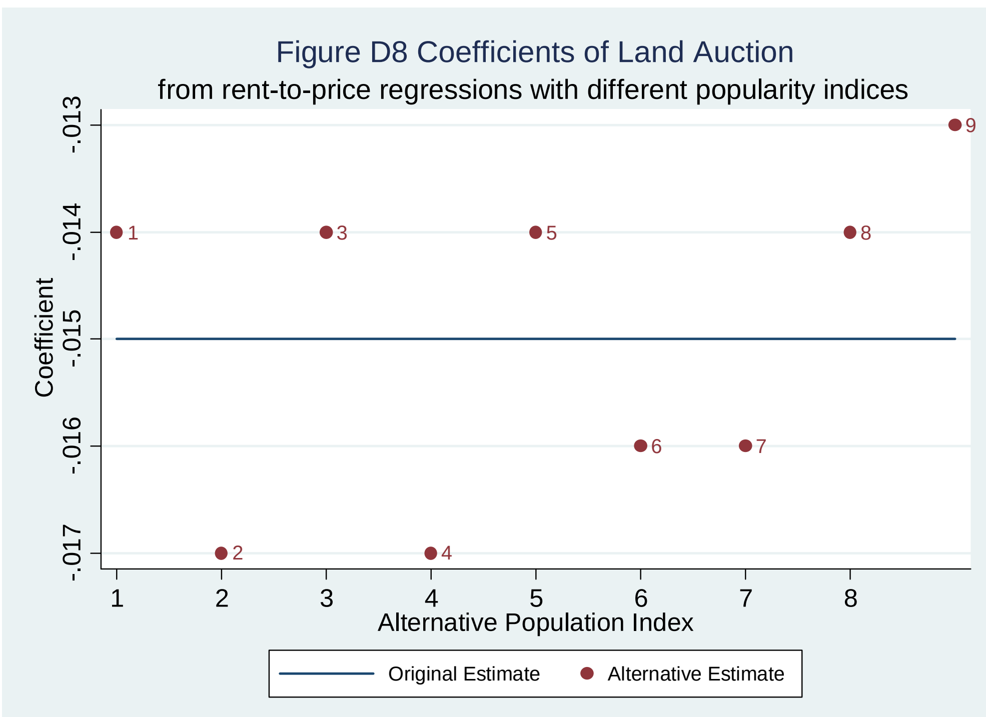


Figure D8 Coefficients of Land Auction
from rent-to-price regressions with different popularity indices
Coefficient
-.017
-.016
-.015
-.014
-.013
1
2
3
4
5
6
7
8
9
Alternative Population Index
Original Estimate
Alternative Estimate

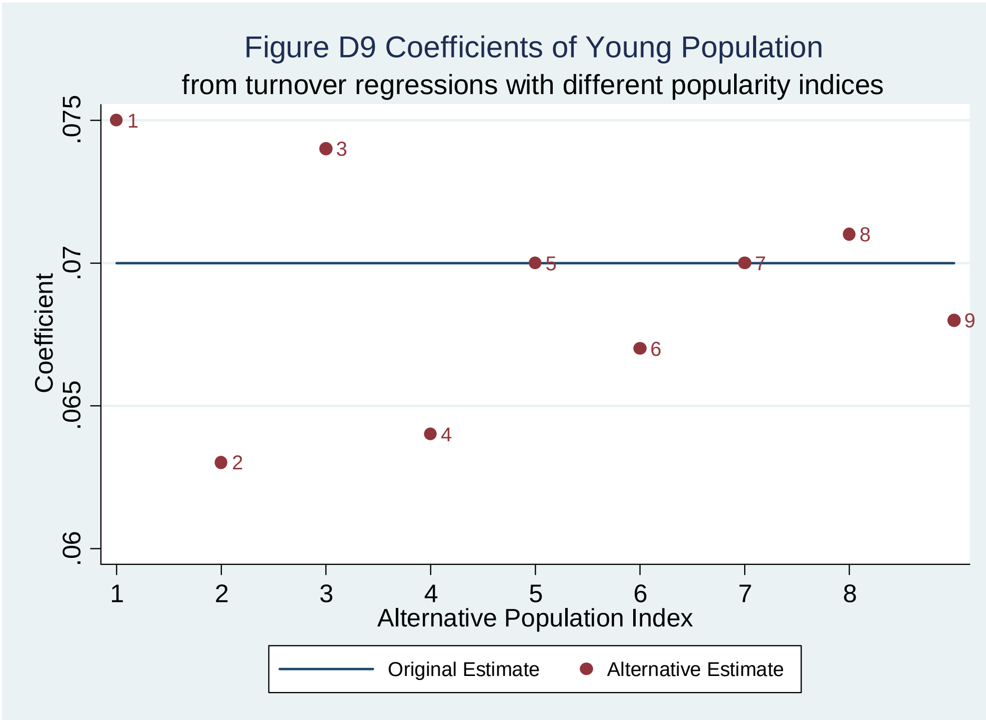


Figure D9 Coefficients of Young Population
from turnover regressions with different popularity indices
Coefficient
.06
.065
.07
.075
1
2
3
4
5
6
7
8
9
Alternative Population Index
Original Estimate
Alternative Estimate

