The Subsidized Hong Kong Property Market: Public Utilities, Proximity Effects, Price Indices and Policy Impact

Yifan (Victor) Ye

Duke University

Honors thesis submitted in partial fulfillment of the requirements for Graduation with Distinction in Economics in Trinity College of Duke University; Duke University, Durham, North Carolina; 2015

Faculty Advisor: Charles M. Becker

Abstract

This paper presents a geospatial analysis of the government-subsidized, semi-commercial sector of the Hong Kong real estate market. Using a time dummy hedonic regression model, size, age, seasonal, floor-level, coastline-distance and commuting effects are investigated. Significant elevation gradient and new-station anticipatory effects are found for apartment proximity to the metro system. The paper also finds evidence of differentiated density spillover influences from various HKSAR housing sectors, positive for commercial housing and negative for the public rental unit market. In addition, a hedonic price index is constructed for the subsidized market. Two examples are included to demonstrate the index's policy-side applications.

JEL: C21, N95, R20, R31, R38, R41, Z18

Introduction[1,2]

From an economics standpoint, property markets are perhaps unique in their dual nature. The need for shelter at a minimum, satiated either through renting or purchase of property, is nearly comparable in its inflexibility to that of food or water. Supported by young adults moving out of their parents' residences, ongoing housing demand is no doubt robust to some degree in all societies. However, the value of a piece of property is also highly influenced by macroeconomic conditions and specific attributes related to its location and design. Hence, houses and apartments are attractive as investment vehicles with price volatility as well as potentially large demand-supply imbalances caused by surges in speculation.

This double identity makes housing markets interesting subjects of economic research. By separating the characteristic-driven part of housing price fluctuations from the speculation-driven, statistical tools can be applied to examine not only how individuals make decisions to purchase property but also how shifts in property markets affect the larger economy. From a finance perspective, speculation-related housing price trends can be used to predict behavior in investment markets. For policy-makers, understanding consumer preferences for houses and the impact of government presence in property markets leads to better regulatory practices.

Property price indices can be set up to evaluate changes in such factors over time, providing meaningful guidance to social-economic institutions and private investors alike. However, considering the slow turnover rate of property and relatively small number of transactions involved, it is difficult to reach significant conclusions without data sampled from an extended timeframe that includes a large share of transactions. To account for such disparities, economic models that can adjust for both time-related and characteristic-related price variations between individual transactions must be created. Such models, in their ability to selectively reflect the influence of certain factors and not that of others, are ideal candidates for the investigation of issues such as long-term preference changes and policy shifts across time.

The hedonic regression model is a well-established method in consumer market research, generally used to separate quality-dependent price changes from market-oriented fluctuations. In housing market research, the method is often used to determine quality factors of a house that influence consumer decision-making and to analyze the extent to which such factors contribute to property prices. Despite the evident difficulty of generating regression-based data in real-time, quality-adjusted prices are often used to create housing price indices. However, little research has

[1] The idea for this paper came from a research report I completed for the Hong Kong Monetary Authority (HKMA). Following several years of rapid price increase in property markets, the HOS program was officially rebooted in 2013 in an effort to provide affordable residencies. However, at the time there were concerns about its long-term sustainability and influence on the commercial real estate sector, and HKMA economists decided that an in-depth look at the subsidized market was needed. In this thesis I have greatly expanded upon the results of the original report and introduced new objectives in the hope of achieving a more comprehensive analysis of the subsidized housing market.

[2] This project could not have been completed without the help of my mentors and colleagues at Duke University and in Hong Kong, to whom I am greatly indebted. I would like to express my deepest gratitude to Prof. Charles M. Becker of the Duke University Department of Economics for his insightful advice and suggestions during the completion of this thesis. I would also like to thank Dr. Raymond S. Yuen, my supervisor and mentor at the HKMA and professors Michelle Connolly and Tracy Falba, instructors of the Duke Economics thesis workshop courses. A special thanks goes to Miss Xinshu Sui, who has wholeheartedly supported my studies and been an invaluable companion ever since I arrived at Duke University.

been conducted on specific subsets of the real estate market such as subsidized housing purchases. The reason for this omission is obvious: although many, if not most, city governments provide some variation of rent subsidies to low-income communities, very few opt to subsidize the purchasing of property.

In this regard, Hong Kong stands out as a peculiar example with its extensive government involvement in property markets in the form of purchase subsidy programs. With capital inflows from mainland China, tight land policy restrictions, and a rapidly growing population, Hong Kong's commercial real estate prices are notoriously high. Even during years with deflationary pressure in the market, commercial residences are typically far beyond the financial means of ordinary individuals. In order to provide eligible citizens with a chance to own property, the Hong Kong Special Autonomous Region (HKSAR) government has created housing purchase subsidy programs such as the Home Ownership Scheme (HOS) and Tenant Providence Scheme (TPS). As a result, a significant segment of the domestic, purchase-based housing sector, up to 25% by market share, involves substantial government participation and may behave distinctly compared to the commercial market. Important as this segment is, it has thus far eluded in-depth, quantitative research. As of this writing, available price indices for Hong Kong property markets do not take into account any government-subsidized transactions.

I would like to fill in this gap. Based on data collected from a number of institutions including the Hong Kong Housing Authority (HKHA), the Hong Kong Ratings and Valuation Department (HKRVD) and the Centaline Property Agency, this paper examines in detail potential factors influencing the value of subsidized apartments in Hong Kong, particularly those related to their quality and geographic properties. I shall also propose an alternative, quality-adjusted housing price index used to illustrate time-related trends in the subsidized housing market. The first section introduces the history and current conditions of the subsidized Hong Kong housing market. The second section includes a literature review as well as theoretical and empirical frameworks for the model. The third section summarizes datasets and data methods used in the paper. The fourth section analyzes these statistics in detail in an effort to interpret and understand the market's properties and trends, the results of which are concluded in section five.

Section I:

1.1 Housing Subsidy Programs in Hong Kong

One of the common issues associated with island nations and city-states in general, housing shortage is in many ways a chronic problem for Hong Kong. The region rapidly developed as a manufacturing hub during the 1960s, seeing substantial growth in population, income levels, living standards and land use. Housing demand, particularly demand for housing suitable for middle-class families, rapidly increased as large segments of the Hong Kong population found new, stable sources of income. However, by 1960 the housing market in Hong Kong largely remained as it was in the early 1940s: a small number of high-end, private estates coupled with extensive rent subsidy programs, most of which focused on satiating the need for minimal shelter. (Smart, 2006) Apartments provided by these early programs, such as the Low Cost Housing Scheme (LCHS) and Hong Kong Model Housing Society (HKMHS) were typically small and cramped, often with limited access to water and electricity(Huang, 1999).[3]

Government officials decided that circumstances called for greater state involvement in the housing market. A third, "semi-commercial" housing market would be created to bridge the gap between the high-end commercial market and rent subsidy programs designed for low-income families(Liu, 2001). The term "Home Ownership Scheme" (HOS) was coined in 1970, at first conceived as a small project to supplement existing rent subsidy programs. The program quickly became a stand-alone government-led housing initiative. Specifically, HOS was designed to accommodate the so-called "sandwich class" – citizens who desired to own property yet could not afford commercial estates. Designers of HOS also argued at the time that encouraging higher-income rental unit tenants to move out and purchase apartments at lower-than-market rates could alleviate pressure on public rent units, which were generally in high demand(Er & Li, 2008).

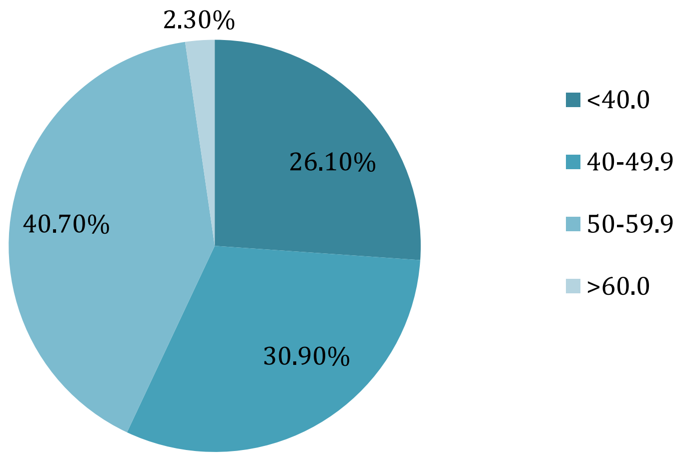
Funded by a HK$1.39 billion pledge from the Hong Kong Legislative Council (LegCo), the first HOS apartment groups, or "courts," were completed between 1978 and 1982. The program, initially managed by the HOS Public Fund and later directly run by the HKHA, would soon become self-sustaining, reporting annual profits for over 15 consecutive years(Liu, 2003). True to the Hong Kong government's hands-off approach to markets, HOS was designed from the beginning to be highly privatized in nature – courts were built and, to this day, listed for sale by private real estate developers. The HKHA's role in the program is highly limited, mostly involving the regulation of secondary markets, managing of HOS funding and establishing of eligibility criteria. Similar programs such as the Tenant Providence Scheme (TPS), Flat For-sale Scheme (FFS) and Sandwich Class Housing Scheme (SCHS) would appear in the following years, but HOS remains to be the largest housing subsidy program in Hong Kong. As of the program's first cancellation in 2003, 219 courts, with a total of some 314,000 apartments, were built under HOS.[4,5]

[3] As of 2013, the average per person living area for public rent unit residents is 12.9 m^2. (Hong Kong Housing Authority, 2014) The averages for 1999 and 2004 are 10.4 m^2 and 11.5 m^2, respectively. (Hong Kong Legislative Council, 2005) The actual per person area as of 1960 is most likely much smaller than that of 1999 and the conditions conceivably worse.

[4] A court is a group of apartment building sharing a name and street address.

[5] As of 2002 the total housing stock in Hong Kong is some 2,224,000 apartments, of which 14.1% are units built under the HOS title. Source: HKRVD

Fig.1 Size distribution (m^2) of existing HOS apartments[6]



In 2003, the HOS program was "permanently" terminated in response to the housing market slump that followed the Asian financial crisis of 1997. Secondary market transactions continued after the termination, but no new land was allocated to HOS and other housing subsidy projects (Zhao, 2005). Policy-makers at the time argued that across-the-board deregulation and dis-involvement of the housing market would create inflationary pressure on commercial estate prices, providing aid to homeowners struggling with loan payments amid low property values. This approach, along with other deregulatory initiatives with similar goals, appeared to be ineffective, as housing prices saw further decline after 2002. The program was eventually restarted in late 2013 in response to rapid housing price growth between 2009 and 2012. However, as of this writing all new HOS projects are still either in the planning phase or under construction, and hence at present there is no primary market for HOS apartments in Hong Kong.[7]

Only so-called "eligible citizens" are allowed to purchase apartments sold under HOS in the primary market. Requirements include family size, income and the absence of ownership of other property. These requirements have been revised multiple times and, as a general trend, gradually loosened. The process itself is actually seen as an integral element driving price fluctuations in both primary and secondary HOS markets (Liu, 2003), since the removal of rigid purchase criteria greatly increases potential demand for such estates. Eligible citizens are separated into two classes, commonly referred to as "green-form" and "white-form" individuals. The former type of eligibility is usually only granted to those previously residing in government subsidized rental units, while the latter can be given to any individual that meets a number of requirements.[8] Preferential treatment is given to "green-form" individuals, including lower down payments and interest rates.[9] Of primary market apartment supply, 60% is always allocated to "green-forms" and 40% to "white forms," despite that white-form applicants are typically greater in number.

[6] Source: *25 Years of HOS: Changes and Developments*, Guoyu Liu, 2003

[7] The first new HOS apartments will be completed in 2015. Source: HKHA

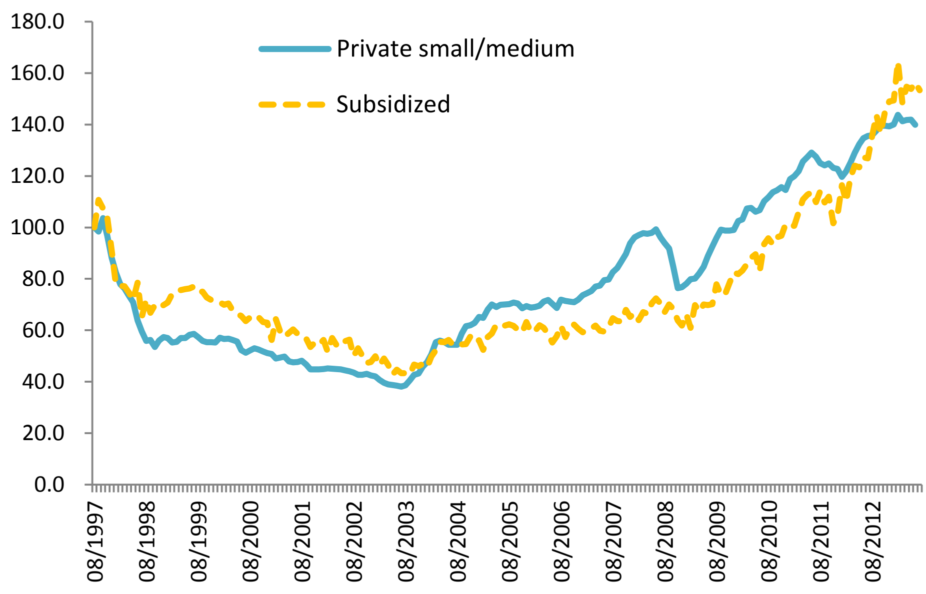
[8] Source: *Eligibility of applicants to the HOS program*. http://www.housingauthority.gov.hk/sc/home-ownership/surplus-hos-flats/eligibility/

[9] Minimal down payment for green form individuals is 5%, and for white form individuals 10%.

Those who are eligible could, depending on the demand of the specific court in question, either directly enter into a purchase contract or submit their name to a lottery with a small pledge(Er & Li, 2008). A specific, prior-quoted discount would then be placed on the purchasing price of the apartment in question. In other words, primary market apartments are quoted at a "commercial" rate determined by the HKHA, from which the subsidized price is accordingly discounted. Note that this does not mean that buyers of HOS are not eligible for housing subsidies of alternative forms. All individuals who are allowed to purchase HOS apartments are also automatically granted access to low-interest, long-term mortgage contracts provided by the HKHA.

True to its market-oriented nature, HOS apartments can, after a certain amount of time since the original purchase date, be freely traded either between eligible applicants or on an "open market" as commercial units. An HOS secondary market was planned in 1996 following public demand for the ability to resell HOS apartments, and the first transactions occurred on August 1997. Transactions between eligible individuals are appraised at a mutually accepted price and made with the original discount applied as a "continuation" of the subsidy. Transactions on the open market happen at market prices, but a percentage of the total transaction amount equal to the original subsidy on the apartment is paid back to the HKHA as a "refund" of the original subsidized amount. This practice not only ensures that the initial subsidy can only be enjoyed by those who meet the criteria for HOS, but also provides the HKHA with a steady revenue stream.

Fig.2 Price indices, commercial market for small and medium apartments/ HOS secondary market, 1997-2013 (Aug 1997 = 100)[10]

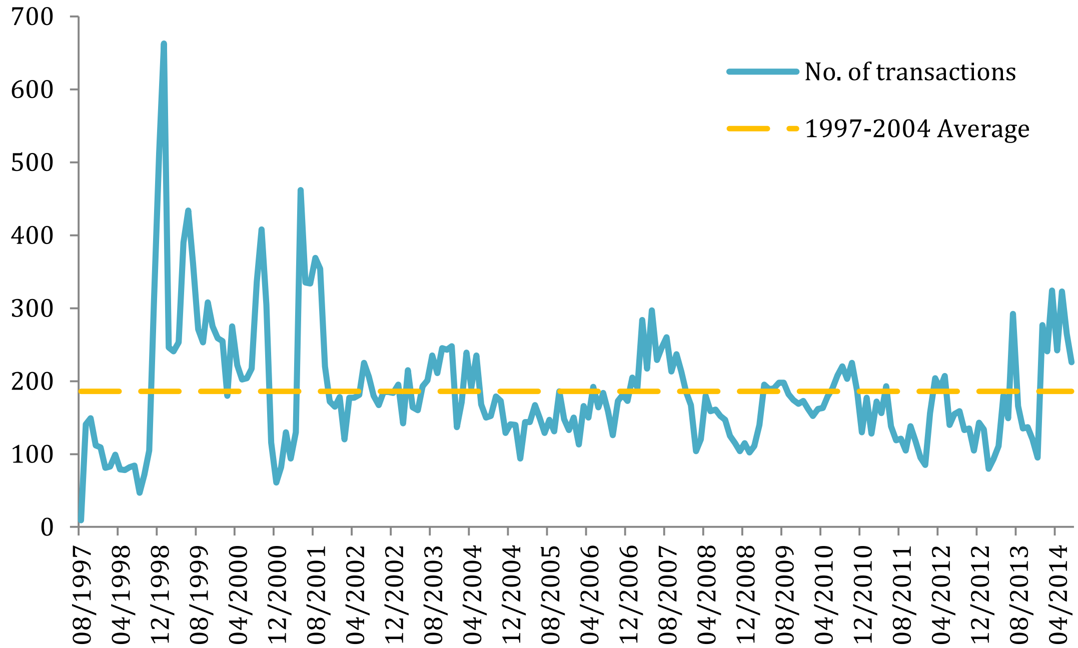


Since its inception, the HOS secondary market has been subjected to a wide range of criticism. The nature of the secondary market, especially the open-market transaction system that caters to individuals that would otherwise not be eligible to own HOS apartments, puts it in direct competition with the commercial market for small and medium sized apartments. Fig.2 shows that the price index for the HOS secondary market, while displaying less volatility, tracks the general trends of the Hong Kong Bureau of Statistics (HKBS) index for small and medium apartment

[10] Private market price index data is obtained from the Hong Kong Ratings and Valuation Department (HKRVD).

sales.[11] Because of the low profit margins of HOS apartments, they are usually listed on the secondary market at prices lower than similar property pieces listed by real estate dealers, causing deflationary pressure in the commercial housing sector. The market itself is managed directly by the HKHA and does not rely on private sector intermediaries, a practice seen as a sign of the government competing with the private sector. Despite often being used as an example of government overreach in Hong Kong, the secondary market has remained robust, with more than 2,200 transactions per year on average between 1998 and 2013 (Fig.3).

Fig.3 No. of transactions per month, HOS Secondary Market, Aug 1997-Jul 2004[13]



Even though discounts are typically large and can in uncommon cases reach 60%, HOS apartments are still expensive items for the income level of the typical eligible applicants: loan contracts can be as long as 25 years, with monthly payments as high as 40% of net family income(Liu, 2003).[14,15] Under most circumstances loans are not government-backed, the exception being guarantees on some transactions that the HKHA will buy back the apartment at the original price within the first two years of purchase. Secondary market listing of HOS apartments is also prohibited until two years since the original date of contract, with further restrictions on the buyer's identity until five years after. Only after such a period can apartments be freely listed on the secondary market and sold to any types of buyer, whether eligible or not. Hence, entering into such contracts is probably a difficult decision to make for many HOS applicants.

11 Small/medium here is defined as Property classes A, B and C, which includes all apartments smaller than 100 m² in gross floor area.
13 520 transactions occurred between August and December 1997, and 1898 transactions occurred in 2014 as of July. Source: HKHA
14 Mean discount rate for all sample transactions is 41.5% and the median is 43%. The maximum discount rate in the dataset is 67% and the minimum discount rate is 6%
15 Income cap for eligible applicants is 40,000 $HKD/month for families and 20,000 $HKD/month for single individuals, as of September 2014.

The impact and size of housing purchase subsidy programs such as HOS and TPS make them stand out among similar government initiatives. China's Economic and Comfortable Housing Program (ECH), established in 1994, was heavily influenced by the success of these Hong Kong programs. As of 2010, however, only 3.8% of all housing starts by area in Mainland China come from subsidized purchasing programs such as ECH (Zou, 2013). In contrast, over 15% of all housing units in Hong Kong fall under some kind of purchase subsidy scheme (Fig.4) as of the cessation of new HOS court constructions in 2002(Liu, 2003). Singapore's Housing and Development Board (HDB) projects may be vastly larger in scope, but their nature is so far removed from the closely market-oriented subsidy schemes of Hong Kong that it is difficult to establish any sort of comparison between housing policies of the two city-states.[16]

Fig.4 Hong Kong housing market composition, 1991 and 2001 compared[17]

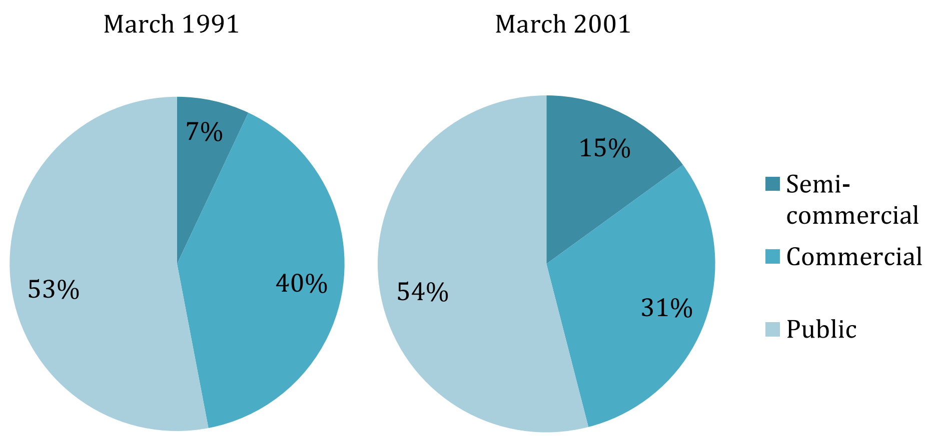
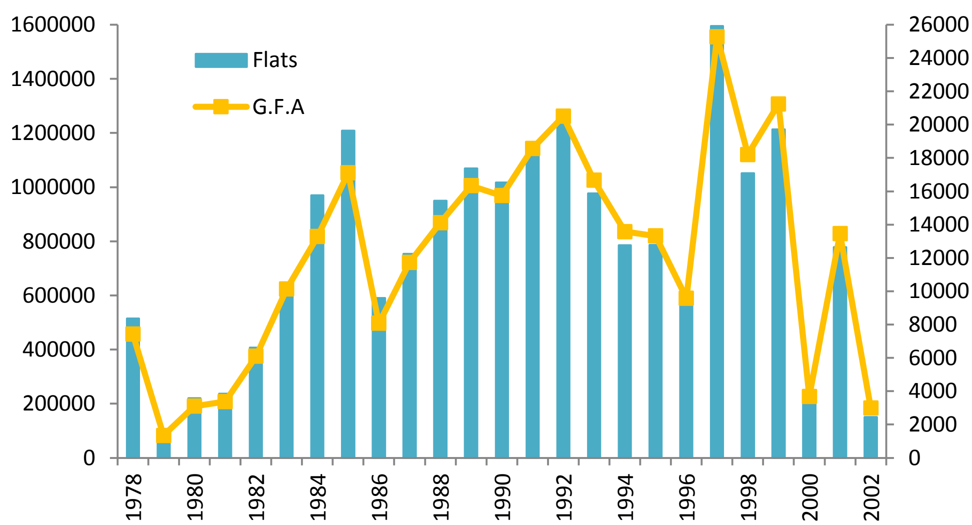


Fig.5 New apartment units/new supply by area (m^2), HOS market, 1978-2002[18]



[16] As of 2014, 81.9% of all residencies in Singapore fall under HDB public housing programs. Source: Singapore Department of Statistics, http://www.singstat.gov.sg/statistics/latest_data.html#20
[17] Data Source: *Property Market Statistics: 1992*, Hong Kong Rating and Valuations Department. *HKHA Performance and Statistics: 2002*, HKHA
[18] G.F.A denotes Gross Floor Area. Source: HKHA

Despite being smaller in absolute scope compared with both the commercial property market and the public rental housing market, the semi-commercial, subsidized purchase market is nonetheless an integral part of the Hong Kong real estate system. It represents one of the most significant government-led efforts to provide accessible ownership of property to middle-income, working class families in densely populated, land supply-constrained regions. There is little doubt that many Hong Kong citizens have greatly benefitted from subsidies provided by programs such as HOS and TPS, yet claims that these programs cause market distortion and deflationary pressure on the commercial housing sector may also hold more than a grain of truth. As the HOS program reboot unfolds following an all-time property price peak in Hong Kong, it is useful to take a closer look at the dynamics and influence of such a market.

In addition, high-quality data available for the HOS secondary market can be employed to study urban housing market characteristics in general. There is little reason to believe that individual preferences for attributes of apartments, as revealed by the hedonic methods in this paper, will differ for this particular market because it is subsidized. Truly low-income communities, such as those that occupy government-provided rental units, may not share quite the same tastes and patterns of demand. However, HOS residents are at least wealthy enough to enter into contracts for home ownership, and are therefore much more likely to have reasonable views of the value of amenities and apartment characteristics. Intuitively, individuals would want to live in bigger homes, apartments with better views that are closer to subway stations and shopping centers regardless of discounts applied to their purchase of property. The broader appeal of this paper lies in the ability to relate findings in the HOS market to other metropolitan housing markets, subsidized and unsubsidized alike.

Section II:

2.1 Literature Review

Much of the basic methodology of this paper is similar to the approach proposed in the Australian Central Statistics research report *Exploring Hedonic Methods for Constructing a House Price Index*, which explores options for conducting housing price index-oriented hedonic regressions. In that paper it is noted that the long purchase-resale cycle makes it essentially impossible to carry out conventional match sample adjustment methods in constructing a housing price index. Such methods pair up an item's transaction in a certain period with transactions of the same item in other periods(Chen, Zhao, Romanis, & Lim, 2004). Considering the inherent difficulty of controlling for all or most characteristics of a house that may influence its final sale price, the paper concludes that regression-based hedonic methods are a better option for housing markets. [19] While hedonic methods are complex and require relatively large amounts of computation, they are in many respects preferable to a match sample adjustment model in terms of accurately accounting for characteristics of housing markets.

Chen and Zhao (2004) outline a general procedure for calculating unadjusted price indices by regions in Australia, involving (1)The removing of extreme values (2)Dividing the dataset by price level, (3)Calculating the un-weighted average prices for each level, and (4) Deriving a weighted average price index by assigning different values to different sub-markets organized by price(Chen et al., 2004). A similar method is used with commercial-market housing indices in Hong Kong. However, the official, government-established Hong Kong housing price indices are separated by flat size instead of absolute price(Hong Kong Ratings & Valuations Department, 2013). The unofficial but widely adapted Centa-city leading index is a comprehensive unadjusted and un-weighted housing index that includes all commercial transactions made by the Centaline Property Agency.[20]

With regard to the choice of hedonic models for housing, Silver and Hervai (2006) demonstrate that time dummy hedonics is an acceptable method of quality adjustment for relatively stable parameters and characteristics sets with little variation over time. Since the use of time dummy hedonics assumes that the set of characteristics influencing the product remains fairly fixed throughout the entire period, rapidly fluctuating parameters or numerous new characteristics can greatly decrease the ability of a model to analyze the influence of a single characteristic. The time dummy hedonic adjustment method is therefore sufficient for property market research as long as parameters contributing to a house's retail price remain relatively stable within the time frame of the regression model.

Haan (2003) and Pakes (2003) further develop this analysis in two purely statistical papers, arguing that the time dummy method should be considered as a special case of the more complex

---

[19] Hedonic regressions can be used to derive the relative impact of each individual characteristic of a piece of property and adjust for the differences between property units. Each transaction essentially becomes adjusted to the price at which a "representative house" would be sold at the same time and circumstances. If the adjustment is perfect, then there should only be speculation or market-related price changes in the long run. However, hedonic model differ greatly in flexibility and complexity.

[20] The Centaline Property Agency (Ltd) is the largest commercial property dealer in Hong Kong. It also offers negotiation services for secondary market subsidized property sales and, when available, listing services for the primary market.

hedonic imputation method, where regression weights are stable and properly chosen. Problems appear, according to Pakes, when the time dummy method is applied to markets with rapidly fluctuating qualities such as consumer electronics. Note that this implies that it is not possible for a standard time dummy hedonic approach to disentangle either stable, long-term quality improvements or individuals' preference changes from changes in the price level. However, there is no reason to believe that the method cannot be safely applied to housing markets, especially for subsidized purchase markets such as HOS where the overall level quality is essentially fixed.

Diewert (2003) develops the theoretic regression models built on Silver and Hervai's work. In particular, Diewert suggests that fully linear hedonic regressions, by simply smoothing out market fluctuations that quality-adjusted price indices ought to reflect, are generally unjustifiable and should be avoided. Traditional match model techniques can in fact be as effective as a hedonic regression, but only in theory and with a sufficiently large number of matches. Also, quantity data such as gross sale amount should be incorporated in hedonic regressions if at all possible. It is concluded that as there is no general consensus on a best practice for employing hedonic regressions, flexible functional forms should be taken into consideration to better approximate changing tastes.

However, there is also criticism towards the general concept of using hedonic regression models to derive quality-adjusted price indices. In an OECD working paper on the hedonic regression methodology, Hill (2011) proposes four potential pitfalls of using hedonic regressions: omitted variable bias, functional form misspecification, lack of transparency, and sample selection bias. Since hedonic regressions are aimed at deriving an essentially "clean" adjustment of qualities, removing the impact of some variables and not that of others can cause the result to be biased. This issue is especially problematic when attempting to analyze variables that are not directly measurable, such as the impact of noise on housing. A larger dataset can potentially alleviate the omitted variable bias problem and significantly improve sample selection quality. However, the second and third pitfalls are not as easily addressed (Malpezzi, 2003). The best functional form of a hypothetical hedonic regression model is impossible to determine, a fact that can be proven by mathematical deduction (Rosen, 1974). However, one can use established transformations such as the Taylor series to reach decent approximations.

It is also true that all hedonic regressions require active choices made by an "index provider" – two researchers given the same dataset will almost certainly come up with different hedonic models (Shiller, 2008). As a result, a hedonic index can never be as transparent or accessible as a direct-weight or matched price index, the creating of which involves no subjective decision. Omitted variable bias may also be especially problematic for housing market indices, since property prices are influenced by a larger group of variables than the typical consumer market. Compared to automobiles or electronics that generally have fewer distinct, quantifiable measurements of quality, the price of a house is determined by both properties of the house itself, such as number of rooms and size, and characteristics of the neighborhood, such as crime and schooling. The effects of these quality factors may neither be predictable nor easily measured.

Melser (2005) identifies another issue with the time dummy hedonic regression approach. According to Melser, hedonic regressions do not satisfying certain mathematical axioms from index number theory. These axioms are crucial to the assumption that the result of hedonic regressions capture both the market assessment of characteristic changes and inflationary or "pure" price changes approximated by the adjusted price indices. (Melser, 2005) More specifically, a time dummy hedonic regression process may not satisfy so-called monotonicity axioms, i.e. that price indices should move in the same direction as the price fluctuations of the second period and the opposite direction as that of the first period in a two-period index, in both cases holding the other period constant. It can be demonstrated that with specific input sets, time dummy regressions will not obey this axiom and derive a lower index number for a price increase from the first period to the second. With regard to the hedonic index figures presented in my paper, the asymmetric percentage difference method is used to ensure that price increases and decreases can be equally represented by the same percentage amount.[21]

There is extensive research involving the application of hedonic models to housing markets. As early as 1977, Goodman (1978) applies time dummy hedonics to data collected from the New Haven urban area and confirms that area-specific hedonic regression can be used to reveal nuanced price structure differences in sub-markets usually obscured by general assumptions about market size and composition. It is therefore advisable to include area-specific variables if the goal is to derive quality-adjusted price levels for an aggregate region. Cebula (2009) conducts regressions with three different models and 24 variables on transaction records from Savannah, Georgia. It is concludes that, holding other qualities constant, houses designated as national historic monuments or located within the Historic Landmark District carry a price premium. He also notes that the use of spatial control variables leads to the observation that there may be seasonal demand and supply shifts that influence housing prices. Quality-adjusted housing indices have also been used to determine demand for clean air by urban residents (Harrison & Rubinfeld, 1978), effects of location-specific characteristics such as commute time to the Central Business District area (Ottensmann, Payton, & Man, 2008), the influence of airport expansions on property values (McMillen, 2004), quality factors for the housing market of Hangzhou, China (Haizhen, Shenghua, & Xiaoyu, 2005) and the effect of school availability on homebuyers (Hayes & Taylor, 1996).

An in-depth study of the impact of subsidized housing can be found in *Housing Market Spillovers: Evidence from the End of Rent Control in Cambridge, Massachusetts* by Autor, Palmer and Pathak (2014). The authors measure negative externalities from rent control programs in Boston by examining their termination and conclude that substantial economic benefits, accounting for a quarter of the total period price appreciation, were generated by the cancellation of such programs. Most of the benefits are not caused by price appreciation of rent-controlled apartments but are in fact from positive spillover effects to residencies that have never been subjected to rent control. This implies that housing subsidies, at least for rental markets, can have large negative "proximity effects" that, by lowering the prices of surrounding residencies, cause overall economic losses. The extent to which such effects are present, if at all, for purchase-subsidy markets is one of the central questions my paper hopes to address.

Also closely related to the theme of this paper, Chow (2011) uses time-panel data purchased from the Centaline Property Agency to construct a hedonic model for the commercial Hong Kong housing sector. As data was limited in both time and scope, the aim of the paper was only to identify potentially significant quality variables for the Hong Kong housing market. Chow concludes that there are a significant positive relationships between apartment price and factors including floor area, absolute height in stories and school network strength. A significant negative relationship is found between apartment selling price and distance to the nearest Metro station. The relationships for age and number of bedrooms are both negative but not statistically significant. These observations illustrate a starting point in the search for potential regression variables for my

[21] The asymmetric percentage difference method adjusts individual price indices figures using an equation that equalizes inflationary and deflationary periods. See theoretical framework section for details.

paper. Despite the limitations of available data, Chow notes that short-term price fluctuations reflected in the unadjusted index are smoothed out when the price index becomes adjusted. In other words, a significant amount of price movement in the currently used price indices is in fact noise generated by disparities of apartment qualities for transactions of different periods. One of the key points of my paper is the further development of the quality-adjustment process.

2.2 Theoretical Framework:

The main attraction of the time dummy hedonic regression method is its ability to separate the price influence of quality-oriented, sample-dependent variables from price changes originating from shifts in the aggregate market. A standard linear-linear regression for a single time period t is represented by the equation:

$$(1) \qquad P_{it} = \delta_0 + D_{it}\delta_v + \sum_{j=1}^{k} B_j X_{ijt} + \varepsilon_{it}$$

P_{it} is the transaction price of item *i*, and δ_0 is a constant intercept term for the base period time effect. The term D_{it} is the time dummy vector, which equals 0 or 1 depending on whether the observation is in time period *t*. δ_v is the time effect of being in period *t*, X_{ijt} a vector of quality variables in the regression input, and B_j the corresponding vector of quality coefficients for observation *i*.

The regression therefore produces two results: a vector of coefficients that describes how different quality variables affect the price of a given item and a constant term, which approximates the base-period time effect. To generate an adjusted price level for period *t*, the expected value of δ_v at *t* or expected time effect of period *t* is derived with the equation:

$$(2) \qquad \exp(\hat{\delta}_v) = \frac{\prod_{i \in S_t} (P_{it})^{1/n_t}}{\prod_{i \in S_1} (P_{i1})^{1/n_1}} \exp\left[\sum_{j=1}^{k} \hat{B}_j (\bar{X}_{j1} - \bar{X}_{jt})\right]$$

The left part of the right side of the equation is $(\prod_{i \in S_t} (P_{it})^{1/n_t}) / (\prod_{i \in S_1} (P_{i1})^{1/n_1})$ simply a ratio of geometric mean sample prices between period *t* and period 1. This ratio is then adjusted with the right part, which compares the quality variables of the samples in period *t* with those in period 1 and removes the influence of the discrepancy in quality. Quality difference is therefore adjusted to a standard, average level, and would not interfere with market-level price changes *if* the adjustments are perfect. By performing the regression over periods 1 to *t*, this equation produces a price index with *t* data points, with the first period price being normalized to 1. In terms of describing the housing market, the result can be considered as the price changes of a single, representative apartment over time.

Note that a fully linear regression may not be the best choice for this type of work. A more practical approach would be using a by-period log-log regression or regression that incorporates log-terms instead. Linear regressions may not be worse compared to logarithmic regressions in evaluative strength, but log-log regressions enable the direct comparison of the influence of different characteristic variables, which may not have the same unit or magnitude of size. More importantly, log-log regressions reduce the severity of heteroskedasticity problems in a regression model. Furthermore, for relationships between apartment price and geographical variables such as the driving and walking distance estimates in this paper, percentage changes are intuitive and relatively easier to articulate.

It is obvious that the choice of the size of period t is crucial to the applicability of the time dummy hedonic. By generating more sample indices within a given timespan, shorter period lengths allow for greater resolution in observing temporary market fluctuations but also introduce a larger amount of statistical noise caused by the small individual regression sample pools, weakening the significance of each single estimate. Hence, there is a trade-off between explanatory accuracy of the individual index data points and the power of the index in describing short-term changes of the market in a time dummy regression-generated index with limited data. Running regressions with differently sized time periods therefore could be beneficial if multiple goals are to be achieved with the same dataset.

As a final note, a price index created using a dataset with many time intervals and multiple periods of unidirectional price changes must be adjusted with asymmetric percentage difference. If for period t the price level is P_t and P_{t+1} for period $t+1$, the recorded percentage change between the two periods cannot simply be formulated as P_{t+1}/P_t. To balance out the impact of numerically equal price increases and price decreases, the percentage change C must be approximated as $2 \cdot (P_{t+1} - P_t) / (P_{t+1} + P_t)$, or:

$$C \approx 200 \cdot (P_{t+1} - P_t) / (P_{t+1} + P_t)\,\% \tag{3}$$

Using asymmetric percentage difference, a 25% price increase followed by a 25% price decrease is the same as a 25% decrease followed by a similar increase. In other words, the price decreases between periods are not discounted as in a traditional ratio calculation. For an index that only increases or decreases across all time periods, the P_{t+1}/P_t approach may not cause any problem. However, for a dataset that can potentially map housing price fluctuations across decades, asymmetric percentage difference is an effective tool in accurately accounting for periods of similarly sized price hikes and drops.

In practice, alterations can be made to the standard time dummy hedonic regression model to suit specific needs of research. If there are systemic differences between sections of the same market, separate regressions for sub-markets may be preferable compared to using additional variables to adjust for sub-market differences. A combination of dummy variables and quality variables is often used, with interaction effects between quality variables and measurements for time and geographic locations.

Section III:

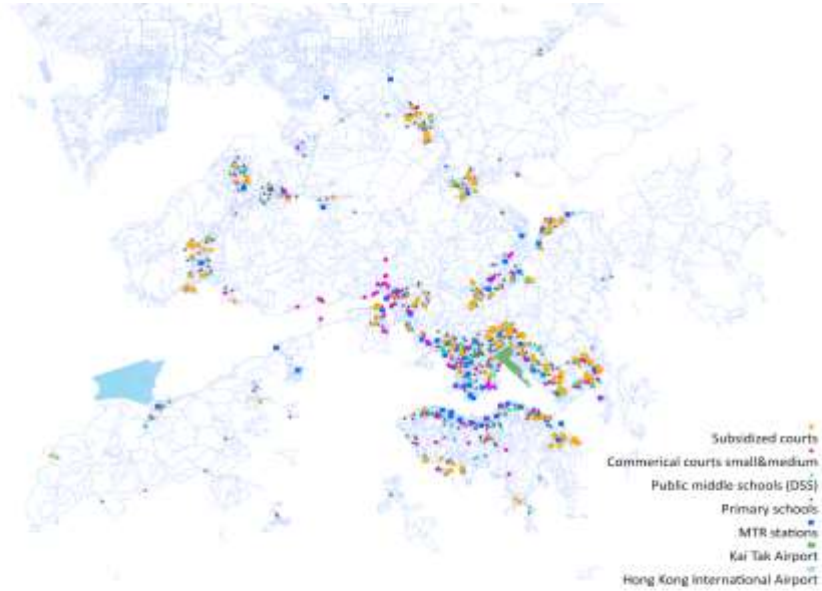
3.1 Empirical Approach

Broadly speaking, the objectives of this paper can be separated into two groups. The first group, which involves the studying of influences of various quality factors of an apartment on its price, appeals to a general audience of urban economics literature. In particular, I shall investigate the possibility of negative spillover effects related to government-subsidized housing and methods to evaluate the influence of metro networks on urban housing markets. The second group addresses questions specific to Hong Kong and is intended as reference for policy-makers in the region. By creating an adjusted hedonic price index, behavior of the subsidized housing market and its relationship to privatized, commercial housing markets of the HKSAR can be analyzed. Since HOS is ultimately a public initiative, one of the focal points of this paper will be the HKHA's role in the subsidized market. To this end, I shall consider past policy changes that have occurred in the HOS secondary market and examine their short and long-term impact on price levels.