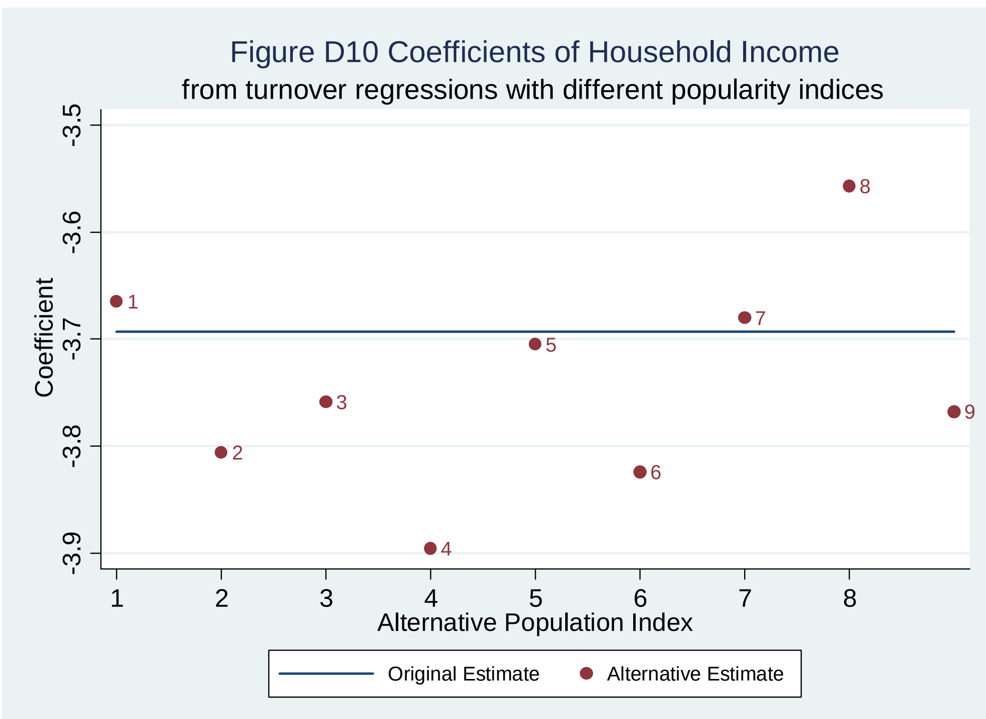


Figure D10 Coefficients of Household Income
from turnover regressions with different popularity indices
Coefficient
-3.9
-3.8
-3.7
-3.6
-3.5
1
2
3
4
5
6
7
8
9
Alternative Population Index
Original Estimate
Alternative Estimate

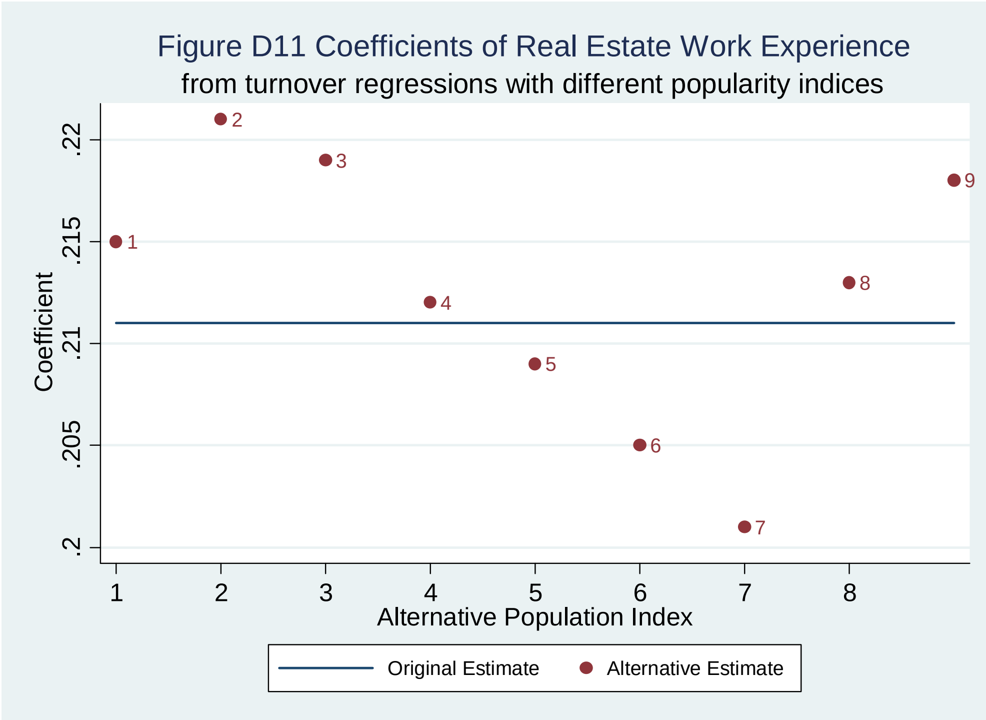


Figure D11 Coefficients of Real Estate Work Experience
from turnover regressions with different popularity indices
Coefficient
.2
.205
.21
.215
.22
1
2
3
4
5
6
7
8
9
Alternative Population Index
Original Estimate
Alternative Estimate