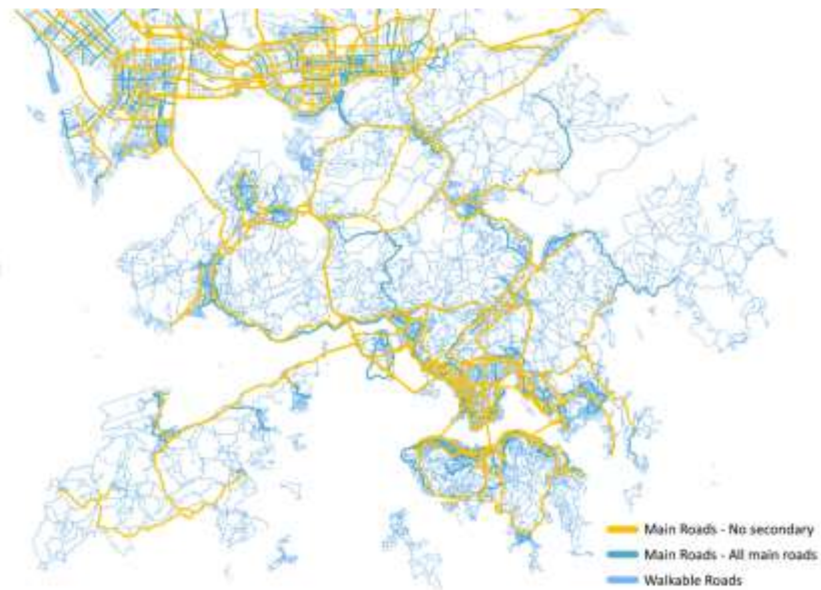
All of these goals critically depend on the ability of the time dummy hedonic regression model to accurately account for quality differences between apartments. The base dataset for this paper comes from the HKHA, which keeps comprehensive records on transactions of flats in the secondary market. While this dataset contains quality descriptions for the apartment involved in each sale such as its size and age, little information about qualities related to the apartment's physical location is included. Yet, these factors are crucial parts of any viable model of urban housing: intuitively, people prefer houses closer to public utilities with better environments. For centralized cities such as Hong Kong, there is most likely also a strong preference for apartments closer to the central business district. For a hedonic regression model to be valid, the influence of these factors on selling price must be adequately accounted for.

Fortunately, there are solutions to this issue. Software packages based on Google Maps API, Bing Maps Representational State Transfer (REST) Services, and ESRI Geographic Information Systems (GIS) are used to create a detailed digital model of the HKSAR, incorporating local amenities, road networks and transit availability conditions. The Google Maps geocoder API provides a method of generating precise geographic coordinates and basic descriptive information for lists of data points according to their names and approximate location. The coordinates are then imported into GIS for the creating of layer files on which a variety of analysis tools can be applied. This approach allows for large-scale, high-accuracy estimations of geographical quality factors of apartments, such as approximations of the walking distance to the nearest metro station or primary school within a 50-meter level of tolerance.

ArcGIS model of Hong Kong with point datasets[23]



ArcGIS road network model of Hong Kong[24]



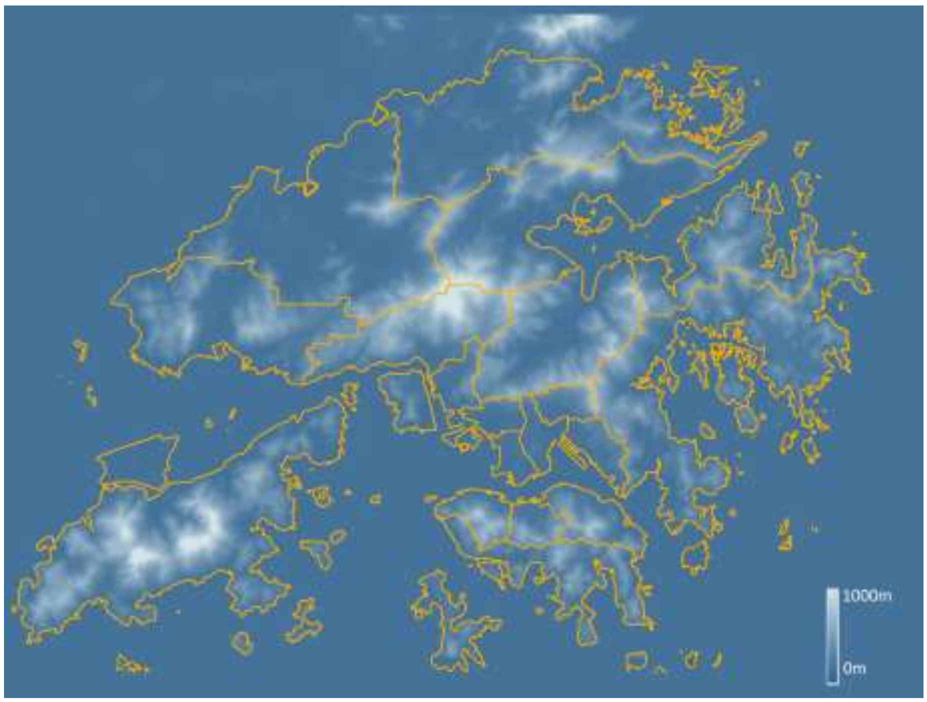
Furthermore, elevation data from a 30-meter-level resolution Digital Elevation Model (DEM) of Hong Kong are used to adjust most of the distance estimates generated with GIS. Although the addition of elevation adds difficulty to the modeling process, I find that there are substantial gains

[23] Data Sources: HKHA, HKBS, Hong Kong Education Bureau (HKEB), mtr.com.hk

[24] The term "Main Roads" denote roads with greater or equal to three lanes in total. "No Secondary" roads are regional highways or roads with greater or equal to four lanes in total. Walkable roads do not take into account roads without specific pedestrian provisions, such as walking under an elevated highway. Data source: The OpenStreetMap project, organized by *MetroExtracts.com*

involved in using terrain data for research on a city with uneven landscapes such as Hong Kong. The level of incline and the walking distance to public utility locations can be independently determined, allowing for separated examinations of effects related to distance and those related to elevation gradients. Knowing the relative altitude of each apartment group can also be useful, since people may have a preference for flats with cleaner air, less noise and better views, all of which are positively correlated to greater elevation. Also, very low elevation is associated with flooding during summer storms, which is highly undesirable for residencies.

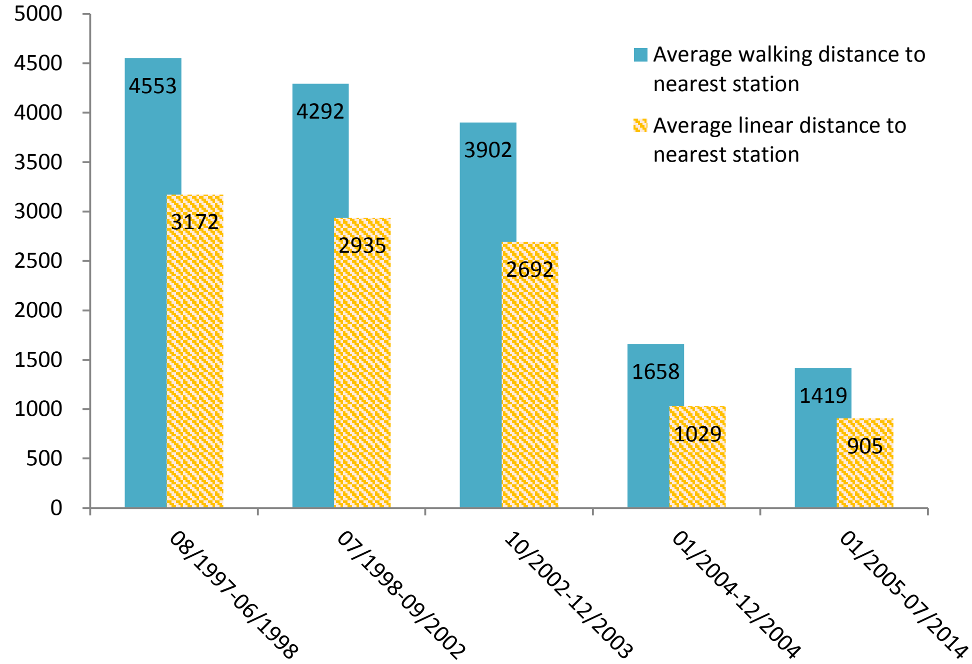
ArcGIS 30-meter resolution DEM for Hong Kong[25]



Considering the relatively long, 18-year time span of the dataset, it is also of great importance to properly adjust for time-related changes of the estimates derived from the GIS model. It is, for example, unwise to use a single metro network model to approximate the availability of subways for the entire dataset because of ongoing metro line constructions between 1997 and 2014. Datasets for schools and road networks are also only accurate for the description of recent conditions, and could be biased when used to adjust transactions that occurred in the past. In the GIS model I have created five different metro network layers to reflect the development of the Hong Kong MTR, with each layer including one new subway line completed after Aug 1997. This guarantees that metro accessibility of the apartment in each transaction is evaluated for the specific metro network conditions at the time period that the transaction occurred.

[25] Coverage of the digital elevation model is limited to the Hong Kong Peninsula, Lantau and Hong Kong Island for cost reasons. Data source: Intermap Technologies, formatted and supplied by *MapMart.com*

Fig.6 Average period walking/linear distance from HOS apartments to the Hong Kong MTR network (meters), Aug 1997-Jul 2014[26]



However, it is somewhat impractical to employ similar methods to school and road network data because of a lack of information about their historical condition. Even if it is possible to know which highways were built during a certain year, road network influences on apartment prices are gradual and difficult to estimate because, unlike metro lines, highways are typically built in small segments and opened to the public one segment at a time. For such variables, I shall include their interaction effects with time indicators in the regression. Through these interaction terms, the influence of the current dataset is allowed to vary for observations in different periods. This option accounts for time-related changes in price effects, albeit to a less extensive degree compared to the modeling-based approach used for the metro data.

The most identifiable potential issue of the regression model used in this paper is a lack of crime data. While crime data exists for the HKSAR region, figures are only released to the public at the multi-district aggregate level. While this might be a major problem for a study on cities in the US, housing prices in Hong Kong are most likely not substantially influenced by local crime rates. While crime events are not rare in the region in general, Hong Kong has an extremely low violent crime rate as well as burglary and robbery rates that are a fraction of comparable rates of the US.[27,28] Rates of petty crimes, most notably pickpocketing and shoplifting, are not as low in

[26] This distance is derived from a closest facility analysis performed on courts and metro stations based on a polyline layer of walkable roads. Each subway station is buffered at 50m and intersected with pedestrian-friendly roads to approximate station entrances. The 50m is added back in after the analysis as a proxy for the distance from the station entrance to the boarding area.

[27] In all years from 2002 to 2013, Hong Kong has consistently ranked among the top 15-30 safest regions in the world by United Nations Office on Drugs and Crime (UNODC) published murder rates (victims per 100,000 inhabitants), with levels comparable to those of Iceland, Sweden and Singapore. The highest recently reported annual murder rate is only approximately 1.0 per 100,000 (2001, 2002). Source: UNODC Global Study on Homicide. http://www.unodc.org/gsh/

[28] Burglary rates in Hong Kong have remained below 130 incidents per year/100,000 inhabitants and is only approximately 57 incidents per year/100,000 inhabitants in 2013. In comparison, the 2013 US burglary rate is approximately 700 per 100,000 inhabitants. Between 2000 and 2014, robbery rates peaked in 2002 at 52.1 incidents per

relative terms.[29] However, such crimes are not likely to influence housing market purchase decisions in the way that violent crimes do. As far as home property and human life is concerned, the average Hong Kong resident ought to feel much safer than the average US citizen, even more so compared to US urban residents. The inclusion of crime data is no doubt beneficial and desirable, but its omission most likely does not substantially weaken the model's explanatory power.

A second, potentially more problematic shortcoming of the model is the lack of window orientation data for apartments. Many subsidized apartment buildings are designed with three to four units per floor. Window orientation is therefore an apartment-specific factor significantly related to selling price. Apartments primarily facing east will sell at a price premium compared to those primarily facing north and south. They will in turn sell at a premium compared to those facing west, which receive almost no direct light until the afternoon. Window orientation data was never collected by the HKHA as part of the quality factor set for secondary market HOS transactions, and is therefore permanently lost. Hence, at least some portion of the total remaining variance unexplainable by the regression model can be explained by different window orientations of the apartments.

However, with a dataset average of 172 transactions per court, it is quite possible for the price variation between transaction apartments caused by window orientation to be largely smoothed out in the whole model.[30] Since apartments must necessarily be located on all sides of a building, with a sufficiently large number of transactions per court the price premiums of apartments on a certain side or corner should cancel out with lower prices of apartments located on the opposite end. If this is indeed the case and, assuming that window orientation is not highly correlated with size, age, floor level or any of the geospatial variables, the coefficients of these variables should not be significantly affected by the lack of window orientation data. The loss of overall explanatory power of the model, measured in remaining variance, is nonetheless unavoidable.

One might also notice the lack of variables concerning the number of bedrooms or bathrooms in apartments. While this is common practice for investigating markets of stand-alone residencies and commercial apartment markets, I believe that there is no need to include these variables for the investigating of the Hong Kong subsidized market. With only 2.3% of all apartments built under HOS being larger than 60m^2, it is highly unlikely than any HOS apartment unit have more than two bedrooms or a single bathroom. Single-bedroom configurations are not suitable for the "family with child" demographic that most HOS apartment residents fall under. While there are no doubt some HOS apartments with only one bedroom – particularly those smaller than 30m^2 – the apartment size variable should be sufficient in controlling for their relative price differences.

With all of the factors mentioned above taken into account, I hope that the modeling approach in this paper is robust enough for a comprehensive hedonic analysis. A total of 37,749 observations over a period of 17 years provide an average of 185 transactions per time period using a by-month method or 555 for a by-season regression. These are decent sizes considering the number of variables involved. There are no biases related to the selection of sample, as the HKHA dataset I am using is not a sample but the complete records of the entire secondary market for subsidized

100,000 inhabitants and was as low as 6.9 in 2013. In the two respective year, the US robbery rate was 146.1 and 112.9 per 100,000 inhabitants. Source: Hong Kong Police Force, data obtained from www.police.gov.hk/ppp_en/09_statistics/, FBI US Crime Report 2013(http://www.fbi.gov/about-us/cjis/ucr/crime-in-the-u.s/2013/crime-in-the-u.s.-2013)

[29] The net theft incident rate is approximately 439.6 per 100,000 inhabitants as of 2013. The shop theft and pickpocketing incident rates are 120.1 and 20.4, respectively. Data source: Hong Kong Police Force (HKPF).

[30] The whole model includes 37,749 observations over 219 different courts.

housing. Biases may still exist because of the nature of the secondary market, but it is not very likely that they are large enough to significantly affect the outcome of the model.[31] Most variables are evaluated at the court level and, with 219 courts in the dataset, have at least 200 degrees of freedom. By removing variance in selling prices across transactions caused by differences in individual apartment qualities and qualities related to locations and distances, the remaining market expectation-oriented variance can be derived. Both the influence levels of qualities and the adjusted market values are examined in detail in the following section.

3.2 Research Objectives

The first goal of this paper is to further of our understanding of housing markets and their response to quality changes in general, with a particular focus on the following issues. First, it has been demonstrated that there are negative spillover effects on housing markets associated with rent subsidy programs in the US. Houses close to areas with rent subsidies lose value compared to houses further away. Flats of the subsidy programs also suffer from their own high relative density(Autor, Palmer, & Pathak, 2014). While the subsidy levels associated with HOS are somewhat comparable to the equivalent subsidy level of the Boston rent control programs, the net purchase and home ownership nature of HOS and its relatively expansive scope may cause it to behave differently in terms of proximity-based effects.[32] I would like to investigate the extent to which such density spillover effects, both from HOS apartment density and that of other Hong Kong housing markets, influence prices in the subsidized market.

Specifically, I would like to evaluate the significance of a compound term which evaluates the relative density of surrounding apartments after adjusting for other qualities of a flat. A significant negative relationship will signal some kind of depreciative price effect associated with living in an area with a relatively large number of apartments from a particular market. On the other hand, if the relationship is not statistically significant, significantly positive or of a substantially smaller size compared to the influence of other quality variables, such negative spillover effects most likely do not exist. Not only are the effects of subsidized apartments on themselves investigated, price effects of the density of commercial housing and the low-end, public rental unit market are also examined and compared with that of the subsidized market.

A second potential question is the relationship between discount rates and the price of apartment transactions. Since for secondary transactions the discount rate serves no real meaning other than informing the buyer of how good a deal the original purchaser of the house received, any positive correlation between the discount rate and the transaction price[34] indicates that the possibility of some sort of irrational decision-making process of buyers is at play. On the other hand, a negative relationship may suggest that the discount rate is signaling other, less desirable aspects of an apartment's quality. This is a fairly simple test and can be addressed by including the discount rate

---

[31] It is nonetheless possible that apartments being sold share undesirable characteristics that led to them being listed on the secondary market. However, the average size of the secondary market apartment listings is almost exactly the same as the average size of all apartments under HOS (49.9 m^2), and there does not seem to be any particular court with a significantly greater amount of transactions for its size. Therefore, it is probably the case that selling on the secondary market is not predominantly caused by unhappy owners but because of demand for better housing or need for liquidity.

[32] The average equivalent subsidy of Boston rent control programs is about 45-50% as estimate by Autor, Palmer and Pathak. The average subsidy level of the HOS secondary market transactions is 41.5% and the median subsidy level 43%.

[34] The reported transaction price is discounted for all observations in the dataset.

of apartments in the regression as a control variable. Of course, the issue can also be used to investigate capital gains of the original apartment owner that originate from the initial subsidies applied to apartments.

The third question I would like to look into is the possibility of seasonal fluctuations of the property market. With relatively long delays between searching and a successful purchase as well as very high prices compared to a typical family's annual income, housing markets should not be subjected to the usual type of seasonal volatility observed in markets such as consumer electronics and clothing. Even if there are demand-driven effects associated with seasonal variations, in theory arbitrage should remove all profitable market impact of such effects. However, established research has found signs of cyclicality in housing markets (Cebula, 2009). The evidence is still far from conclusive, and I would like to examine data on the Hong Kong subsidized market to see if a stronger connection can be established.

With regard to housing policy, I plan to use the dataset to create a quality-adjusted, hedonic price index for the subsidized Hong Kong property market. The secondary market's nature as well as the substantial lag between primary market and secondary market availability largely isolate it from supply-side shocks, i.e. introduction of newly constructed flats. As a result, this index will, under ideal conditions, track and only track changes caused by shifts in consumer expectations, such as those related to newly enacted policies and macroeconomic conditions. In practice, I hope that the vast majority of quality differences between apartments involved in secondary market transactions can be removed by the time dummy regression, resulting in high-quality estimates of market-oriented, demand-side fluctuations.

The generated price index can be compared to established consumer price indices for housing markets in Hong Kong. Candidates include the Centa-city leading index, created in 1997 by transaction data from private housing agency *Centaline*[35], and the standard commercial housing index provided by the Hong Kong Ratings & Valuations Department (HKRVD). The latter includes indices for different tiers of the property market separated by apartment size. While none of these indices incorporate quality-adjustment measures, results obtained from comparing them to the subsidized market index could still be of some interest. Potential questions that can be answered by such an analysis include the leading or lagging relationship between private and subsidized markets, and the degree to which different private markets separated by apartment size follow the trends of the market as revealed with hedonics.

Of course, a quality-adjusted housing index is a highly powerful tool for policy analysis. As a market partially driven by the amount and eligibility criteria of subsidies, the extent to which such forces distort general market trends is no doubt a question of great importance to designers of housing policies. By using a by-month time dummy regression, policy-related price changes in the housing market can be observed. The resolution provided by the regression and modeling process is enough to examine not only the long-term market impact of policy changes but also short-term shocks that may only be visible for months after a new policy's implementation. Can a quantitative relationship be established between the various HOS eligibility criteria and market price levels? Is there a way to predict the size of price shocks caused by policy revisions based on the specific policy and revision in question? These are but two examples of questions that will no doubt be of

[35] This index sets the housing market price level on August 1997 as a base level of 100 and is a direct comparison of the current price level of Centaline transactions with base level prices.

great interest to Hong Kong policy-makers. Hopefully, I shall begin to address issues such as these in the next section of this paper.

3.3 Data Summary

This section details the sources, formatting methods, approximations and potential issues of the variables used in the regression, starting from the base dataset provided by the HKHA. The HKHA records for secondary market transactions include the discount-rate adjusted selling price for each transaction, the name of the court where the flat is located and its district, the time of the transaction accurate to the month, the actual usage size of the flat by ft^2 and m^2, the discount rate attached to the flat and a floor level indicator.[36] Apartments with floor levels below the 14th, between the 15th and 26th and the 27th or above are respectively denoted as low (L), medium (M) and high (H). Dummies are used in the regression model for the medium and high apartments, with low apartments as the base level.

This dataset is combined with available data about subsidized housing projects to obtain a number of other regression variables. Most notably, the age of each flat as of the date of transaction is derived from subtracting the opening date of the court from that date.[37] Age numbers are rounded to the closest year in the regression model. It must be noted that some apartment buildings may have either been vacant for extended periods before the first sale occurred or received refurbishment in recent years. Also, some of the largest courts may have a significant time difference of up to 2-3 years between the opening date of the earliest building and that of the latest. These scenarios are not accounted for in the age data.

Considering local traditions, the regression also includes indicators for whether an apartment has the numbers 4, 8 and 13 in some section of their listed address. The appearance of the number 8, which sound similar to “get rich” in Chinese, is a sign of fortune and thus noted with a dummy variable for “good luck.” The appearance of the number 4, which sounds similar to “death,” is noted with a dummy variable for “bad luck” along with the appearance of the number 13.[38] The full address of each transaction is obtained by using a Google Maps batch search package to generate location information according to the names of the apartments in the dataset.

The Google Maps API is also used to gather part of the location information for each transaction. In particular, an apartment’s driving distance to city central, driving time to city central, and public transit time estimate are generated using the standard Google Maps pathfinder.[39,40] Although Google Maps often fails to properly account for the influence of traffic in time estimates for large,

[36] The price recorded for transactions between eligible applicants is the actual selling price. The price recorded for transactions on the open market, where part of the selling price is returned to the government as a refund of the original subsidy, is the transaction price but with the “refunded” amount removed.

[37] The opening date denotes the date during which the first flat within a given court is sold.

[38] It is noted that Hong Kong is also a nation with a history of British colonialism and observe many of the traditions of the British. Hence the Chinese culture unlucky number and the Western culture unlucky number are both included, but are not treated differently.

[39] The Google Maps pathfinder takes into account various methods of public transportation including, but not limited to buses, express shuttles, metro networks and ferries.

[40] I used the location of the Hong Kong International Finance Centre as an approximation for the center of downtown Hong Kong and the center of the court group as the location of the apartment. The shortest path is selected for the distance estimates. The path with least expected time is used for the time estimate.

densely populated cities such as Hong Kong, this bias is in a uniform direction for all apartments since the actual travel time is almost certainly longer than Google's projection based on the assumption of little to no traffic. The other two estimates, driving distance and public transit time, should be more accurate – they do not vary much unless rare incidents occur, such as emergency highway repairs or breakdowns of the metro system.

The average per capita monthly income of each District Council District is used in the regression as a general control for district-level conditions.[42] These figures are acquired from the 1996 and 2006 Hong Kong Census as well as the 2001 and 2011 by-census, with numbers for years between these dates interpolated. Data for the years after 2011 are extrapolated from the 2006-2011 trend, since the next general census will not take place until 2016. Significant variation of average per capita income exists between districts – the highest average income level for a district is HK$ 41,346 and the lowest HK$ 17,106.[43,44] Subsidized housing courts exist in all districts, but are somewhat concentrated in those with relatively low income levels.[45]

The GIS model used in this paper is constructed from a variety of data sources. The base map, which contains district outlines, is obtained from the Hong Kong Geological Society's (HKGS) free-to-access database. The street and main roads layers are generated from open-source transit network maps obtained from the Open Street Map (OSM) Project, and are designed to account for network breaks such as tunnels and bridges. The elevation adjustments of the model are made using a 30-meter resolution-level TIFF format Digital Elevation Model created by geological survey agency NEXT Maps. Locations of metro stations, current and past airports and schools are collected from the official websites of respective government agencies in Hong Kong.

Schooling availability is measured at both the elementary school level and the 7-9th grade middle school level[46]. The walking distance and incline level between each apartment as well as the nearest full-time primary school are calculated using the GIS road network layer and DEM. In particular, the incline is calculated with the linear distance between the apartment and school and the elevation gap, and is intended as a measurement for the overall difficulty incurred in walking to or from the school caused by a positive or negative slope. The middle school indicator is generated as the total number of DSS middle schools[47] within a 3km radius of the apartment under the assumption that, at the middle school level, the specific closeness of a particular school is not as important as the education quality and availability in the general vicinity of the apartment. Alternatively, a compound education factor can be used for middle school education that retains the 3km distance criteria, but weights DSS schools within the parameter by their walking distance to the apartment with a natural log decay function.

Similar to the education variables, the variables used for metro availability are a combination of distance terms and terms representing the incline level, both calculated with the GIS model.

---

[42] District Council Districts are the basic administrative unit in Hong Kong.
[43] The highest-income district is Wan Chai district, at approximately US$5,331 per annum.
[44] The lowest-income district is Kwun Tong district, at approximately US$2,206 per annum.
[45] Transactions from the Islands district are omitted since there is no land path to the main Hong Kong region and hence no derivable driving time information or reliable public transit time estimates. There is one HOS project, Peng Lai Court, in the Islands district with only 3 transaction records between 1997 and 2014.
[46] Schooling availability data is collected from the Hong Kong Education Bureau (HKEB).
[47] DSS denotes the Direct Subsidy Scheme. This is an education program in Hong Kong that subsidizes private middle schools to provide scholarships for students from low-income families. Schools are selected by the HKEB based on academic merit. They therefore represent affordable, high-quality education for families in the income bracket of HOS apartment owners.

However, unlike the education sector[48], the Hong Kong MTR system has experienced massive changes between 1997 and 2014, with a huge decrease in the average walking distance from HOS courts to subway stations. The metro system also plays an enormous role in the day-to-day lives of Hong Kong residents. Extra attention is therefore given to investigating its influence on property prices – data from the Hong Kong MTR website are used to recreate the metro systems during different periods between 1997 and 2014, with new layers representing the opening of new lines. Using multiple layers, the walking distance and average incline for each apartment involved in a given transaction as of the time of the transaction can be approximated. Details of the construction of metro availability variables and the theoretical aspects of modelling public utilities are discussed at length in section 3.4 below.

Another geographical variable is the linear distance of courts to the airport in meters, adjusted by elevation. The old Hong Kong Kai Tai Airport was closed in July 1998 and replaced by the new Hong Kong International Airport, a change that is reflected in the model. Transactions that occurred before the date are estimated according to their distance to Kai Tai Airport, and transactions that occurred after are estimated according to their distance to Hong Kong International. To further improve accuracy of this estimation, both airports are evaluated in ArcGIS not as points but as polygon areas including the runway and terminals.[50] The Kai Tai Airport, being located in a densely populated area close to downtown Hong Kong, is expected to have a much greater overall negative influence on HOS apartments, but there are also a limited number of apartments close enough to the Hong Kong International Airport to potentially have noise-related issues. Note that no distinction is made here, by design of the variable, between the negative price influences of the two airports.

One might expect airports to also have a positive influence with regard to distance, in the sense that being closer to the airport means lower time costs when traveling by air. However, considering the demographics of the potential owners of HOS apartments, it is not very likely that their income level allows them to enjoy frequent trips or, when alternatives are available, consider air travel as a favorable option as opposed to train or long distance buses. Therefore, I will not take into account any potential positive benefits of being close to airports or quality differences between Kai Tai Airport and Hong Kong International. The closure effect of Kai Tai, if there is any, is simply represented by a change in the airport distance variable for transactions that occurs before and after the date of closure.

Other geographical variables in the model include the linear distance of courts to the coastline adjusted by elevation, linear distance to the district border and linear distance to closest major road. Two definitions of major roads are used. One includes only highways and roads with greater than three lanes in total, while the other also includes three-lane roads and smaller district-connection roads. A distance to major road variable is generated using each definition; in the model it is used as a single proxy for noise, light and air pollution as well as local-level convenience of transportation. These effects work in different directions; hence, the net effect is unclear.

[48] The DSS school and elementary school dataset is from 2013. It is possible that some of the schools in my dataset did not exist when the transaction occurred or, though very unlikely, have already shut down when post-2013 transactions took place. In the regression interaction terms between education availability and time are used to account for this possibility.
[50] Kai Tai airport data are from scanned blueprints of the Hong Kong Department of Planning (HKPD). Hong Kong International Airport data from www.hongkongairport.com.

3.4 Modelling Public Transit Systems

This section outlines the construction procedures of the metro system availability variables used in the hedonic regression model. These variables are designed to take into account two distinct effects: first, the non-linear price effect of the walking distance to the closest subway station, and second, the expectations effect of future, closer stations on current apartment prices. This effect varies not only with the time gap between the transaction date and station opening but also with the relative gain in walkability provided by a new, closer station. Particular attention is given to the use of such anticipation effect adjustments for their potential in analyzing price associations of many types of amenities and utilities in urban markets.

As is the case with many public utilities, metro network expansions exert influence on apartment prices well before the actual stations are put into operation. It is common practice for such large-scale projects to be announced years before construction starts, at which point the future locations of stations begin to become public knowledge. Some segments of the population will hear about such projects earlier than others, either because they have insider information or intensely follow the news, but word would no doubt eventually spread. Also, at least for the Hong Kong scenario, it is virtually impossible to not notice subway network construction sites near one's home. These intuitions suggest that there are non-constant anticipatory effects caused by a gradual diffusion of information throughout a city's population.

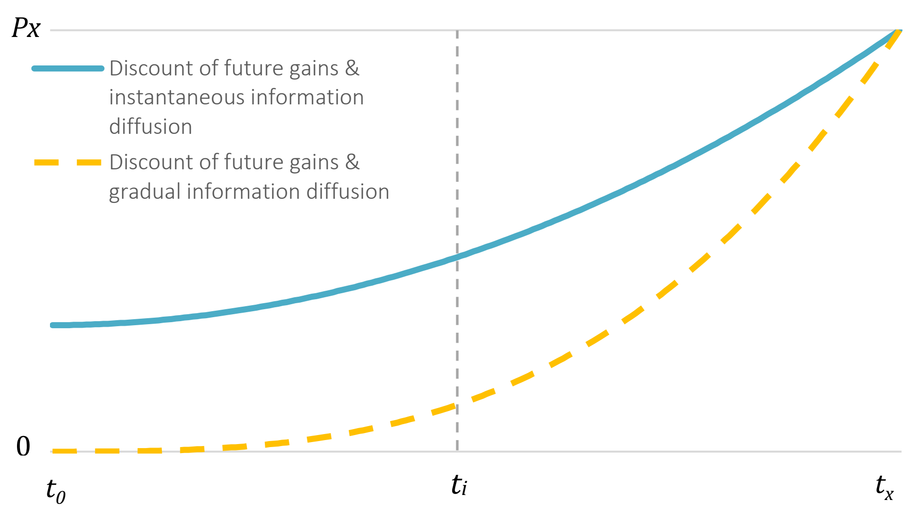
Independent of information diffusion effects, the value of future subway stations should be discounted by their "futureness" or, in other words, the amount of time a potential buyer has to wait until he or she can enjoy the benefits of a future station. Intuitively, if one were to sell a flat a month before the opening of a subway station close by, one would certainly include almost all of the additional value introduced by the future station in the present selling price. Also somewhat intuitive is the fact that this effect does not extend indefinitely – a subway station available ten years in the future should not have much, if any, influence on today's apartment prices.

Note that the size of the price effect of a future station should also vary by *how much* extra utility it brings. If the current closest station is far away and the future station very close, the price effects of the future station should be greater than that of a scenario where the walking distance decrease of moving to a new station is small. This calls for interaction effects between the gain in walkability, measured as the distance difference between the current closest and future closest station, and future-discount and information diffusion effects, measured by the time gap between the transaction date and future station opening date.

Furthermore, it is also possible that closer future stations are discounted less because they are highly desired by residents in the current time-frame, leading to higher-order interaction effects between distance and time gap. However, the modelling approach of this paper simplifies the issue by assuming that individual time-associated valuation of future stations is constant no matter how much extra utility the future station brings. This may seem like a somewhat questionable assumption, but given that new station lines are mostly built in response to the needs of residents and that apartments tend to cluster around established road networks and business facilities within a district, it is unlikely that many subsidized apartments only experience a marginal "new station benefit." In other words, most apartments located in districts that have experienced subway line expansions after 1997 are in fact located fairly close to the new lines, and stand to gain similarly large amounts of economic benefit from the new stations. However, it would nonetheless be beneficial to include in future research provisions that allow time-gap preference curves to change according to distances of the current and future station.

Quantitatively formulating the effects described above, one could simplify the system by assuming that the current closest station has no effect on the apartment price level and the future station at distance x has a positive price effect of P_x. Therefore, the utility gain at any point in time after the construction of the future station is announced and before the new station is put into use is greater than 0 and smaller than P_x. With the endpoints fixed, the functional form of the utility-time plot in between the two points can then be determined. This is fairly straightforward, since ideally both diffusion of knowledge and discounting of future utility should assume an upwards-sloping, second-order convex shape as illustrated in Fig.7.[52] Assuming full information diffusion at the time of announcement, the price premium of the future station is immediately applied with a future gains discount and gradually increases until P_x. The gradual information diffusion scenario has greater curvature because the average price premium at each point in time is determined by both informed sellers and buyers trading apartments at the premium-applied price and uninformed ones who buy and sell at the original, no-effect price.

Fig.7 Time gap-price response of future station, announcement to completion[53]



However, there is no need to account for instantaneous price shocks in the model for this paper. New metro lines are regularly planned well over three decades prior to the construction date in Hong Kong. All of the metro system changes that took place within the time span of the dataset were planned and announced well before the subsidized housing secondary market even existed. Therefore, all dataset observations with a closer future station occur at some time point t_i between t_0 and t_x. With endpoint price levels determined by actual observations, the same basic, convex

52 This is true assuming compound interests and gradual diffusion of information. In reality there may be sudden bursts of the publicity of such programs. It is also possible that information diffuses so fast that it can be seen as perfectly known by all buyers and sellers within a short time after the initial announcement.

53 The initial date of announcement is t_0 and the date of completion is denoted as t_x.

functional form can be used to describe anticipatory price effects, with no need at all for any specific assumption about information diffusion speed.[54]

To this end, the functional form choice for this investigation must take into account both time and distance changes of future stations. This is accomplished in the regression model by including a single "equivalent" distance estimate. This distance estimate adjusts the distance change caused by the completion of a future station with the time gap between the transaction date and new station opening date. The term is given by equation (4), where *D* denotes the equivalent distance estimator, d_c is the walking distance of the current station, and Δd the distance difference between the current and future stations. Δt represents the time gap, which can be measured in either months or years. t_{total} is an arbitrary term that defines a "sufficiently long" time gap: the value of a subway station arriving t_{total} time units after a transaction occurs is assumed to have zero or close-to-zero influence on the apartment's current selling price.

$$D = d_c + \Delta d \cdot \left[1 - \left(1 - \Delta t / t_{total}\right)^k\right] \tag{4}$$

In the equation, the goal is to find a value of *k*, which describes the amount of curvature of the function between t_i and t_x in Fig.7. Using different estimations of *k* and observing changes to the model's statistical strength, one can speculate that a certain function form best fits our valuation of future stations. This can be observed with maximization of the explanatory power of equivalent distance term *D*. A value of *k* greater than one is associated with a convex, second-order positive function, similar to that of Fig.7. A value smaller than one, on the other hand, implies that the function is second-order negative and price premiums of future subway stations are realized earlier than a linear association between price and time gap would suggest. It must be noted that this test is not useful unless it can be demonstrated that using anticipatory effects in the regression model is significantly better than not using them. This question, as well as the best-estimate value for *k* as suggested by the regression model, is discussed in detail in section 4.2.

Switching the focus to current stations only, the relationship between walking distance and price also should not be modelled as linear. Intuitively, the 1-kilometer decrease from six kilometers to five would not be valued equally by potential buyers as the 1-kilometer decrease from two to one. Extra distance most likely matters very little if an apartment is beyond typical walking distances to a subway station. Tests for the extent to which preferences for subway station proximity is non-linear can be carried out either by applying non-linear transformations to walking distance estimates of apartments that are not affected by subway system expansions, or by including extra higher order terms in the regression model.

[54] This implies that, at point t_i, a price premium level somewhere between that given by the instantaneous and gradual information diffusion scenarios is observed in the dataset. Therefore, the curve connecting this point to P_x will always have a similar convex shape and can be modelled as such.

Section IV:

Considering the breadth of topics address in this thesis, the following section is separated into five sub-sections. Section 4.1 discusses regression output using the time dummy hedonic methodology, focusing on addressing the statistical validity of the regression model and evaluating the relative impact of different regression variables. Sections 4.2, 4.3 and 4.4 cover the three central research themes of this paper, namely price effects associated with the availability or lack of public transit options, similar effects associated with the relative density of various types of apartments in Hong Kong, as well as the creation of a hedonic, quality-adjusted price index for the Hong Kong subsidized market. Section 4.5 discusses the relevancy and broader potential application of elevation gradient effects observed in the regression model.

4.1 Regression Results

A best-fit regression model using a 2/3rd power transformation and 88 variables, including dummies for month, year, district of transaction and a select number of indicators for outlier apartment groups reports high overall significance.[56,57] Interaction terms with the time period of transaction are used with a variety of geographical variables, and interaction terms with steepness of slope are applied to elementary school and metro proximity measurements.[58] The model sufficiently explains most of the variance between prices of apartments in the dataset, generating an R^2 of 0.930. After the inclusion of district and apartment group dummies, outlier elimination is limited to the removal of 4 observations with residuals greater than two standard deviations away from the predicted mean.

The statistical strength of the model can be examined through a variety of methods. The model reports an RMSE value of 75.94, considerably smaller than the standard deviation of adjusted price (286.6). The adjusted R^2 of the model is 0.9298, suggesting that over-fitting is not a significant issue. In comparison, a reduced model with only year and month dummies reports an R^2 of 0.647 and an RMSE value of 170.3. A slightly more inclusive model with time terms and transaction-specific terms reports an R^2 of 0.704 and an RMSE of 156.[59] Therefore, it can be established that while a large proportion of the variance in price between transactions can be explained by time period variations and effects of size and age, geospatial effects also substantially influence the selling price of any given apartment.

The effectiveness of geospatial adjustments used in the regression model can be illustrated by a comparison with a model using no geospatial effect terms but indicator variables for each court, 218 terms in total. Since all of the geospatial variables in the actual model are based on court-specific geocoding methods, the maximum amount of quality adjustment offered by using these

---

[56] The 2/3rd power transformation means that $P_{regression} = P_{actual}$^(2/3).
[57] Indicator terms for court no.94 "Kornhill," court no. 190 "Tung Yuk Court" and court no.33 "Yu Shing Court" are added in the regression. The first two courts can be explained as outliers for their geographical location, which are, respectively, extraordinarily favorable and unfavorable. The third court does not seem to display any characteristics that may negatively impact its price. It is assumed that there are certain local level effects not captured by the model.
[58] Time period terms are expressed as a series from 1 to 204, denoting the specific month that a transaction occurred.
[59] The transaction-specific terms for this model are size, age, floor level indicators and discount rate of each flat.

variables should be approximately equal to having spatial and fixed quality effects of each court be represented by the coefficient of the court's indicator. Using the same time dummy variables and transaction-specific variables, the court indicator variable model generates an R^2 of 0.940 and an adjusted R^2 of 0.9396. This implies that with essentially optimized geospatial adjustments, only approximately 1% more of the total variance of the model can be explained beyond what is explained by the actual regression model. The remaining 6% is likely caused by variations in the quality of furnishings, designs of apartment layouts and window orientation. This portion cannot be accounted for without greater knowledge of the conditions of individual transactions. Hence, while there is room for improvement in the design of the model and the geospatial adjustment approach, gains at this stage are fairly limited at best.

Residuals of the regression are plotted against a variety of factors including total price of apartments, time period of transaction, size, latitude and longitude. With the exception of a slight bias towards greater negative residuals for apartments with a gross price of below 2,000,000 HKD, in all cases residuals seem to be fairly evenly distributed across the range of the variable in question.[60] There appears to be no particular latitude/longitude combinations or apartment sizes within the scope of the dataset that display substantially greater residual sizes than the average level. Therefore, it can be expected that the predictive power of the regression model stay relatively consistent across geospatial variations as well as apartments of different quality. Stata-generated residual plots for these factors can be found in the appendix.

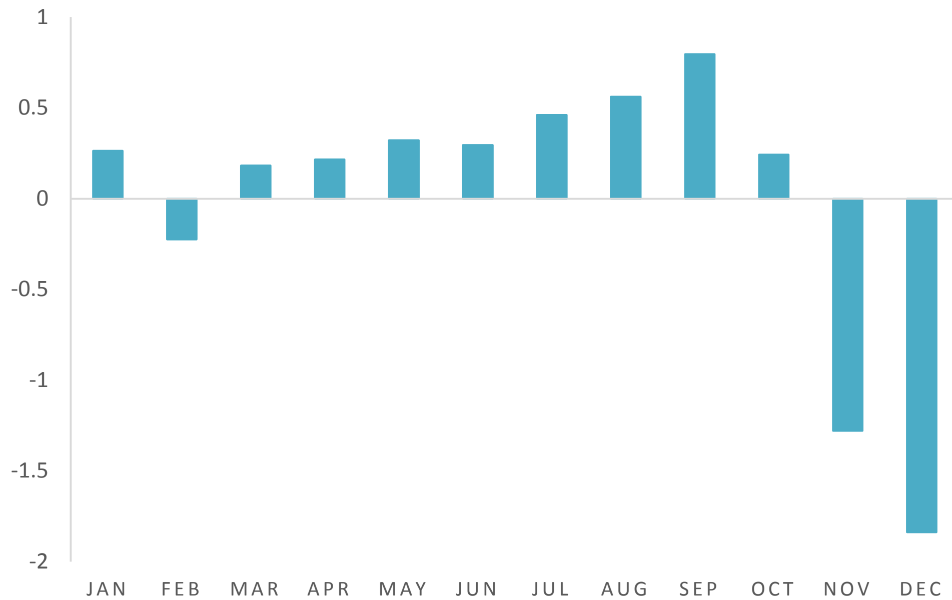
It must be noted that given the nature of the transformation used in the regression model, I have elected to report the influence of most of the variables by price elasticity evaluated at the mean price level of all transactions.[61] This approach gives a readily interpretable result of the connection between quality changes and price at a typical price level, but should not be taken as the extent of a variable's influence on any actual, individual apartment. The influence of quality variables on individual apartments of a given period of transaction should always be considered on a case-by-case basis. Percentage price change predictions from a one-unit variable change nonetheless should be fairly accurate for any period, considering the model's high overall explanatory power and even distribution of residuals across time. Other variables in the regression, in particular those with statistically significant higher-order terms, will require discussions in greater detail.