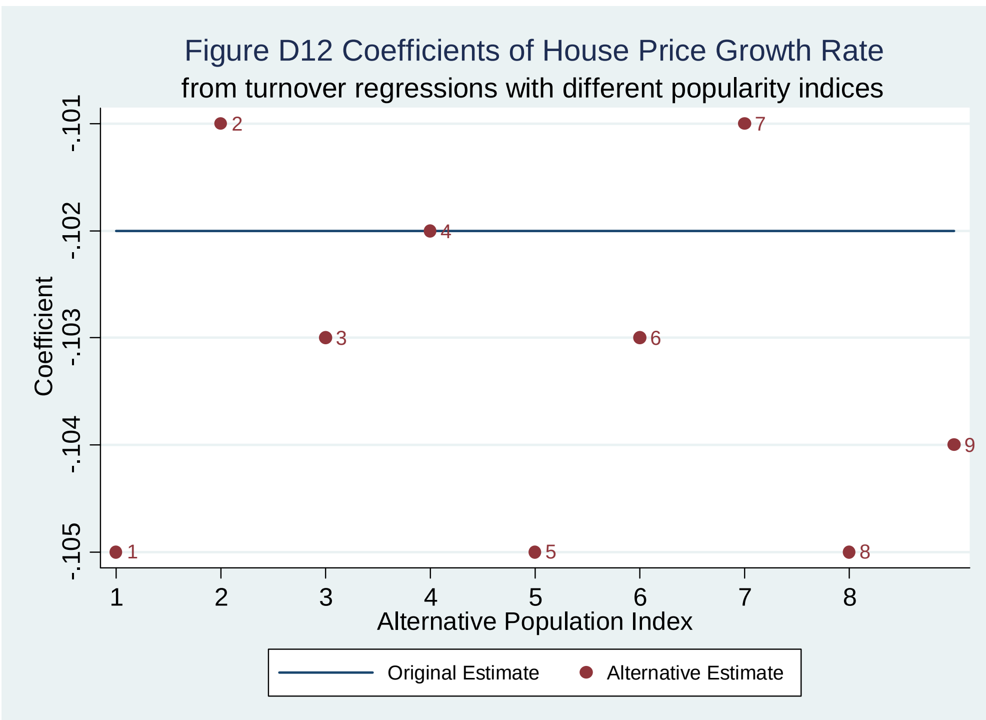


Figure D12 Coefficients of House Price Growth Rate
from turnover regressions with different popularity indices
Coefficient
-.105
-.104
-.103
-.102
-.101
1
2
3
4
5
6
7
8
9
Alternative Population Index
Original Estimate
Alternative Estimate

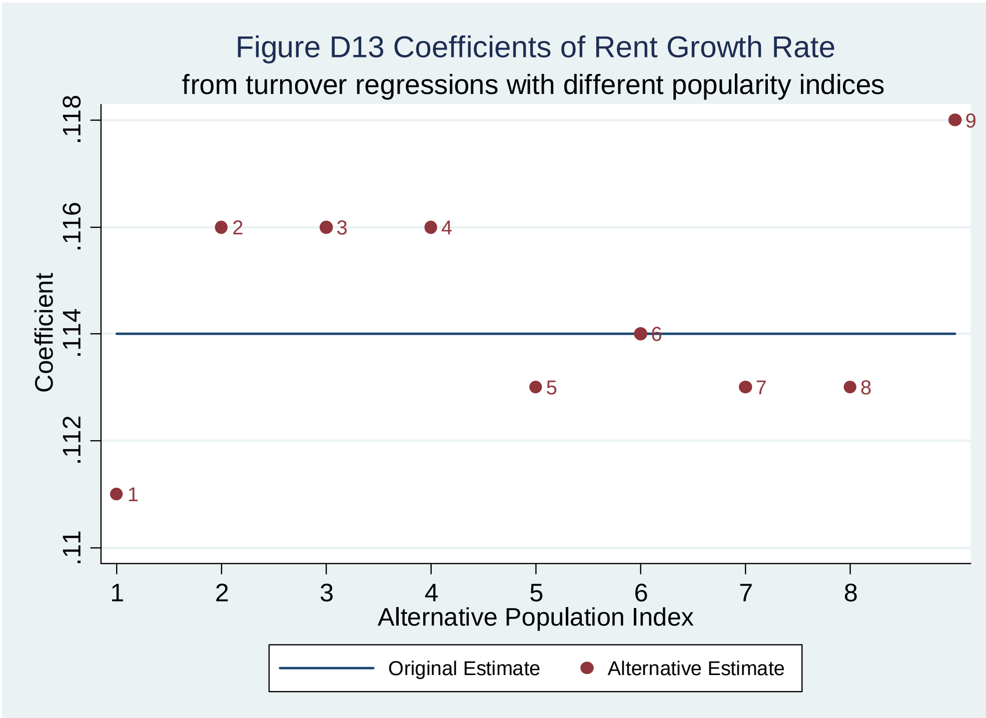


Figure D13 Coefficients of Rent Growth Rate
from turnover regressions with different popularity indices
Coefficient
.11
.112
.114
.116
.118
1
2
3
4
5
6
7
8
9
Alternative Population Index
Original Estimate
Alternative Estimate