Several interesting observations can be made from the regression results. Even though the year-and-month indicator approach cannot be used to examine seasonal effects, an alternative modelling approach using time trend terms and the same quality control variables can be utilized to separate within-year market price trends and month fixed effects.[62] Modelling the overall time trend of prices as a quadratic function, significant price effects are found to be associated with being in certain months of the year.[63] Month fixed effects for being in September, November and December are significantly different from the January base level.[64] The effect for being in February is also different from that of January at a significance level of 90%. Fig.8 plots the fixed effects of each month with the average level of effect adjusted to 0.

---

[60] An intuitive explanation is that low-price apartments are not cheap without good reason. Quality effects unique to the single unit (interior damage, previous incidents, etc.) cannot be captured by normal methods of adjustment.
[61] Approximately HK$ 26,149 or US$ 3,352.4 per m^2, the corresponding adjusted 2/3rd power value is 881.05.
[62] The year and month indicator method leaves the month indicators to account for within-year market trends, which can block out the relatively small influences of seasonal effects.
[63] The second-order term is used because of the shape of the price trend in the dataset. Both the 1st and 2nd order terms are significant at the 99.9% level. The cube term is not statistically significant ($P \approx 0.89$)
[64] $P \approx 0.023$ for February, $P<0.001$ for November and December.

Fig.8 Price differences between transactions across months (percent), average = 0



As shown in Fig.8, all month fixed effects are small in magnitude, especially compared to many of the quality effects discussed later in this section. The maximum positive deviation from the 12-month average is 0.80% for September, and the maximum negative deviation is the effect for December at -1.83%. Arbitrage should not allow for the possibility of substantial and regular fixed time-price effects, which is in line with the observed pattern. The pattern itself is perfectly explainable from an intuitive standpoint – prices increase in the summer months leading up to September because parents want to make sure their children move in by the start of the new school year. Prices are low in November and December because of the holiday season and public breaks, and the February drop in price could be caused by Chinese New Year. Holiday season demand and supply are most likely both below the average level, but in the housing market scenario supply might be slightly more rigid because of time-conscious individuals selling to gain liquidity or finance new home purchases. If this is the case, then the mismatch between supply and demand could drive seasonal price fluctuations.

Despite the relatively small variation between the ages of subsidized apartments, there appear to be highly significant, substantial price effects associated with age at multiple orders.[65] Higher-order terms of age in years up to the 6th power report significance at 99.9% and cause statistically significant impacts when omitted from the model.[66] The 7th order term is not significant at the 90% level. While it is intuitively reasonable to expect some non-linear behavior of the effect of age on price, there does not seem to be a straightforward explanation for the high significance of so many higher-order terms.[67] Using all six terms, a ten-year-old apartment can be expected to enjoy a very large, 23% premium over a 20-year-old apartment, which in turn enjoys a 19% premium over an apartment with 30 years of age.

[65] All subsidized apartments are less than 35 years old as of 2014

[66] All terms report P<0.01 using a two-model f-test.

Fig.9 Age influence on apartment selling price, 0-32 years (percent/years)

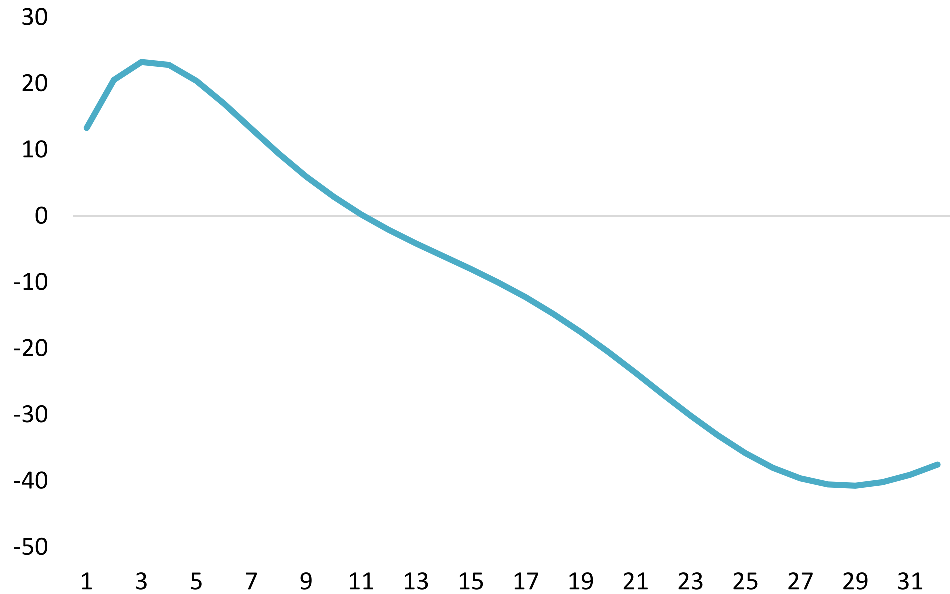


Fig.9 plots expected price influence in percentage terms against apartment age in years. Premiums associated with age are most positive when apartments are three years old and most negative at 29 years. In general, while newer apartments are strongly preferred to older ones, the newest and oldest apartments in the dataset seem to reverse this trend. One possible explanation is that ages greater than 30 years suggest that the transaction likely occurred fairly recently, with a fully recovered housing market and high demand for secondary market subsidized apartments. Conversely, an apartment less than three years old could only have been sold before 2006, while housing prices have yet to return to pre-1997 levels.[68] It is somewhat unlikely, but perhaps not impossible that these effects are not fully captured by time indicator terms in the regression model. This may be the case, for example, if perceived value associated with apartment age actually changes with market performance of the secondary market.

In addition, the oldest apartments in the dataset are more likely to have been renovated either by the owners or as a part of public renovation initiatives, and therefore have on average higher quality than their somewhat newer counterparts. On the other hand, new apartments are much more susceptible to quality problems created during the construction and furnishing phases. While such issues are usually quickly reported and addressed, early owners must bear the costs and endure the inconvenience of repairs. If apartments are not expected to be relatively problem-free until some years after completion, average selling prices of the newest apartments could be lower than that of slightly older ones. Very young apartment could also have vacancy issues and a lack of local infrastructure that could lower their appeal to buyers.

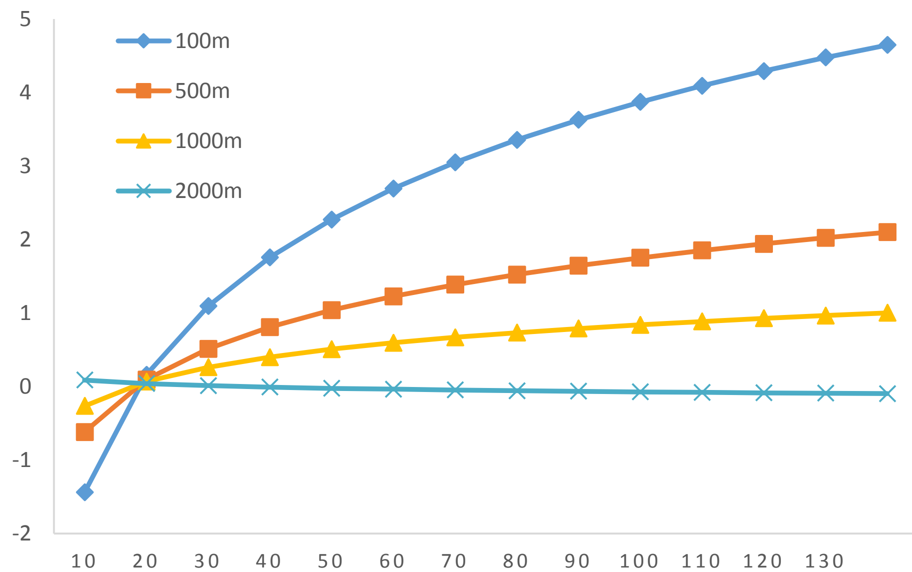
Both apartment size and floor level indicators are significantly associated with transaction price.[69] Holding other factors constant, increasing the size of an apartment by 10% results in a 2.4% price premium. The overall influence of size is likely not substantial given the relative homogeneity of apartments in the dataset with respect to size. Apartments of 14 to 27 floors of height report a 6.3% premium over apartments below 14 floors, and apartments higher than 27

68 The last of the currently existing HOS courts was completed in 2002.
69 All P-values are smaller than 0.001.

floors of height have a 2.4% further premium. Independent of floor level effects, doubling the elevation level of an apartment results in a small but nonetheless statistically significant 0.39% premium.[70,71] Holding other factors constant, an apartment with an elevation level of 5m would sell for 1.6% less than an apartment with an elevation level of 150m.

The possibility of distance to coastline influencing the effect of elevation is tested by adding an interaction term between logged elevation and logged coastline distance. Without considering interaction effects, there is a 0.67% negative influence on price for being twice as far away from the coastline, significant at the 99.9% level. The added elevation interaction term is also significant at the 99.9% level and improves the significance of both elevation and coastline distance terms. Fig.10 plots the elevation influence on price at four different distances to the coastline. As shown below, elevation effects are strongest for locations closest the coastline (100m) and virtually nonexistent for a 2-kilometer distance from the apartment to the coastline. Holding other factors constant, an apartment at 100m above sea level and also 100m from the coastline is at a 5.3% price premium compared to an apartment 100m from the coastline and 10m above sea level.

Fig.10 Elevation-price effects at different distances to coastline (percent/meters)[72]



There is a weaker positive price trend with higher elevation levels at the 500 and 1,000-meter coastline distance levels. At the 500m distance level, an elevation increase from 10m to 50m translates to a 2.37% price premium, and an elevation increase from 50m to 100m translates to a 0.71% price premium. Intuitively, if people mostly desire higher elevation for better views, only buyers of apartments sufficiently close to the coastline will care about elevation level. With 1,000-meter level elevation-price effects being less than 1% on both the negative and positive side and 2,000-meter effects close to zero, the model seems to confirm such intuitions, suggesting a preference of elevation based on good scenery instead of other potential benefits.[73] Negative price

[70] All elevation figures in the model are formatted with sea level as a baseline of 0m.
[71] The elevation effect coefficient is significant at the 99% level with P-value ≈ 0.010.
[72] "0" on the graph reflects price effect of the average level of distance to coastline and elevation from sea level (1.57km and 39m , respectively).
[73] Other benefits of high elevation could include cleaner air, less noise and less moisture.

effects of very low-elevation apartments, particularly those less than 20m above sea level, are mostly likely caused by the threat of flooding and other storm-related damages.

The existence of significant changes in elevation effects at different distances to the coastline raises another interesting question: are there different coastline-distance effects for being on different floor levels? If view effects are a part of the benefits of being on higher floors, it is only reasonable to assume that high floor level premiums are greater when an apartment is closer to the sea. To this end, interaction variables between the two floor level indicators and coastline distance are introduced to the regression model. Interaction terms between floor level indicators and elevation are also included under the assumption that high elevation may enhance the premium of being on higher floor levels.[74]

Fig.11 High/medium floor level indicator premiums at different linear distances to coastline (percent/meters)[75]

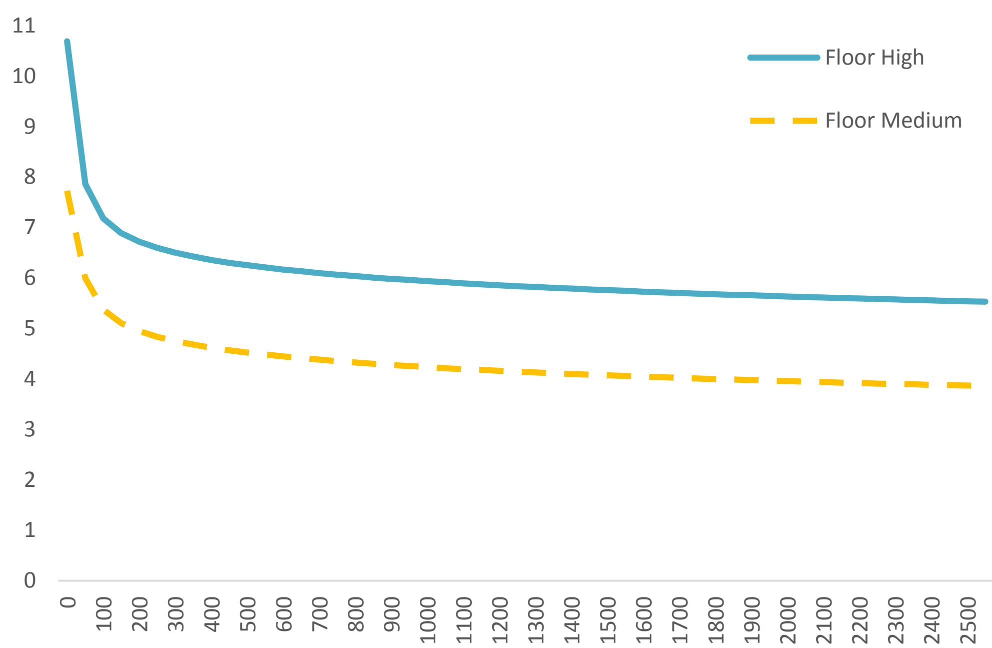


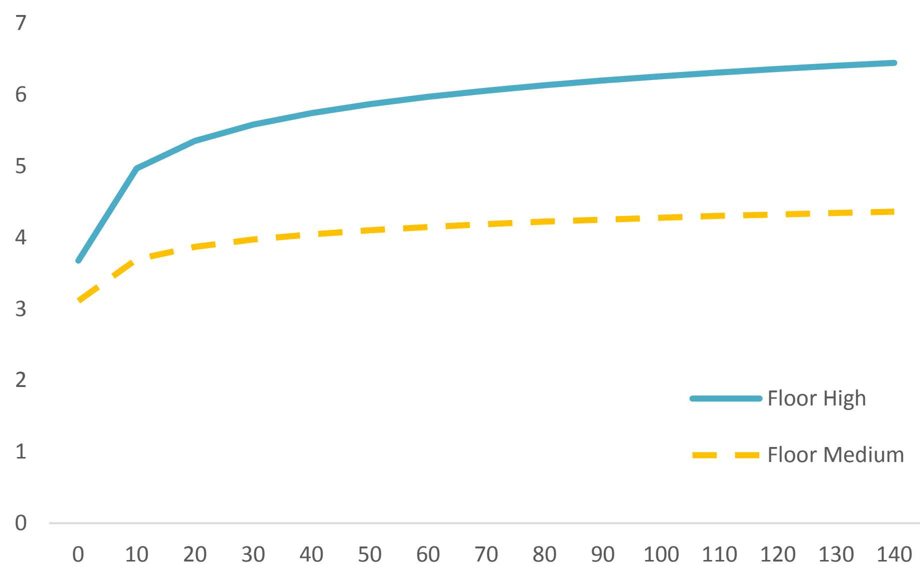
Fig.11 plots the fixed price effect of being at high and medium floor levels for apartments at different distances to the coastline.[76,77] For both indicators, there is a significant boost in price premium for being very close to the coastline. At 100m from the coastline, an apartment above 27 floors is expected to sell at a 7.2% premium compared to an apartment below 14 floors. In the same scenario, the premium of an apartment at floor levels between 27 and 14 is 5.4%. At 2,500m,

[74] Because of time constraints of the thesis program, interaction effects between coastline distance, elevation and floor level indicators are not included in the regression model used in other parts of this section and the investigations in sections 4.2-4.5. Other coefficients in this section and conclusions of other sections are based on the original model, which does not account for these factors. The two models are not substantially different – the original model reports an R^2 of 0.9300 and an adjusted R^2 of 0.9298, while the model with the extra interaction terms reports an R^2 of 0.9303 (adjusted R^2 = 0.9301). Outputs of both regression models can be found in the appendix.
[75] Price effects of floor level indicators are evaluated at the average apartment elevation level of 38.8m. "High" floor levels are floors 28 and above. "Medium" floor levels are floors 14-27.
[76] All interaction terms included are highly significant with $P<0.001$.
[77] Note that these are relative premiums for each distance level. In other words, the baseline in Fig.11 represents the price level of the low floor level apartment (below 14 floors) at the respective coastline distances.

the two respective price premiums are 5.5% and 3.9%. Unlike elevation-coastline distance effects, effects of higher floor levels do not seem to completely disappear at greater distances to the coast. At the average coastline distance of all apartments in the dataset (1,568m), the price premium of a high floor level apartment over a low floor level one is approximately 5.7%. The corresponding price premium of a medium floor level apartment is 4.0%.

This observation underlines a basic difference between elevation effects and floor level effects. Being on a high floor level means that an apartment is almost certainly above local streets and hence above human activities in general, which comes with numerous benefits beyond view effects. Factors such as noise, security and sunlight access positively influence selling prices of apartments regardless of location. On the other hand, elevation changes gradually and over a much larger spatial framework. This explains the difference between the coastline distance response of elevation and floor level indicators – view effects of both factors are small unless an apartment is very close to the sea, but floor level indicators have a non-location specific price influence that elevation effects lack.

Fig.12 High/medium floor level indicator premiums at different apartment elevation levels (percent/meters)[78]



There is also evidence that the effect of being on higher floor varies with elevation. Fig.12 plots the price premium of the two floor level indications for the elevation range of the dataset. At low elevation levels, being on higher floor levels has a smaller positive influence on price. This effect is particularly significant for, but not limited to elevation levels below 20m. Evaluated at an elevation level of 10m and the average distance to coastline, the price premium of high floor level apartments and medium floor level apartments over low floor level ones are, respectively, 5.0% and 3.7%. The same premiums at an elevation level of 100m are 6.3% and 4.3%. Conceptually, it is quite possible that being in a low-elevation region diminishes the benefits associated with higher floor levels. For example, noise could still be an issue for such apartments if there are local roads at higher elevation levels. Apartments with very low elevation levels could also face the risk of flooding and high winds during storms. If elevator systems are easily damaged by flood water,

[78] Price effects of floor level indicators are evaluated at the average distance to coastline of 1,568m.

owners of high and medium floor level apartments in low-elevation regions could actually be at a distinct disadvantage.

Indicators for unlucky and lucky numbers are both highly significant with $P<0.001$. Other things equal, an apartment with numbers 4 or 13 in the address will sell for 2.9% less and an apartment with the number 8 in the address can be expected to sell for 1.5% more. The difference in the absolute sizes between the two effects could perhaps be explained by biases involved in assigning building codes to apartment groups. 17.7% of all apartment groups have the number 8 in their addresses, yet only 6% have either the number 13 or 4. With "lucky" apartments being three times as numerous as "unlucky" apartments, it is quite likely that the former will be less sought after and the latter be seen as even more undesirable.

Other notable effects include a 0.06% price premium for being one extra kilometer closer to the airport.[79] An extra 1,000 apartments within the same court as the transaction apartment lowers its selling price by 1.2%. Raising the discount rate of the apartment by 1% lowers the selling price by 0.6%. A possible explanation for this substantial effect is that larger subsidies were assigned to certain apartments because, for some reason, they were not expected to sell well. Therefore, high discount rates signal negative influences related to the location or design of certain apartment groups. Even if such factors have been mitigated since the apartments' completion, buyers may still act out of experience or intuition and consider high rates of subsidies a sign of potential quality issues with the apartment.

Although it may not be possible to obtain window orientation data for HOS market transactions, it could still be possible in future work, with access to building plans or blueprints, to construct indicator terms for different types of basic apartment designs. Beyond the current quality adjustments for metro station distance and schooling availability, variables for distances to the nearest hospital and police station could be added to improve the overall model design. There are also alternatives for a general proxy of the level of local business activity, for example the number of restaurants in the vicinity or distance to the nearest major shopping center, to the currently used district boundary-income interaction approach. The interaction effects between coastline distance and elevation can be extended to floor level indicators or a combined "total elevation indicator" that includes both elevation from sea level and approximated floor height. Furthermore, if at all possible, the coastline distance variable should be refined to reflect the "quality" of the respective closest point on the coastline – surely proximity to a beautiful beach and palm trees would be valued much more than proximity to a shipyard or cargo bay.