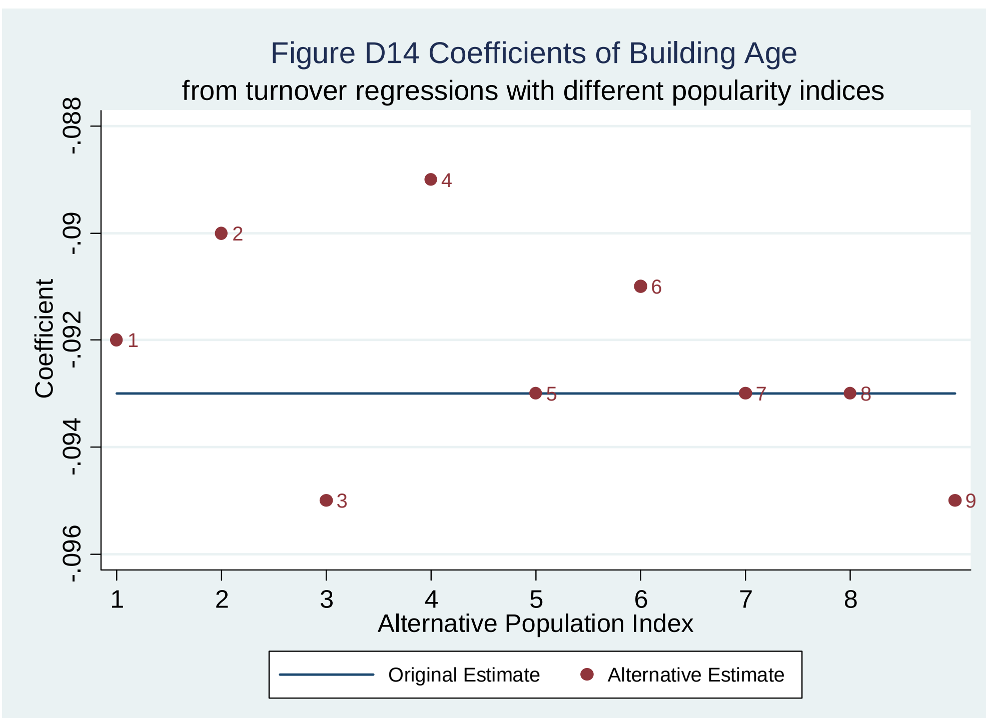


Figure D14 Coefficients of Building Age
from turnover regressions with different popularity indices
Coefficient
-.096
-.094
-.092
-.09
-.088
1
2
3
4
5
6
7
8
9
Alternative Population Index
Original Estimate
Alternative Estimate

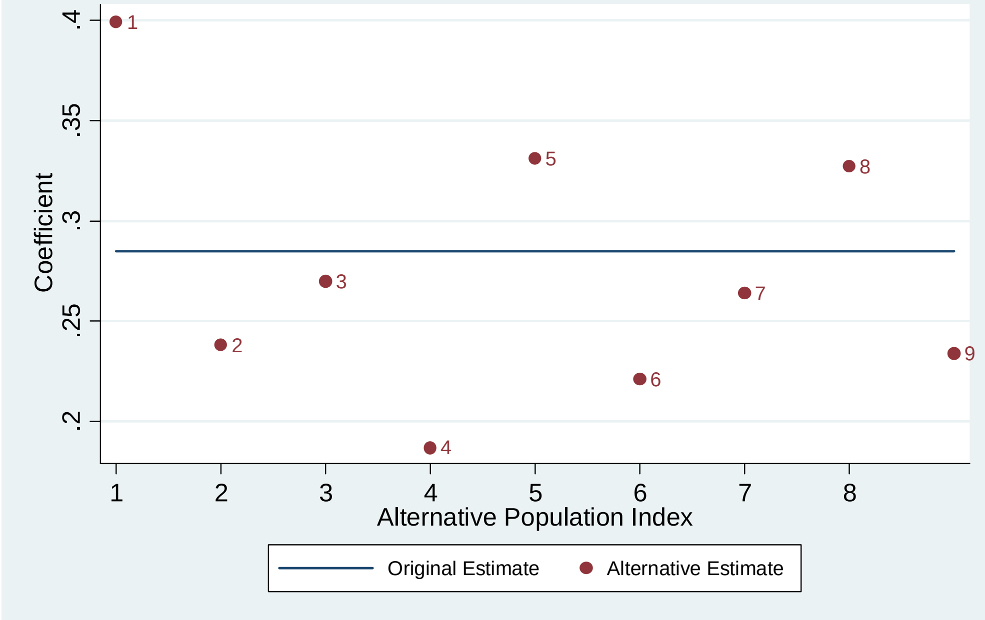


Figure D15 Coefficients of RED Size
from turnover regressions with different popularity indices

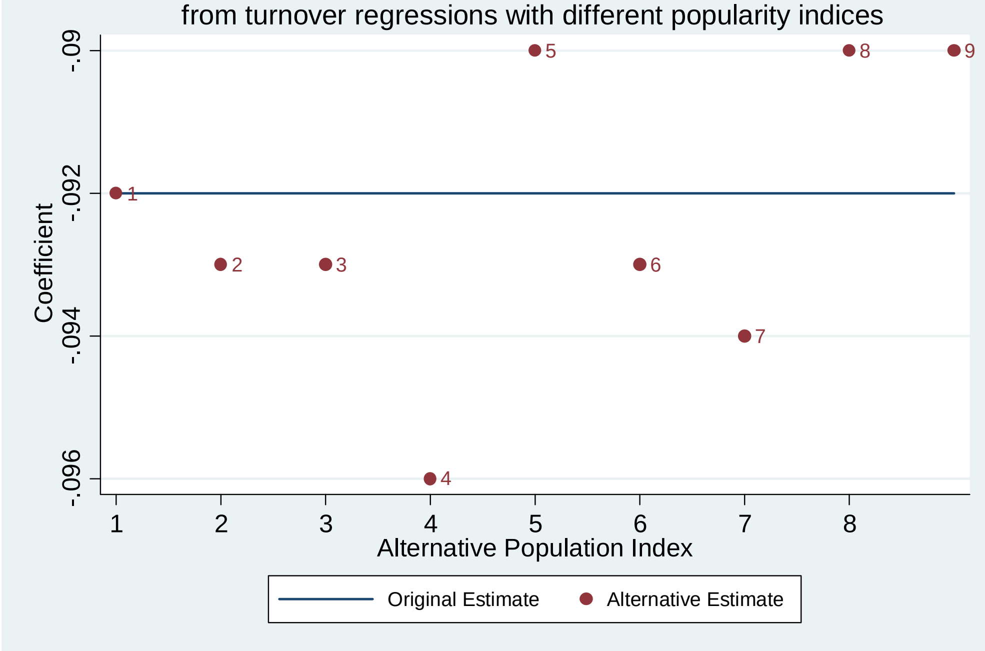


Figure D16 Coefficients of Land Auction
from turnover regressions with different popularity indices

Appendix E: More on our popularity index

This appendix provides a correlation table among different variables and some discussion of the construction of our popularity index, and provides a complete list of that index on 132 major RED in Hong Kong.