This sub-section has provided an overview of the results of the regression model used in this paper. Effects related to public transit and density of apartments of various housing markets are discussed in the next two sub-sections. Conclusively, the model displays a high level of explanatory power and good standings in statistical terms. Combining the benefits of an inclusive, whole-market dataset with a thorough investigative approach using geo-spatial adjustments

[79] $P \approx 0.01$. A natural log or higher-order transformation is not used here because only the original form is significant at above the 95% level. This goes against intuition since airports are usually associated with noise and traffic. A somewhat probable explanation is that given the high population density of Hong Kong, the noise floor is high enough that only apartment units right next to the airport will see any kind of negative price influence caused by noise. The vast majority of apartments will not be in such extreme proximity to the airport. Therefore, holding other factors constant, it might actually be beneficial to be a bit closer to Hong Kong International for the expediency of traveling. It is also possible that air quality and ventilation conditions are better for apartments closer to the airport, which could be highly valued in a city as crowded as Hong Kong.

provides enough resolution to explore small changes in the effects of quality variables associated with location or time. The dataset itself is perhaps uniquely advantaged in this regard because of its homogeneity as apartments built under the regulations and guidelines of a single government initiative. Apartment quality in terms of furnishings and court-provided amenities is most likely much less varied across HOS flats than across commercial properties in general, lending greater credibility to the coefficients of geospatial terms in the regression model. Most of the independent variables are significant well beyond the 95% level and can be directly linked to price predictions of individual apartments based on their qualities.

While these variables are not the primary focus of this paper, there is little doubt that the ability to quantify their influence on apartment selling price will be of great interest to local regulators and policy-makers. A deeper understanding of geo-spatial effects on housing in Hong Kong could lead to better policies involving the market or optimized planning of new subsidized apartment groups. With the development of a series of new HOS projects well underway as of this writing, these results are perhaps particularly relevant. Many of the effects observed in this model could also conceivably be compared to results of established research on housing markets of other cities, contributing to our understanding of preferences related to housing purchases.

4.2 Metro Availability

Anyone who has ever visited Hong Kong can attest to the importance of the metro system to life in the city-state. One of the most densely populated cities in the world, extremely hostile to bicycles because of the uneven terrain, Hong Kong's reliance on metro networks and the importance of subway stations in the life of the average Hong Kong citizen cannot be overstated. This is perhaps even truer for residents of subsidized apartments who, limited by income and work options, may not only have to seek employment that is further away from home but also not be able to afford cars or access parking space.

Before utilizing the approach discussed in section 3.4 to examine anticipatory price effects of future subway stations, the actual usefulness of considering future stations in the model at all must be tested. A streamlined model with a single term for distance to the closest current station is compared to a full model with an extra term for distance to future station disregarding the time gap between the transaction and the opening of the future station.[80] Using a two-model F-test, the full model is revealed to be significantly better than the streamlined model.[81] The results confirm that future subways station have an "expectations" effect on current housing prices that should not be ignored. Therefore, evaluating public transit availability using only data for the closest current station is insufficient in determining their relationship with the selling price of a flat.

For function (4), 90 months is selected as t_{total} or, effectively, a sufficiently long time span between the transaction and station opening that results in no price effect. The number is chosen because of two reasons: that the longest actual time gap in the dataset is 89 months and, intuitively, that a future station almost eight years away should not be a significant factor in determining current selling prices. Values for k between 0.1 and 5 are tested in the regression with step size 0.1,

[80] If there is no future station closer than the current station, the same distance is used for the two variables.
[81] The two models are different at a high significance level with P<0.001.

resulting in a best-fit value of 1.5.[82] The k value implies that the influence on the selling price of a new station increases faster as the new station approaches completion, but at a moderate rate of second-order growth. In fact, within any one-year period the price appreciation of an apartment associated with a future subway station can almost be regarded as linear in nature. Fig.13 shows the equivalent appreciation curve of the price value of a future station using k=1.5, assuming no current station price influence and that the full price influence of the future station is 1.

Fig.13 K=1.5 equivalent effect of new station (full = 1, percent/months before opening)

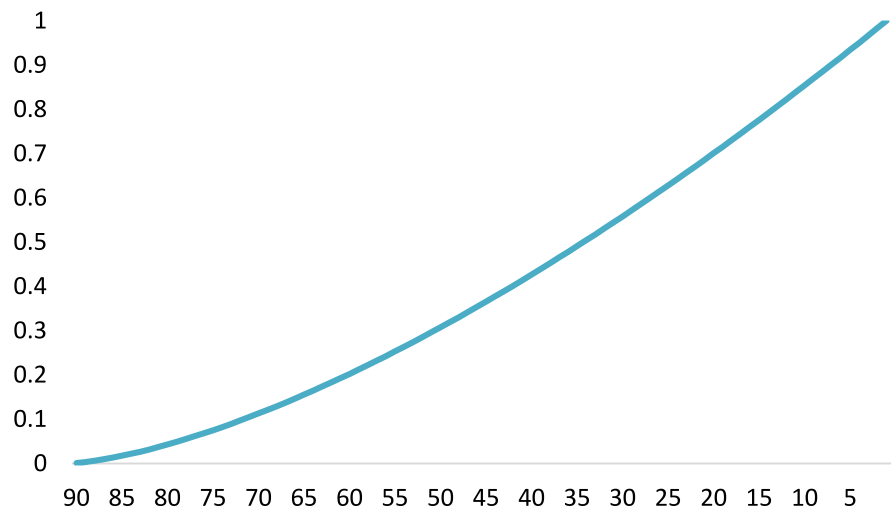


Fig.14 Best fit equivalent distance, new station distance by months before opening[83] (meters/months)

meters\ months	12	24	36	48	60	72
100	273.9	434.8	581.7	713.0	826.8	919.5
200	354.5	497.6	628.2	745.0	846.0	928.5
300	435.2	560.4	674.7	776.8	865.3	937.4
400	515.9	623.2	721.1	808.7	884.5	946.3
500	596.6	686.0	767.6	840.6	903.8	955.3
600	677.3	748.8	814.1	872.5	923.0	964.2
700	758.0	811.6	860.6	904.4	942.2	973.2
800	838.6	874.4	907.0	936.2	961.5	982.1
900	919.3	937.2	953.5	968.1	980.8	991.0

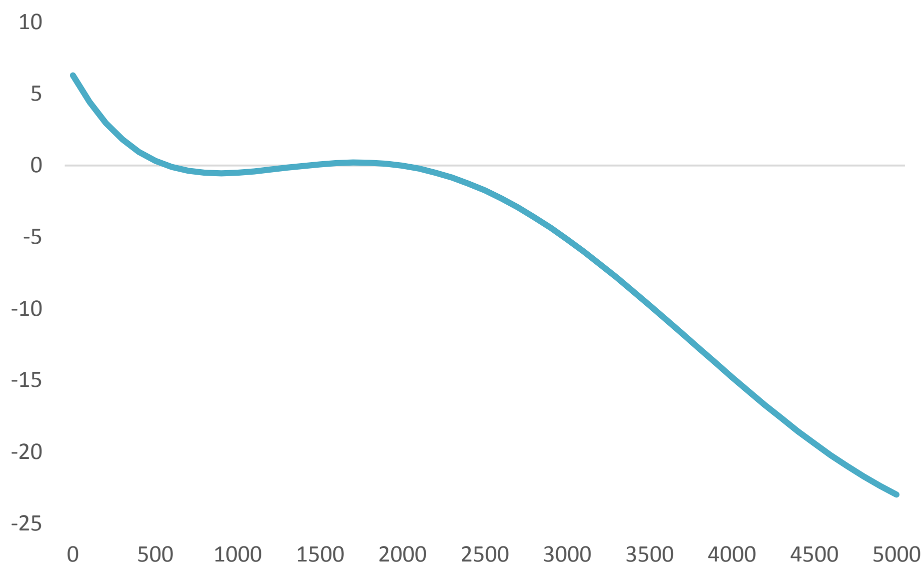
Fig.14 shows an equivalent distance matrix, assuming that the current station is exactly 1 kilometer away from the apartment. A new station that is 100m away and a year in the future will cause the apartment to be priced equivalently to an apartment with a current station 273.9m away. The same new station six years in the future will cause the apartment to be priced as if the current

[82] k = 1.4 and k = 1.6 are not statistically significantly different from k = 1.5, but report smaller R^2 values.
[83] Figures derived assuming that the current station distance is 1km.

station is 919.5m away. These effects are not insubstantial and are in line with the intuition that people who have obtained information about future subway expansions will attempt to take advantage of the situation and increase selling prices on their properties accordingly.

If one removes the influence of future stations, how much do Hong Kong residents value the metro system? The same regression model as described in section 4.1 is performed on the 30,528 transactions that did not experience a local MTR expansion between 1997 and 2014 which resulted in a closer station. The model reports higher-order terms of walking distance estimation variables up to the 8th power, all significant at the 99.9% level. Fig.15 shows the estimated price influence of a metro station at distances 0 to 5,000m, assuming that the average-distance station has no value.[84] It is apparent that being in close proximity to a subway station is highly valued, yet the drop in value associated with greater distance slows down between approximately 700m and two kilometers. Intuitively, while living right next to or, more likely for Hong Kong, directly on top of a subway station is hugely beneficial, being slightly further away when one's apartment remains within walking distance of a station does not matter much.

Fig.15 Closest station distance relationship with apartment selling price (percent/meters)



Holding other factors constant, an apartment that is 100m away from the closest metro station is at a 4.9% premium compared to an apartment that is identical in every other way but a kilometer away. An apartment that is a kilometer away is at a 4.6% premium compared to an apartment 3 kilometers away, which is in turn at a large, 17.8% premium compared to an apartment 5 kilometers away. It certainly seems that there is a "threshold of walkability" present in the public perception of subway stations in Hong Kong: apartments that are within the threshold command a hefty premium over those that are not.

Thus far, the model has assumed that there are no elevation factors at play. However, common sense indicates that elevation gradients or differences in altitude can have significant effects on the perception of distance – walking uphill or downhill will, quite intuitively, feel more laborious than walking on a perfectly level plane. Considering the extremely uneven nature of Hong Kong's

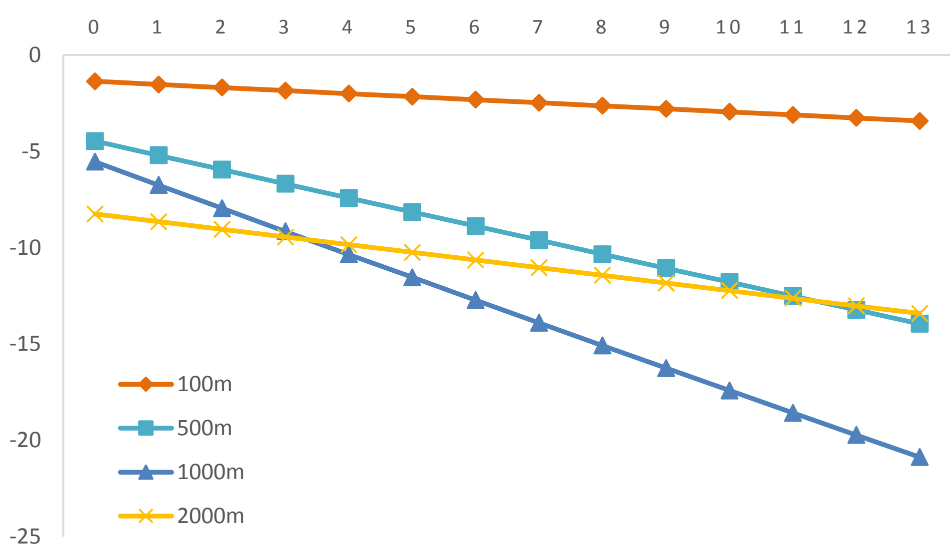
[84] The average closest station walking distance for apartment with a single closest station throughout the time duration of the dataset is 1422.6m. Fig.15 assumes that this distance has zero effect on the selling price of an apartment.

terrain, it seems only reasonable to assume that, for public utilities, distance parity alone does not equate to price parity.[85]

With this issue in mind, a new model can be constructed where the influence of elevation gradients on distance is assumed to be linear. This may be a questionable assumption given the possibility of threshold effects similar to those with station distance, but it allows for a relatively straightforward description of the relationship between different elevation gradients and price effects associated with metro station distances. It is also assumed in this model that there is no pure elevation gap effect. In other words, the effect of the altitude difference between the closest metro station and the apartment is set to zero for zero distance. The reason for this limitation is quite simple – if a station is directly beneath or above one's apartment, one ought to be able to simply take the elevator.

Note that all stations are assumed to be homogenous in service quality, local amenities and, given that an individual is within the entrance estimation range of 50m from the platform, provide identical convenience of transportation. Furthermore, no distinction is made between the apartments in question being uphill or downhill from the metro station. The reasoning is straightforward – it is assumed that when an individual walks to the subway station he or she must eventually return to the apartment. Therefore, the "total experience" of a one-time use of the subway system will always be one walk uphill and one downhill. It is nonetheless still possible that one might prefer to walk downhill or uphill first or at the start of the day, which would not be accounted for in this model.

Fig.16 Price effect of elevation gradients at different fixed walking distance levels, assuming linear response (percent/degrees)[86]



Using the current station distance model, which only includes observations without local new station constructions during the time coverage of the dataset, an additional interaction term between walking distance and slope is significant at the 99.9% level. Fig.16 plots the price influence of four given walking distances to the nearest subway station at different elevation

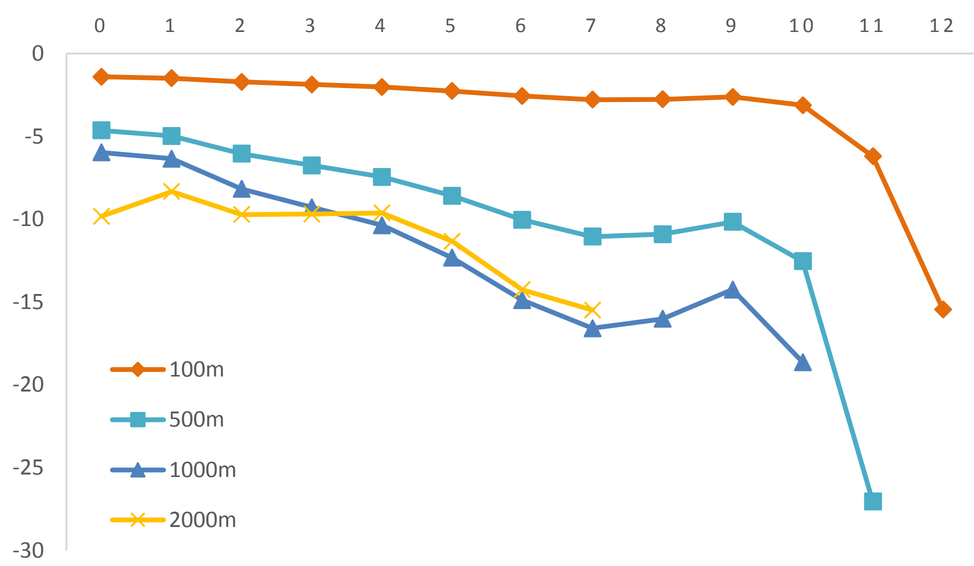
[85] A quantitative comparison of Hong Kong's elevation variance to that of other cities can be found in section 4.5.

[86] To provide a singular reference frame for price effects of different distance levels, a closest subway station with zero distance and zero elevation difference from an apartment is assumed to have zero effect on price for this graph as well as Fig.17. Note that this is different from Fig.15, which is equalized at the average walking distance.

gradient levels. As shown in the graph, price effects of a fixed distance can change significantly when slope is involved. Compared to a flat ground scenario, introducing a 5-degree slope to a 1-kilometer distance causes a 6% negative impact on selling price. Increasing the slope to 10 degrees, close to the maximum level of the dataset, translates into an 11.9% price decrease from the 1km flat ground distance. Distances of 100m and 500m exhibit similar trends, albeit to a smaller degree. A 2-kilometer distance scenario also slopes downwards when plotted against elevation gradients, but with milder price effects at higher elevation levels. It is possible that with great enough a walking distance, people pay more attention to the distance itself and less to elevation differences. Also, assuming a fixed maximum elevation difference, distance is inversely related to average slope. If few or no apartments have both large closest station walking distances and elevation gradients, the fitted values for elevation gradient interaction effects may not be useful in predicting price effects of such combinations.[87]

To consider the possibility of non-linear effects of gradient, several significant higher-order terms are added to the model.[88] Fig.17 plots the results as price influences of distances 100, 500, 1,000m and 2,000m at various gradient levels up to the maximum elevation gradient level at each respective distance. Overall, elevation seems to have a significant effect on price influence of a given distance, which increases dramatically past 9-10 degrees. In particular, an apartment with a closest subway station walking distance of 500m sells for 7.9% less if there is a 10-degree gradient involved. An 11-degree gradient decreases price by a further 14.5%, though evaluations at extremely large gradient values are expected to be not as accurate as those at smaller values. Net price effects of elevation gradients seem to be strongest at 500-1000m levels, which translates into slightly further but nonetheless walkable ranges. However, they can also be observed from the 100m and 2000m curves.

Fig.17 Price effect of elevation gradients at different fixed walking distance levels, with nonlinear elevation effects (percent/degrees)



87 The maximum slope for an apartment with an actual metro station walk distance of greater than 2,000m is 7.1 degrees. The maximum slope for apartments with walking distances greater than 1,000m and 500m is 10.2 degrees and 10.9 degrees, respectively.
88 Terms of 2nd to 5th power are significant at the 99.9% level, and the 6th power term is significant at 95% level.

Note that since the data involve aggregated elevation gaps but not slope level estimates of actual roads, a gradient level of greater than 10 degrees may indicate much steeper real slopes. Roads designed for pedestrians have no real slope limit – a 10-degree overall elevation gradient could come from paths with slopes of 30-40 degrees. At such slope levels there are stairs and most likely also steep turns involved, which could explain the large negative price effects of elevation gradient levels beyond 10 degrees. Furthermore, the limited 30-meter resolution of the DEM could smooth out elevation gaps, leading to elevation gradient estimates that are slightly smaller than actual levels involved.[89] Therefore, real elevation-price effects are likely to be marginally weaker than those suggested by the result in this sections.

In short, it is clear that there are observable, significant nonlinear gradient effects associated with the perception of distances to public utilities such as subway stations. These effects are independent of elevation levels of both the apartment and station, confirming the intuition that people tend to avoid walking either up or down mountainous terrain. If such effects are not taken into consideration when planning or constructing subsidized apartments, there may very well likely be unnecessary value loss caused by building apartments too far up the hillside. Furthermore, if private housing agencies in Hong Kong determine purchase or sale prices by algorithms that do not take into account gradient effects, there could be regular trends of paying too much for houses with large elevation gradient effects or too little for houses with small effects. Other broader implications of such effects are discussed at length in section 4.5.

Using a hedonic, time dummy regression model and period-specific subway walking distance estimates, this sub-section investigates the influence of metro availability on transaction price of apartments. It is concluded that there are significant price effects associated with the distance to the closest station, future availability of subway stations as well as elevation gradient levels between apartments and stations. In particular, future stations as far as four to five years away can still have non-negligible positive effects on current selling prices. These relations are predominately non-linear, with higher-order interaction terms at high statistical significance.

A more in-depth analysis of the topic might include data for the second-closest station, fixed preference for particular stations and lines, effects associated with different waiting times of different stations, and differentiated preferences for uphill and downhill walks. Second-closest station data may be important if there are large disparities between the expected wait times of different stations, or if shopping centers and other amenities vary greatly between stations. It is also possible that individuals place value not only on actual walking distance but also on the difficulty of navigating, estimable by number of turns or the ratio between linear and walking distance. Using a distance-slope profile instead of a single mean slope estimate could result in more accurate gradient effect estimates.

4.3 Density Spillovers

This sub-section investigates the relationship between the three main housing markets in Hong Kong from the perspective of the subsidized apartment market. Established research suggests that highly subsidized housing markets have negative price effects related to density, lowering both

[89] Slope levels of the first and last 30m of a walk from the apartment to the station cannot be accounted for with a 30-meter resolution DEM.

their own values and the values of houses in the vicinity (Autor et al., 2014). While this issue is well-documented for rent subsidy markets, there is little literature on purchase-subsidy housing markets such as HOS. Theoretically speaking, subsidizing purchases would be expected to have smaller negative effects, since ownership of property would give individuals greater incentives to maintain their residencies and contribute to the local community. Home ownership is also empirically associated with better child outcomes (Haurin, Parcel, & Haurin, 2002) and lower crime rates (Ni & Decker, 2009; Rohe & Stewart, 1996). However, subsidy levels on HOS apartments are comparatively large even by rent subsidy standards, which could imply a greater income gap between subsidized apartment residents and commercial apartment residents. Hong Kong is also well-known for high levels of social inequality, potentially contributing to a negative public perception of subsidized apartment owners.

To better understand the relationship between the public rental unit market, subsidized apartment market and commercial market, the respective proportion of units of these markets among total units within a series of fixed radii is included in the regression model. To control for price effects related to absolute density, the total number of units of all markets in the vicinity of each subsidized apartments in the dataset is also included, using search radii of 1-5 kilometers. These variables, along with the existing, extensive set of geospatial control terms, should almost completely account for local area effects.

I would like to note that all density estimates used in this sub-section are fully controlled for individual time periods. In other words, an apartment density estimate for a transaction on a particular date will only contain public, subsidized and commercial apartments completed on or before the previous year. The one-year lag is used because some buildings have different opening dates for different floors or lack data on the specific month of first sale. The lag ensures that apartment groups considered in the density measurements are most likely both in place and relatively well-populated during the time period of the observation in question.