We first provide the correlation table, Table E1, which is also known as the correlation matrix. Notice that *not* all variables are highly correlated. It should not be surprising, as Table E1 mixes variables which represent housing characteristics and variables which represent household characteristics. A more sensible comparison would be to focus on the correlations among housing characteristics, which is provided by Table E2. It shows that housing characteristics are indeed statistically correlated. One might therefore be tempted to construct principal components (PC) to summarize the co-variations of the variables. Unfortunately, *statistically significance does not always translate to economic significance* and our sample provides a good illustration for this principle. As shown by Table E3, the first principal component explains only 20% of the total variations. We need to include from the first to the sixth principal components to explain 80% of the total variations. However, we only have a small sample of RED and we cannot afford to 6 PC into the regression and at the same time be able to identify different effects given the 3SLS framework. Therefore, *the formative index approach may be constrained optima* in terms of the econometric strategy.

Table E1: Correlation Matrix

Number	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
1																	
2	0.25																
3	0.11	-0.12															
4	-0.14	0.40	-0.32														
5	0.17	-0.33	0.22	-0.89													
6	0.17	-0.36	0.37	-0.79	0.81												
7	-0.43	-0.08	-0.02	0.33	-0.29	-0.20											
8	-0.53	-0.32	-0.06	-0.06	0.03	-0.02	0.61										
9	-0.60	-0.34	0.02	-0.13	0.08	0.07	0.54	0.93									
10	-0.44	-0.03	-0.07	0.12	-0.14	-0.13	0.38	0.34	0.45								
11	0.13	-0.21	0.03	-0.33	0.24	0.23	-0.35	-0.21	-0.25	-0.42							
12	-0.43	-0.09	0.07	0.28	-0.29	-0.21	0.55	0.40	0.41	0.28	-0.28						
13	-0.08	-0.36	0.33	-0.49	0.41	0.47	-0.03	0.20	0.25	-0.05	0.17	0.15					
14	0.47	-0.25	0.25	-0.57	0.56	0.65	-0.50	-0.29	-0.20	-0.33	0.23	-0.41	0.42				
15	0.13	-0.06	0.97	-0.26	0.20	0.34	-0.01	-0.09	-0.02	-0.04	-0.07	0.06	0.24	0.21			
16	0.23	0.22	-0.15	-0.02	0.14	-0.04	-0.33	-0.40	-0.47	-0.25	0.28	-0.36	-0.14	-0.05	-0.10		
17	0.14	0.17	-0.06	0.05	-0.03	-0.06	-0.17	-0.23	-0.23	-0.15	0.05	-0.42	-0.25	0.10	-0.03	0.15	
18	-0.26	-0.08	0.01	0.07	-0.18	-0.16	0.21	0.15	0.22	0.25	-0.19	0.35	-0.09	-0.12	-0.01	-0.41	0.03

Note: Variables with correlations lower than 0.4 to other variables are left out to save space.

Key:

Number	Variables
1	Mean R/P of the RED
2	Turnover rate %
3	Population Resided in the RED
4	Young population %
5	Middle Age Population %
6	Median Age of the Residents
7	College education level %
8	Monthly Median Income of Residents HKD
9	Median Monthly Mortgage Payment HKD
10	Mortgage income ratio %
11	Home Ownership Rate in the RED %
12	Popularity of the RED
13	Average Net Gross Ratio %
14	Average building age
15	Estate Scale
16	Distance to CBD M
17	Distance to Seafront M
18	# Top 100 Middle School

Table E2: Correlations among Housing Characteristics

	Variables	1	2	3	4	5	6	7	8	9
1	Affordability									
2	Access to MTR within 500 Meters	0.01*								
3	Distance to the CBD	-0.32***	-0.02***							
4	Net Gross Ratio	0.16***	-0.02***	-0.19***						
5	# of Top 100 Middle Schools	0.11***	0.25***	-0.33***	-0.06***					
6	Developer Reputation	0.00	-0.09***	-0.02***	0.02***	-0.14***				
7	Access to Waterfront within 500 Meters	0.23***	-0.22***	-0.24***	0.20***	0.05***	-0.07***			
8	RED Has Club House	0.13***	0.05***	0.09***	-0.24***	0.12***	-0.03***	0.14***		
9	RED Has Swimming Pool	0.03***	0.05***	0.24***	-0.19***	-0.01*	0.05***	0.04***	0.49***	