Assuming that spillover effects of greater relative densities is fixed at different absolute density levels, the public rental unit market is first considered. The percentage of public rental units among total units of all three markets is calculated for radii 1-5 kilometers, representing different levels of what can be perceived as the "neighborhood." These figures are included in separate regression models which controlled for the total number of apartments in all five radii. Data for radii 6-9 kilometers are also collected, but not used in the analysis because the search area sizes involved appear to be far too large compared to the size of the entire HKSAR region[90].

[90] The total land area of Hong Kong is only 1,104km^2 , and some 98% of all apartment groups are within 20 kilometers' linear distance of each other (213 out of 219).

Fig.18 Price effect of 0.1-unit increase in proportion of public rental apartments among all apartments within fixed search radius, 1-5 kilometers (percent)

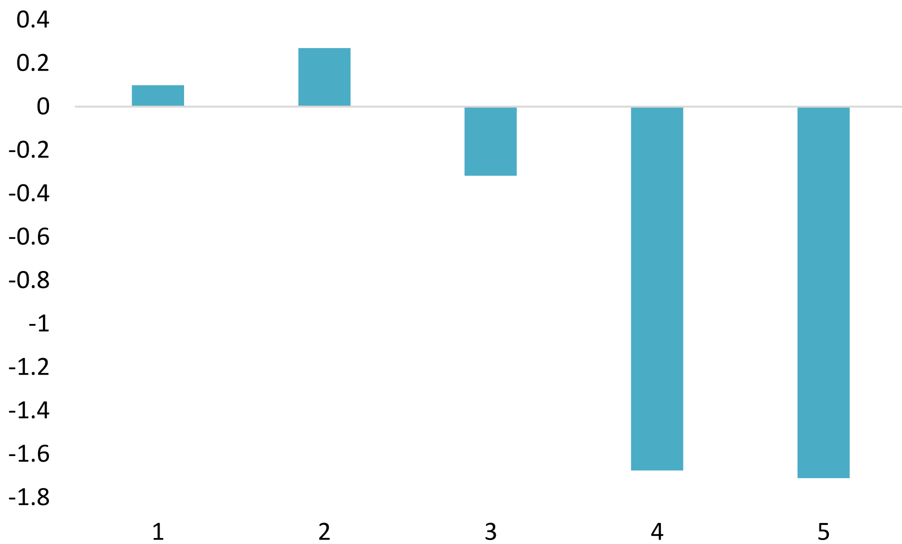


Fig.18 shows the price effect of a 10% or 0.1-unit proportional increase[91] in the percentage of public rent units among total units at different search radii from 1 to 5 kilometers. Smaller radii report positive effects, most likely caused by the model's limitation in considering each apartment group as a single latitude/longitude combination. Public rental unit courts, being relatively few in number but large in size, would be underrepresented in the data at smaller search radii simply because a number of them cover too much land to be accounted for with point estimates and search parameters with such radii.

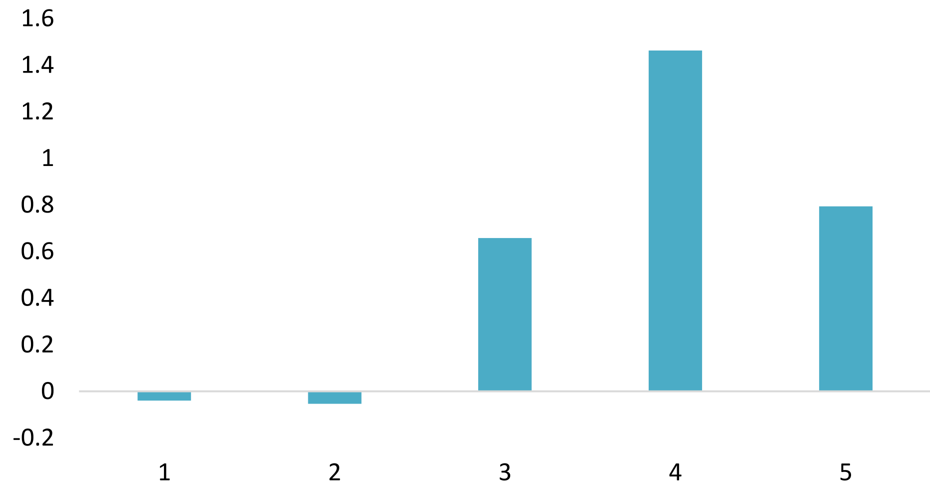
At 3 kilometers and above, particularly for 4 and 5 kilometer search radii, there are significant negative price effects associated with relatively greater densities of public rental units.[92] Evaluated at the maximum-effect radius of 5 kilometers, a 0.1-unit in the proportion of public rental units in the neighborhood translates into a 1.7% price decrease. According to the estimates, a subsidized apartment with the average percentage of neighboring apartments being public rental units, about 32%, will sell for approximately 5.4% lower than an apartment without any public rental units within a 5-kilometer radius.

A similar analysis procedure is conducted for commercial apartments. Fig.19 displays price effects of a 0.1-unit in the proportion of commercial apartments at respective radii.[93] Small search parameters report insignificant negative relationships between commercial apartment ratio and price. However, at search radii of 3 kilometers and above, a larger proportion of commercial apartments in the neighborhood of a subsidized apartment has a significant positive effect on its selling price. The effect is maximized by a search radius of 4 kilometers, where a 0.1-unit in the

[91] Assuming linear influence of increasing relative density. 10% here means a change of 0.1 units in the percentage apartment share of a market. For example, an increase from a commercial apartment proportion of 30% to 50% is a 0.2-unit or 20% change. The three markets are tightly entangled in Hong Kong because of the high population density and, while there is significant variation in the shares of the markets in different locations, very few locations have a single dominant sector of housing. Therefore, although a better assumption to use might be a logistic fit or higher-order transformation, the linearity assumption is more than likely sufficient for estimation over the existing spread of the relative density estimates.
[92] Variables at 2-5 kilometer radii are significant at the 99.9% level. The 1-kilometer variable is significant at 90%.
[93] Variables using 3-5 kilometer radii all significant at the 99.9% level.

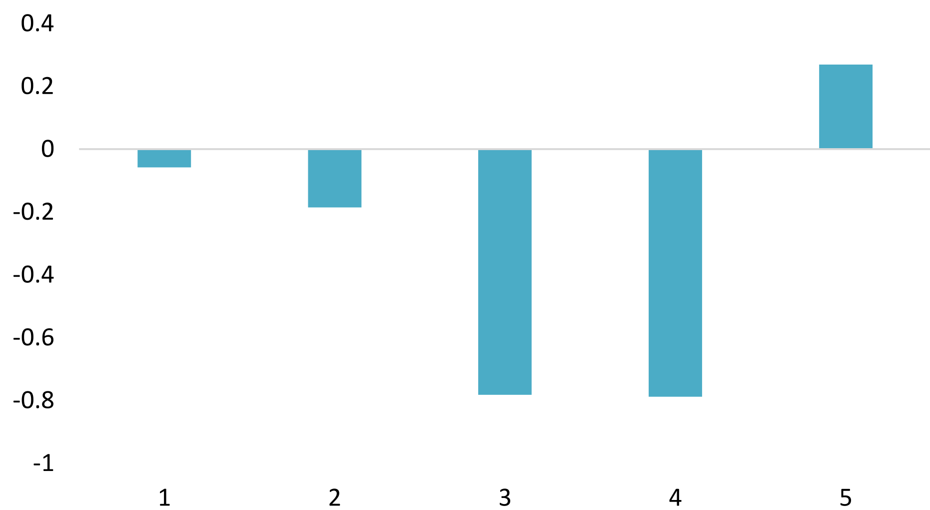
proportion of commercial apartments translates into a 1.46% increase in the selling price. The average price effect of radii 3-5 kilometers is 0.97% per 0.1-unit proportional increase.

Fig.19 Price effect of 0.1-unit increase in proportion of commercial apartments among all apartments within fixed search radius, 1-5 kilometers (percent)



The significant density-price effects of the commercial and public rental unit markets raise the following question: do subsidized apartments have density-related effects on their own price levels? Adjusting for total density and the size of the court that the transaction belongs to, the proportion of subsidized apartments, using search radii of 1 to 5 kilometers, are added to the same regression model as the two previous scenarios. Fig.20 reports the percentage influence on price of 0.1-unit increase in the proportion of subsidized apartments of each search radius. As is the case with the other two markets, search radii of 1 and 2 kilometers report comparatively small values.[94] Considering the relatively small and somewhat ambiguous price effects at such search radii for all three markets, it is likely that they are simply too small for the evaluation of density effects measured in groups of apartment buildings.

Fig.20 Price effect of 0.1-unit increase in proportion of subsidized apartments among all apartments within fixed search radius, 1-5 kilometers (percent)

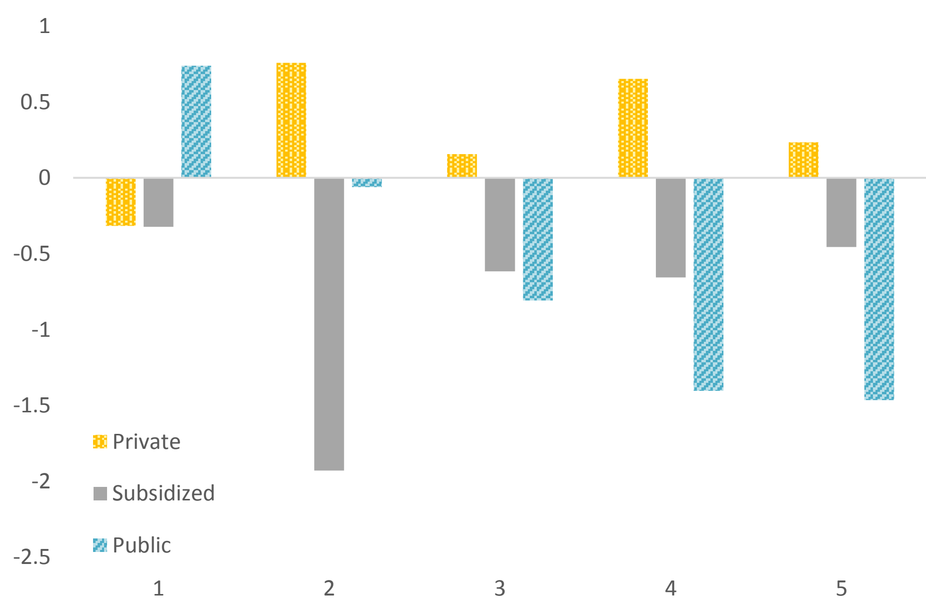


[94] The two terms are nonetheless significant at the 99% level.

However, results for larger radii are more interesting. Radii of 3 and 4 kilometers report highly significant negative effects between density proportion of subsidized apartments and price. At a respective price effect size of -0.78% and -0.79% for a 0.1-unit increase in subsidized apartment proportion, subsidized apartments have a negative spillover effect on themselves comparable in magnitude to that of public rental apartments. The 5-kilometer radius data show a significant positive effect on price. A not improbable case could be made that at small search parameters, proximity to commercial apartments is desirable for their amenities and quality spillover effects. This effect lowers the attractiveness of higher proportions of subsidized apartments at these radii. On the other hand, at larger search radii a high concentration of subsidized apartments signals that the region in general contains business and recreational facilities that are in-line with the purchasing power of subsidized apartment residents. While sounding somewhat far-fetched, this approach would also explain the fall in the positive price effects of commercial apartment density at the 5 kilometer search radius. It must be noted that all density effects presented above are nonetheless small compared to subway station distance effects or elevation-price effects.

Taking the analysis one step further, a more nuanced approach involving interactions between absolute density and market proportion effects can be applied. Using a series of interaction terms between relative proportions of the three markets and total density of a given search radius, effects of each market's relative density can be evaluated at the total density level of each radius.[95] This would provide a straightforward estimation of density spillover effect sizes of the representative apartment while taking into consideration that price effects of the proportion of a certain market may behave differently at different population density levels. For example, at low total density levels density-price effects may also be smaller because residents are less aware of their local population configuration. As is the case with gradient effect interactions, it is assumed in the model that there is no spillover effect at zero density. The regression does not contain a fixed variable for relative density, forcing the series of interaction terms to explain the spillover effects.

Fig.21 Price effect of 0.1-unit increase in proportion of apartments for different markets within 1-5 kilometer search radii, at respective average total densities (percent)



[95] Total density is evaluated as the sum of the number of flats of all three markets at a given radius.

Fig.21 describes the effect of a 0.1-unit proportion increase in the three markets at radii 1-5 kilometers, evaluated with interaction effects between relative density effects and total density. The numbers reflect density-price effects at the average level of total density within the five radii.[96] Similar to the elevation gradient effect investigation, all results are acquired by fitting functions containing market proportion terms interacted with total density terms transformed to different orders. In all cases, the highest-order transformation used is between 4th and 7th power. Search radii 3-5 kilometers uniformly report a significant price premium associated with higher proportions of commercial apartments and lower proportions of subsidized and public rental units. The negative price effect of greater public rental unit proportions is also uniformly stronger than the effect of subsidized apartment proportions at 3-5 kilometer radii. Using the 4-kilometer radius as an example, a 0.1-unit increase in the proportion of commercial, subsidized and public rental apartments in the neighborhood of a given apartment can be associated with a 0.65%, -0.66% and - 1.47% respective price change to its selling price.

At these levels, given that the linearity assumption is valid and holding other factors constant, an apartment with no public rental units within a 4-kilometer radius is at a 4.6% price premium compared to an apartment with the average proportion of public rental units.[97] An apartment with no commercial units within the radius can be expected to sell for 2.5% lower than an apartment with the average neighborhood proportion of commercial apartments.[98] Since a higher relative presence of one market is usually associated with a lower presence of another, the actual effect of a commercial or public rental development project on subsidized apartment prices is most likely an aggregate of price effects of proportional changes of all three markets.[99]

Estimates using 1 and 2-kilometer radii also report statistical significance, but are not consistent with the trend displayed by estimates derived with larger radii. Aside from the aforementioned issue with using single point locations to approximate apartment groups, a second possibility is that the local apartment group of a transaction increases absolute density levels of subsidized apartments by introducing fairly large amounts of flats with zero distance to the search center. While these effects exist for all search radii and are controlled in the model by a separate "own court size" variable, their influence on total density measurements could be particularly great at smaller search radii, causing distortion in interaction terms. Either way, such radii seem to be inherently problematic, at least if the current methodology is to be used.

Using 3, 4 and 5-kilometer search radii, best-fit curves of the price effect of a 10% or 0.1-unit proportional increase in the three markets are plotted for different level of total residency density, measured by number of apartments for a given search area. Fig.22, 23 and 24 describe the effect on selling price of a 0.1-unit increase in the proportion of the three markets at the three search radii. For all three radii, a clear distinction among the price effects of the three markets can be found at higher total density levels, which grows in magnitude as total density increases up to the maximum value for each search radius. In all scenarios, the price effect of the public rental unit market is generally negative and slopes downwards as total density increases. Similarly, larger commercial

---

[96] The average number of apartments for all three markets are 29,185 units, 65,648 units, 104,717 units, 153,889 units and 213,745 units for 1, 2, 3, 4 and 5 kilometer radii, respectively.
[97] The average proportion of public rental units of all transactions in the dataset is 33.2% at a 4-km radius.
[98] The average proportion of commercial units of all transactions in the dataset is 38.9% at a 4-km radius.
[99] Consider an apartment with 10,000 units from each of the three markets. An extra 5000 commercial units increases the commercial proportion by 9.5% and decreases both the subsidized apartment and public rental unit proportion by 4.7%.

market proportions are generally associated with a price premium, only displaying negative effects at small absolute densities using the 3-kilometer search radius.

Fig.22 Density spillover effects at different total densities, 3km search radius (% price/ 0.1-unit proportion change), 0-200,000 total apartments

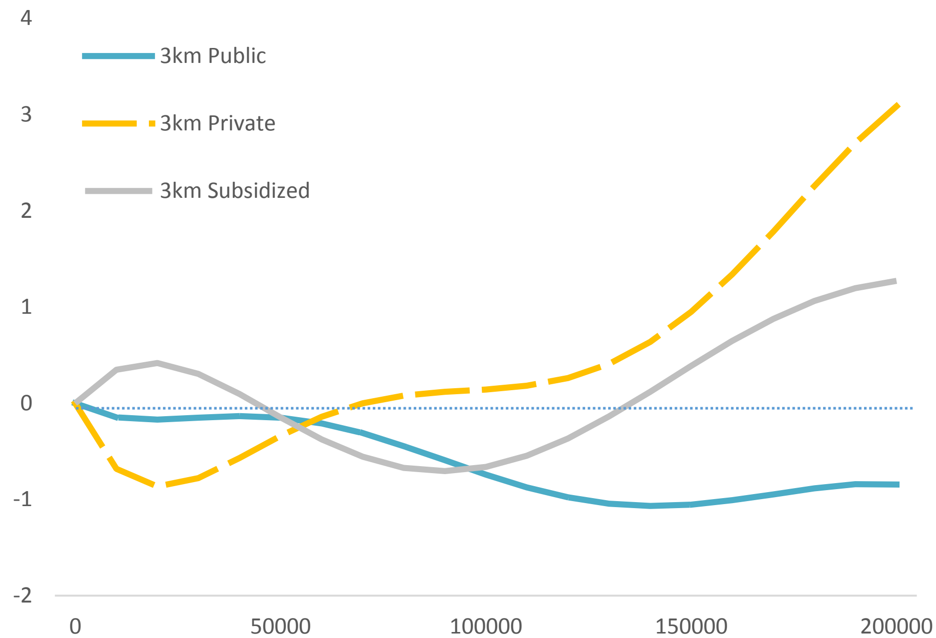


Fig.23 Density spillover effects at different total densities, 4km search radius (% price/ 0.1-unit proportion change), 0-280,000 total apartments

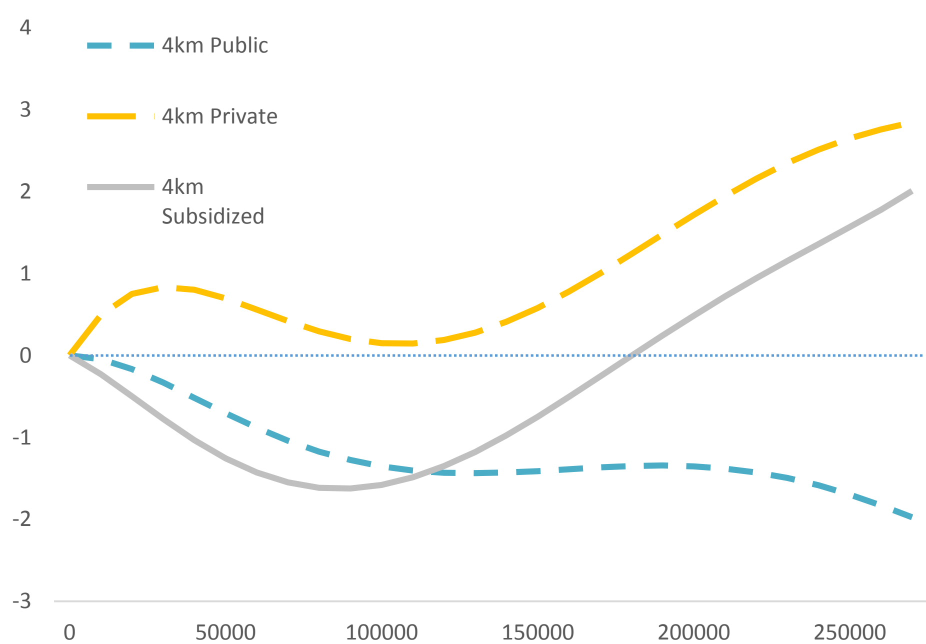
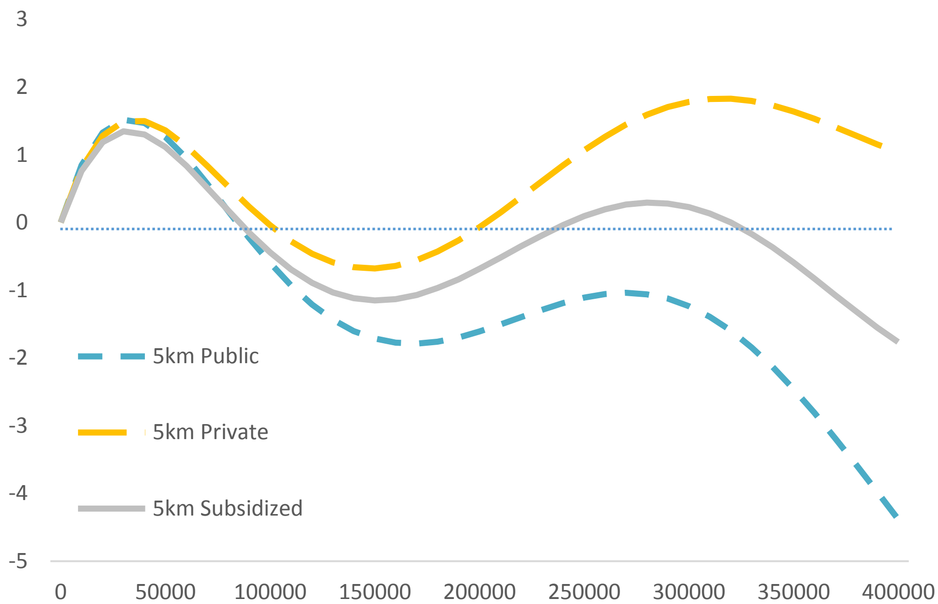