* p<0.05; ** p<0.01; *** p<0.001

Table E3: Principal Component Analysis

Measures	Comp1	Comp2	Comp3	Comp4	Comp5	Comp6
Affordability	-0.347	0.369	0.156	0.215	0.236	-0.719
Access to MTR within 500 Meters	0.057	0.160	-0.610	0.201	0.508	0.080
Distance to the CBD	0.548	-0.196	0.117	-0.193	0.295	-0.016
Net Gross Ratio	-0.429	-0.149	0.167	0.025	0.644	0.389
# of Top 100 Middle Schools	-0.210	0.381	-0.486	0.012	-0.314	0.331
Developer Reputation	0.046	-0.122	0.260	0.877	-0.170	0.236
Access to Waterfront within 500 Meters	-0.322	0.291	0.439	-0.311	-0.055	0.343
RED Has Club House	0.285	0.582	0.150	0.004	-0.004	0.038
RED Has Swimming Pool	0.401	0.440	0.204	0.102	0.233	0.200
Cumulative proportion of variance	0.204	0.388	0.540	0.653	0.752	0.833

Now we need to discuss our "popularity index." While it is beyond the scope of this paper to review the literature, it may nevertheless be instructive to highlight some of the key findings. Measurement models can be divided into two types, reflective, where the term of "scale development" is usually used, and formative, where the phrase of "index construction" is more often employed. Many scholars have discussed that, including Diamantopoulos and Winklhofer (2001) and Jarvis, MacKenzie, and Podsakoff (2003), among others.

In reflective models, the composite variable is cause, and indicators composing it are effects. The reflective measurement model is essentially seeking common information among realized results. Since all effects stem from a common cause, there must be collinearity among them. It follows that data dimension reduction technique such as principal component or factor analysis best applies to the context. In contrast, cause and effect are reversed in a formative model, i.e. indicators produce the composite variable they compose. Researchers attempt to summarize an outcome considering all possible causes. In fact, formative measurement models have been applied in economics and other social sciences. Daly and Cobb (1994) develop the sustainable economic welfare index. United Nations Development Program builds the human development index in 1990. Johnston (1988) constructs the quality-of-life index. In the case of real estate studies, Griliches (1971) builds the hedonic price indices. More examples can be found in Horn (1993).

Clearly, different causes are not necessarily highly correlated. For such reasons, principal component kind of analysis may not be practical. As a matter of fact, we try the principal component analysis on our data. As we have shown, we need to retain as many as six PCs to capture less than 80% of variances in our eleven indicators, which does not seem to be a promising approach given our small sample size and 3SLS procedure.

A simple alternative is the linear combination of formative indicators. In our case, popularity is obviously a consequence variable determined by underlying items, suggesting developing the popularity index as a formative measurement model may be more promising. Specifically, our popularity is a linear sum of equally weighted indicators selected.

Appendix F: 3SLS estimation with district level data

This appendix tries to incorporate some district level data in the 3SLS regression to account for the district specific effect. Such data come from the 2011 Census and 2006 by-census. At first, we attempt to include a number of variables in the specification, such as population and income growth rates between two census periods, population density, average household size, percent of college degree holders, the median rent and median mortgage payment, as well as proportion of foreign population in 2011. We even use four measures to proxy for possible foreigners' effect: new immigrants from mainland China with duration of residence less than seven years, population born outside of Hong Kong, population whose place of domicile is not Hong Kong, and population whose first language is not Cantonese. Yet after experimenting with different subsets of these variables, we find that at district level *only the population density and income growth rate have explanatory power* over property level rental yield. Therefore, only these two variables are introduced in the regression and the results are reported in Table F.1.

We first discuss the rental yield regression results. Comparing the RMSE statistics in the tables in the main text and other appendices, the goodness-of-fit of the rent-price ratio regression slightly improves. The new results are that both higher population density and higher income growth in the district will predict a lower RED-level rent-to-price ratio. On the other hand, the inclusion of the district data tempers both the economic magnitude and the statistical significance of normed transaction and popularity index estimates. Other variables remain significant as well. The numerical values of estimates do not change much.

The goodness-of-fit of the turnover rate equation even degenerates. The statistical significance of other variables also diminishes. The population density and the income growth rate are merely marginally significant. The rent growth rate becomes insignificant as well, which is supportive evidence of our finding that long run rent growth influences rent to price ratio while trading volume is more affected by short-term return in price. Only the income, young population proportion, and lagged price appreciation, are statistically significantly different from zero.

Table F1: Results of 3SLS with District Data

	Coef.	Beta Coef.	Std. err.	pval.
Equation: R/P				
Turnover Rate %	-0.047	-0.315	0.031	0.122
Popularity of the RED	-0.043	-0.116	0.040	0.282
College Education Level %	0.006*	0.147*	0.003	0.066
Ln(Income)	-0.635***	-0.534***	0.137	0.000
Mortgage Income Ratio %	-0.031**	-0.170**	0.014	0.028
Work in Real Estate Consulting %	0.046**	0.193**	0.018	0.012
Last Year HP Growth Rate %	-0.007	-0.091	0.007	0.310
Rent Growth Rate in Past 5 YRs	0.017*	0.143*	0.009	0.051
RED Scale	0.104*	0.133*	0.059	0.077
Land Auction	-0.014	-0.113	0.009	0.136
District Population Density	-0.107**	-0.247**	0.046	0.019
District Income Growth	-0.021*	-0.150*	0.013	0.095
Constant	11. 543***	-0.014	1.796	0.000
Observations	130	130		
RMSE	0.367	0.771		
Chi-squared	111.60	111.60		
P value	0.000	0.000		
Equation: Turnover				
Mean R/P of the RED	-2.735	-0.412	3.656	0.454
Popularity of the RED	-0.192	-0.077	0.332	0.564
Young Population %	0.082*	0.206*	0.044	0.063
Ln(Income)	-3.776**	-0.479**	1.618	0.020
Work in Real Estate Consulting %	0.319	0.201	0.197	0.105
Last Year HP Growth Rate %	-0.095*	-0.177*	0.054	0.079
Rent Growth Rate in Past 5 YRs	0.127	0.161	0.080	0.112
Average Building Age	-0.093	-0.256	0.071	0.191
RED Scale	0.569	0.110	0.611	0.352
Land Auction	-0.109	-0.137	0.086	0.208
District Population Density	-0.855	-0.297	0.554	0.122
District Income Growth	-0.191	-0.202	0.126	0.129
Constant	61.300**	0.003	29.953	0.041
Observations	130	130		
RMSE	2.785	0.881		
Chi-squared	65. 53	65. 53		
P value	0.000	0.000		

Appendix G: Robustness of the variable choices on X and Z.

In the main text of the paper, we put "Mortgage Income ratio" among the R/P equation control variables, and "Young Population %" among the turnover equation control variables. It has been suggested to us that the mortgage-income ratio could also matter for turnover. The idea is that a higher value of the mortgage-income ratio might imply that people are less mobile. To check for the robustness of our results, we move that variable to the turnover equation. To apply 3SLS methodology, we move the young population percentage variable to the R/P equation. Results indicate that both variables become insignificant after switching. Moreover, the overall performances of both equations deteriorate significantly. We therefore maintain our original variable choices in the text.

Table G1: Results of 3SLS with control variable changes

	Coef.	Beta Coef.	Std. err.	pval.
Equation: R/P				
Turnover Rate %	-0.820	-5.442	2.259	0.717
Popularity of the RED	-0.103	-0.276	0.533	0.847
College Education Level %	-0.035	-0.851	0.096	0.711
Ln(Income)	-2.500	-2.104	4.806	0.603
Average Building Age	-0.069	-1.270	0.195	0.722
Young Population %	0.051	0.849	0.257	0.844
Work in Real Estate Consulting %	0.198	0.827	0.372	0.594
Last Year HP Growth Rate %	-0.105	-1.295	0.260	0.687
Rent Growth Rate in Past 5 YRs	0.095	0.801	0.222	0.670
RED Scale	0.221	0.283	0.939	0.814
Land Auction	-0.081	-0.675	0.197	0.681
Constant	36.849	0.075	64.116	0.565
Observations	130	130		
RMSE	2.173	4.565		
Chi-squared	5.79	5.79		
P-value	0.887	0.887		
Equation: Turnover				
Mean R/P of the RED	-11.107*	-1.673*	6.174	0.072
Popularity of the RED	-0.489	-0.198	0.645	0.448
Mortgage Income Ratio %	-0.385	-0.323	0.312	0.216
Ln(Income)	-8.299***	-1.052***	2.964	0.005
Last Year HP Growth Rate %	-0.202**	0.358*	0.081	0.012
Rent Growth Rate in Past 5 YRs	0.246*	-0.376**	0.135	0.068
Work in Real Estate Consulting %	0.570*	0.314*	0.301	0.058
RED Scale	0.959	0.185	1.114	0.389
Land Auction	-0.194	-0.243	0.124	0.119
Constant	140.648**	-0.004	54.318	0.010
Observations	130	130		
RMSE	4.934	1.562		
Chi-squared	17.13	17.13		
P-value	0.047	0.047		

Appendix H: Possible interactions possible interactions between the labor and housing markets

It has been suggested to us that there may be non-trivial interactions between the labor and housing markets. While we agree with this conjecture theoretically, the current manuscript provides limited coverage on that. Here are the reasons. First, our focus is on the interactions between the rent-to-price ratio and the turnover rate. It is not clear to us how the interactions between these two variables would affect the labor market and then feed back to the housing market. Second, we have a *data issue*. Since our empirical works mainly focus on cross-sectional comparison across different real estate developments (RED), it is natural to seek for RED level unemployment rate data. However, in the case of Hong Kong, we *do not even have district level unemployment rate data*. Even if we had the district level unemployment rate data, we may not be able to rule out the possibility that different REDs within the same district exhibit different unemployment rates. Hence, it is difficult for us to conduct much cross-sectional analysis on the unemployment rate and property market rent-price ratio (or property yield). All we have from the official data is a wage index constructed for the whole Hong Kong. We nevertheless explore whether such wage index is correlated to the housing market yield in time series. Recent works such as Chang et al. (2013), Leung et al. (2013) suggest that Hong Kong macroeconomic data may display regime-switching and structural change. If this is indeed the case, full sample correlation coefficient could be mis-leading. Following the literature, we conduct rolling window estimation instead. Figure H1 shows the subsample correlations between the unemployment rate and property yield evolve over time (please see the following pages). Notice that the correlation between the two variables does not seem to be stable. It can exceed well above 0.5 in some periods and drops below -0.5 in some other periods. To model the relationship between these variables may not be easy. Similarly, the subsample correlations between the real wage and the property yield also vary over time. As shown in Figure H2, while the correlations are mostly positive, we do observe several subsamples exhibiting negative correlations. Among the positive correlations, some are small in magnitude and not statistically significant.

Figure H1: Rolling Window Correlation Coefficients of Unemployment Rate and Property Yields

(Rolling Window: 24 Months)

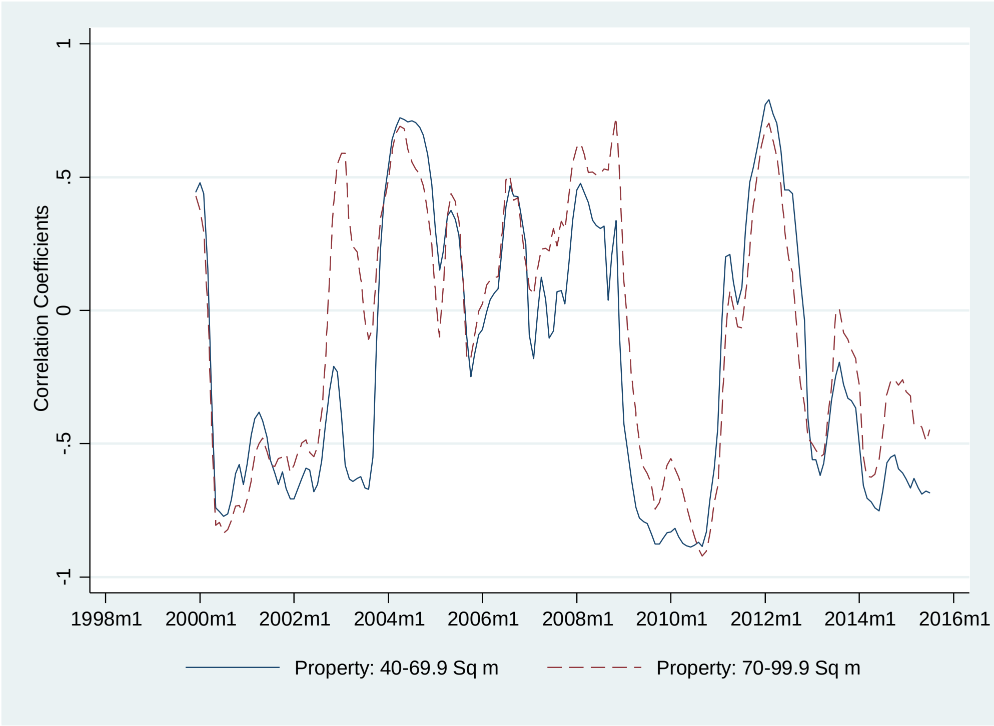
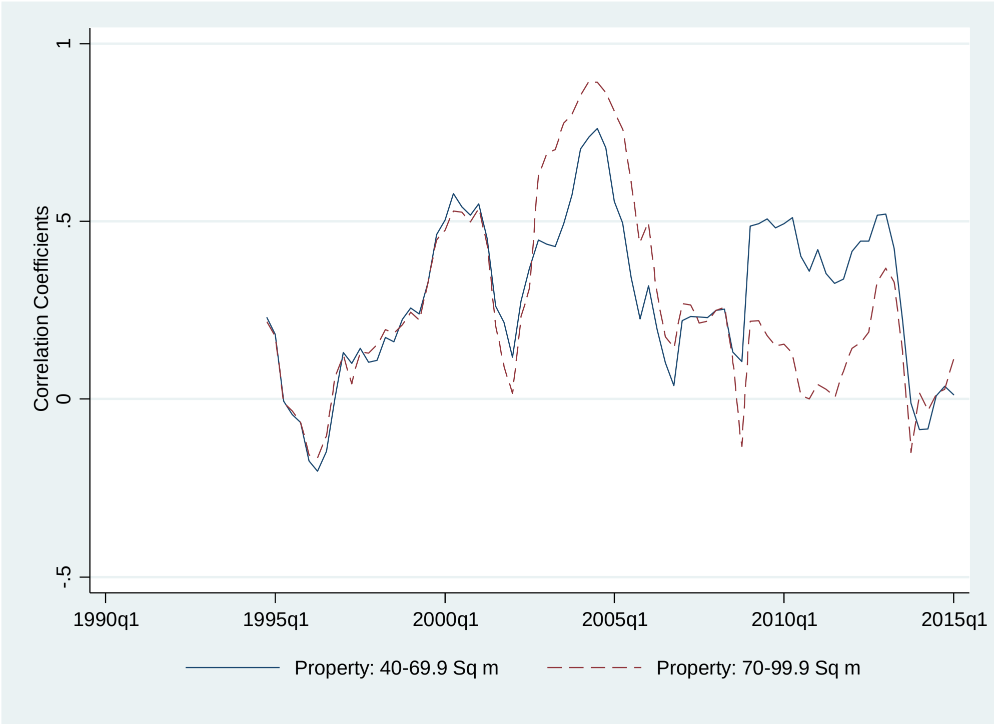


Figure H2: Rolling Window Correlation Coefficients of Real Wage and Property Yields

(Rolling Window: 20 Quarters)



Note:

While the unemployment rate data is in monthly frequency, the wage data is in quarterly frequency. We therefore shrink the rolling window to 20 periods (i.e. 20 quarters) instead of 24. Otherwise, the number of sub-samples will be too small for any meaningful inference.

Moreover, the progress of Hong Kong labor market research in terms of search-theoretic approach may be limited. For instance, based on Krause and Lubik (2007), Krause et al. (2008a, b), Lubik (2009), among others, Lubik (2012) estimates a new Keynesian model with frictional labor market with Hong Kong data. It is found that some parameters *may not be identified very well*. For instance, the range of the posterior estimate of the bargaining power is not very precise (please see Table 3 of Lubik, 2012). Given such imprecision, it may not even be practical to assess whether the interaction between the labor market and housing market is statistically significant.

Clearly, all these findings do not rule out the possibility that there are economically and statistically significant interactions between the labor market and housing market. They do suggest that the relationship may be more complicated than we thought and can only reserve that to future research.