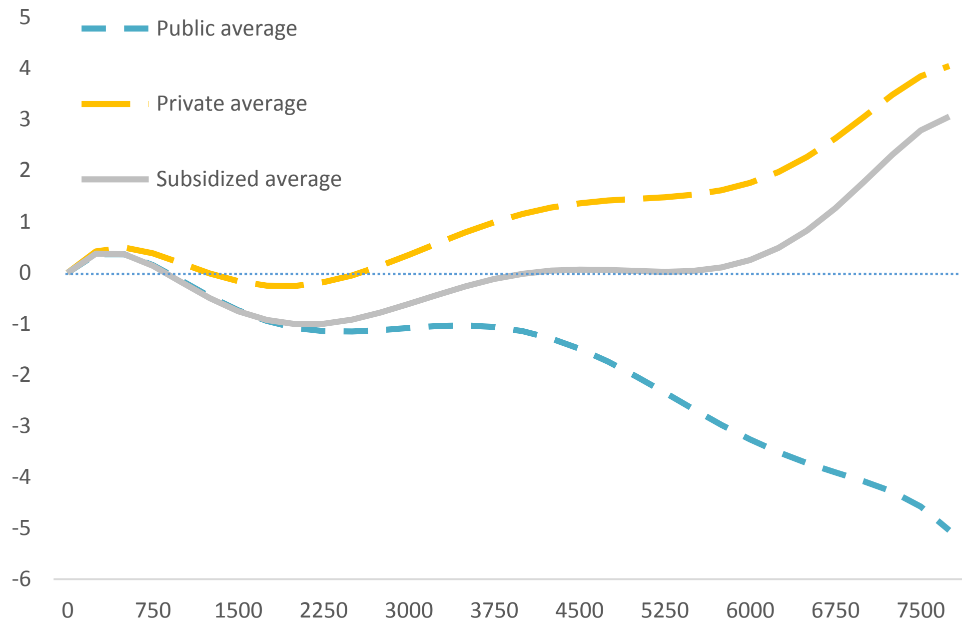
Fig.24 Density spillover effects at different total densities, 5km search radius (% price/ 0.1-unit proportion change), 0-400,000 total apartments



In all three cases, the response to different total density levels behaves in a non-linear pattern. This effect is most pronounced in the 5-kilometer search radius scenario, where all three markets respond in a similar pattern, with similarly-sized positive proportion-price effects for a 0.1-unit increase in market proportion at low total density levels and similar shapes of curvature at higher total density levels beyond 150,000 units. Note that the similarity of the sizes of the effects at low density levels implies that for such densities, there is essentially no overall effect for a proportional increase in any particular market, as increases in the proportion of one market is balanced out by the decrease in relative density of the other two. Effects are not as uniform in the other two search radius scenarios but are nonetheless generally non-linear.

To improve the quality of price effect estimations and reduce potential errors of the fixed search radius models, the previous results are combined by averaging price effects of the three radii at the same relative housing market density, evaluated as apartment unit count per km^2. Fig.25 shows the density-price effects of the three markets evaluated for a range of relative apartment density values up to 8,000 units/km^2. At density levels of less than 1,500 units/km^2, all three markets display similar price responses at less than 1% overall for a 0.1-unit increase in proportionality. This suggests that for low density levels, there is virtually no net effect on selling price of an expansion in any of the markets.

Fig.25 Averaged density spillover effects at different relative densities (% price/ 0.1-unit proportion change), 0-8,000 apartments/km^2



At moderate relative density levels above 2,500 and below 4,000 units/km^2, the commercial market proportion has a positive influence on price. The subsidized market proportion has a negative influence smaller than that of the public rental unit market. Evaluated at the 4-kilometer mean density of 3,281 units/km^2, a 0.1-unit increase in the proportion of commercial units increases the selling price by 0.58%. A 0.1-unit increase in the proportion of the subsidized market and public rental market reduces the selling price by 0.43% and 1.04%, respectively. These figures suggest that while being close to subsidized apartments is still relatively undesirable compared to being close to commercial apartments, the overall effect is much smaller than that of the public rental unit market, and can be ambiguous when measured at typical density levels.[101] At high density levels beyond 4,000 units/km^2, the price effect difference between the commercial market and public market widen. Greater subsidized market proportionality becomes unequivocally positive at density levels greater than 6,000 units/km^2. It must be noted that such density levels, while rare for cities in general, are common for the densely populated Hong Kong.[102]

A detailed examination of the explanations and causes of the observed effects is beyond the scope of this paper. However, it can be concluded that subsidized housing programs such HOS have a far smaller negative spillover footprint than public units of a rent-subsidy nature. While effects of subsidized market density are ambiguous at low total density levels, they are significantly different from public market effects at average levels and clearly positive at higher total densities. Considering the size and price levels of the overall Hong Kong property market,

[101] Assuming that all three markets have 10,000 units. At the 3,281 unit/km^2 price effect levels, an extra 1,000 units in the subsidized market has a net price effect on subsidized apartments of -0.38%. An extra 1,000 units in the public rental unit market lowers subsidized apartment selling prices by 1.06%. At a slightly higher density level of 4,000 units/km^2, an extra 1,000 units in the subsidized market lowers selling prices by only 0.02%.

[102] The maximum 4km^2 radius per-area density level in the dataset is approximately 9,675 units/km^2. The maximum 3km^2 radius per-area density level in the dataset is approximately 11,511 units/km^2.

substantial economic gains and losses may be linked to the general development patterns of the different market segments.

For example, negative spillover effects of the public rental unit market suggest that there are large, hidden externalities associated with public housing in Hong Kong. Compounded by high costs of rent subsidies, the result could suggest a need for public housing programs to be scaled back or substituted by semi-commercial programs such as HOS. Even if such program expansions necessitate higher levels of average subsidies, the net social outcome could still be positive, especially for high-density residence areas. Furthermore, if there is indeed a shortage of amenities in areas with a prevalence of public rental housing, small amounts of gap-filling funding could have significant positive spillovers in the form of reduced price penalties.

However, possibilities of research are limited without access to public and commercial market data at similar detail levels to the subsidized market dataset used for this paper. In particular, net influence estimates are virtually impossible to derive without setting up a cross-market, quality-adjusted data framework. Future research could focus on collecting commercial and public housing data and comparisons of these markets with subsidized programs. There are also ways in which the current model can be refined, such as evaluating density-price effects with a weighted distance model instead of fixed search radii.

4.4 The Subsidized Housing Hedonic Index

The third and final objective of this paper is the creation of a quality-adjusted, hedonic index for the subsidized housing market in Hong Kong. As a market with entry barriers and practically no overall quality change since 2002, one would not expect a hedonic subsidized housing market index to behave drastically differently from a match-model index or even a simple average-price index. However, short-term noise caused by variations in the quality of observations can be removed by hedonic methods, leading to better estimations of the size of fluctuations and the leading or lagging nature of the market. The nature of housing markets also creates a possibility for quality bias during periods of price fluctuations. For example, during periods of housing price booms and economic growth, disproportionately large numbers of low-quality apartments may be listed for sale by individuals wishing to upgrade their residencies. With a hedonic index, such trends can be well-accounted for.

Creating the hedonic index requires apartments of individual transactions to be adjusted to the quality of a "representative apartment" with averaged characteristics. This is completed by first removing the price influence of all transaction-level and geospatial quality terms evaluated at the level of the respective terms for each transaction. The influence of these terms evaluated at the average level of all 37,745 transactions is then added back to the price. For indicator terms such as district fixed effects, the coefficients are averaged to estimate the price effect of being in a single, representative district. The adjusted prices of observations within each month in the dataset are then averaged to provide an estimation of the monthly housing price levels.

The graphs below describe the hedonic subsidized housing index constructed using this approach. Fig.26 plots the hedonic index along with 95% confidence interval boundaries defined by upper and lower 2-standard-deviation bounds for each month's predicted average price value, with the estimated market price level on August 1997 set as a base level of 100. 1-standard-deviation distance profiles to the predicted values are also provided in Fig.26. Fig.27 compares the unadjusted, average-price index of the subsidized market to the hedonic index, both based on index level = 100 for August 1997.

Fig.26 Hedonic index for HOS secondary market, with 1-SD/2-SD boundaries, Aug 1997 – Jul 2014 (Aug 1997 = 100)[103]

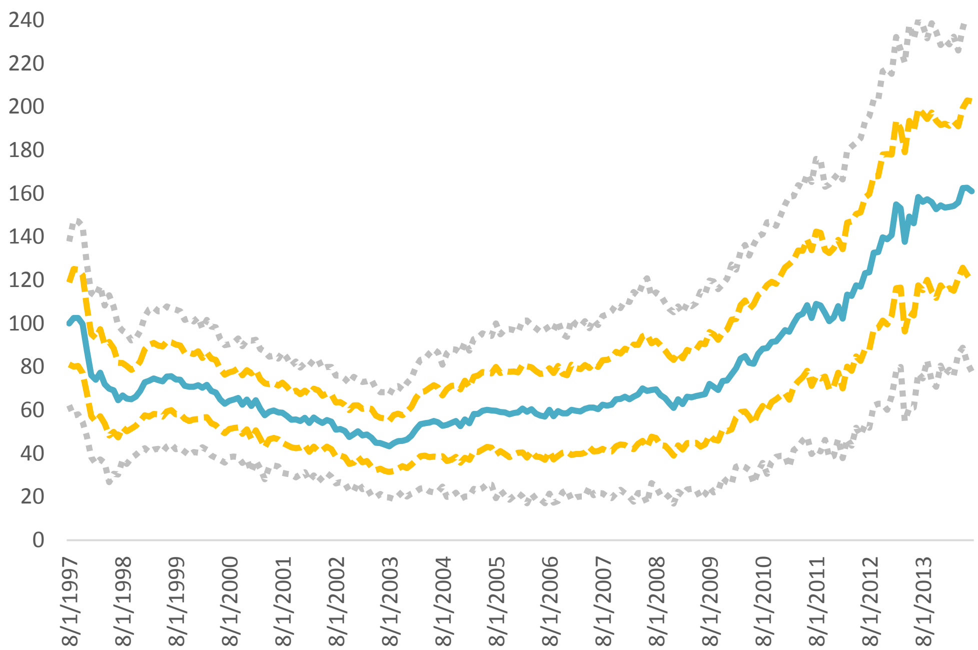
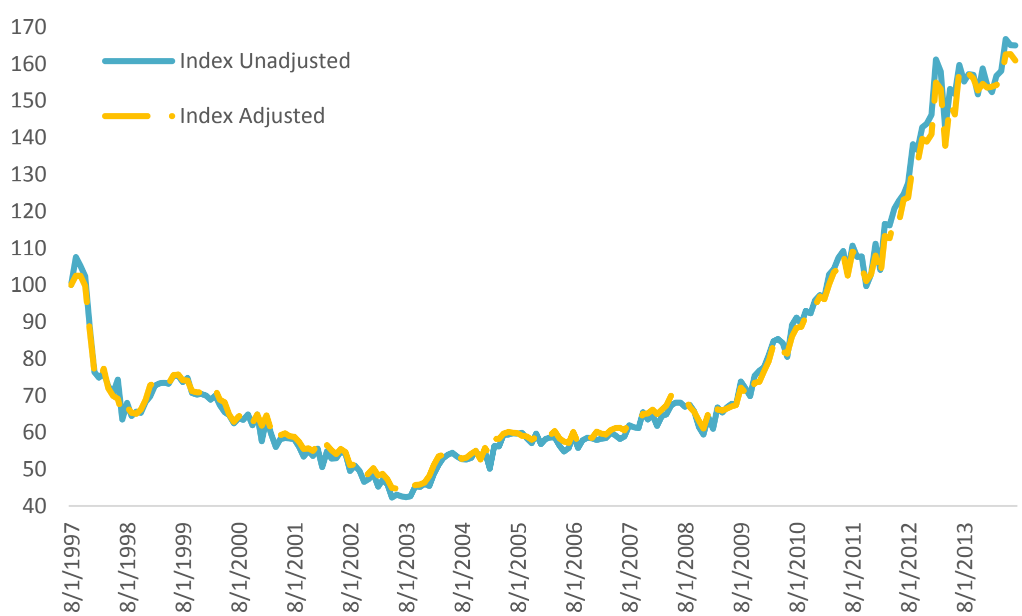


Fig.27 Comparison of adjusted/unadjusted indices for HOS secondary market, Aug 1997 – Jul 2014 (Aug 1997 = 100)

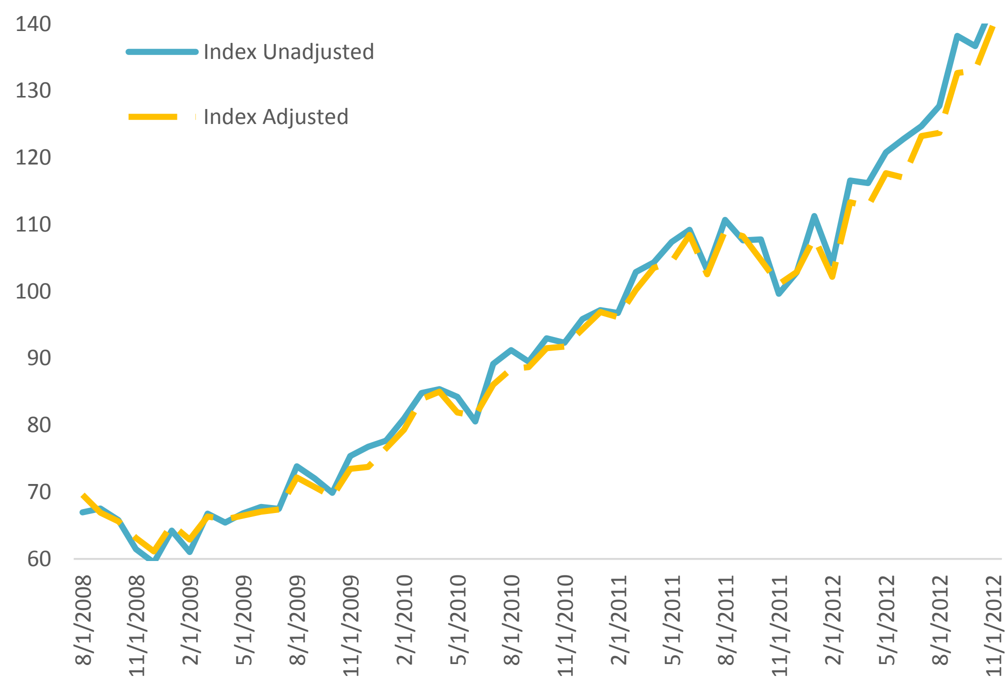


103 Index levels since 2003 can be considered as reflecting trends of the entire subsidized market. Primary market transactions occurring between 1997 and 2002 are not accounted for. New, primary market flats that experienced delays in interior work or did not pass initial quality inspections might have been sold after 2002, though.

To better visualize the smoothing behavior of the hedonic index, Fig.26 displays price trends of a limited 4-year section of both indices between August 2008 and August 2012. While the difference between the two plots is small, the quality-adjusted index shows less variation across observations, with smaller local maximums and larger local minimums in general. From a quantitative perspective, the standard deviation of all indices of the hedonic index is 30.9 compared to 32.4 for the unadjusted index. The average post-quality-adjustment absolute change to the selling price of individual transactions in absolute terms is approximately 13.0%, and the average change to monthly index figures 2.1%.

The difference between these two values can be considered as a rough approximation of the percentage of variation in quality addressable simply by averaging over transaction prices of each month. If we consider the hedonic approach used in this paper robust enough to adjust to near-perfect "representative apartment" price levels, it would imply that around 91.5% of the total quality-induced price variation in the dataset is already removed in the average-price index, with a further 8.5% gain from using the hedonic index. These figures may provide some insight into the accuracy of unadjusted housing price indices of subsidized housing markets in other cities worldwide. However, since housing purchase and rent subsidy markets are usually somewhat homogeneous for a variety of quality factors, most notably house size and age, these estimates might not be good indications of how unadjusted commercial housing market indices perform with regard to inter-transaction quality variance.

Fig.28 Comparison of adjusted/unadjusted indices for HOS secondary market, Aug 2008 – Aug 2012 (Aug 1997 = 100)



While the results do not seem to favor the use of hedonic methods for the Hong Kong subsidized housing market, a 2% improvement to the market-trend describing ability of an index cannot be considered trivial if the goal goes beyond general indication of market trends to specific tests of external shocks or leading/lagging behavior of the market. Applications of such an index are particularly numerous with regard to policy-side inquires of markets with substantial government involvement. A hedonic index could, after adjusting for macro-economic conditions, be used to establish quantitative relationships between regulatory decisions and fluctuations in the price level, or track changes in public perception of such markets.

An in-depth investigation of any such topic is clearly far beyond the scope of this paper. However, I will provide two examples to illustrate the policy-side potential of a subsidized housing hedonic index for the Hong Kong market. After a year of extreme contraction in the real estate sector, the HKHA abruptly cut off land supply to the HOS program in July 1998 as part of a rescue package to increase consumer confidence and stabilize prices. Their efforts did not seem to be effective, and housing price levels in Hong Kong continued to decrease for almost five more years. It has even been suggested that the policy worsened the housing price slump by pushing up prices and generating activity in the subsidized secondary market (Lok, 2000). Literature on the incident describes sellers taking full advantage of the sudden and unexpected boon to liquidate their properties and, as buyers rushed to the subsidized market, sales and prices of commercial units were further depressed.

Fig.29 plots both hedonic and quality-unadjusted indices for the subsidized market against the unadjusted, average-price indices of two sections of the commercial market. All three markets were in a freefall state until Q2 1998, after which they briefly recovered before continuing the downward trend. However, it can be observed that the subsidized market not only experienced the greatest price surge between Q3 1998 and Q1 1999, but also stayed at a higher index level than the other two markets afterwards. Between December 1998 and June 1999 there was a price increase of 9.7% in the subsidized market, whereas the commercial market for small or medium apartments fell by 2.5% and the commercial market for large apartments grew by 1.5%. The size of the “recovery” of 1998-1999, estimated by the maximum price difference between June 1998 and June 1999, is 17.1% for the subsidized market, and only 12.9% and 11.4% for the commercial small or medium and large markets, respectively.

Fig.29 Comparison of housing market indices in Hong Kong, Aug 1997 – Dec 1999 (Aug 1997 = 100)

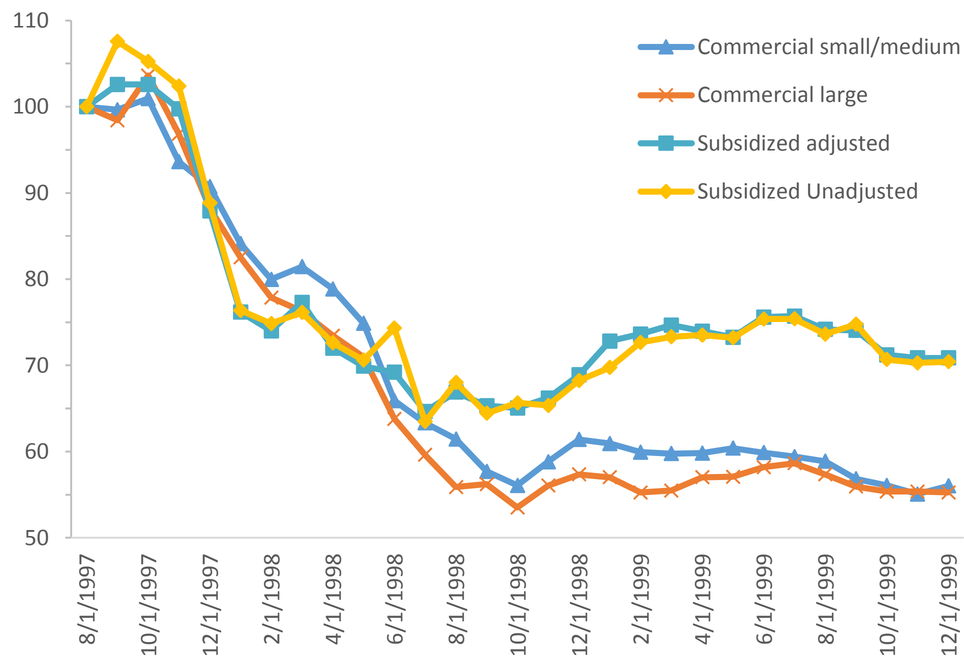


Fig.30 Comparison of housing market 6-month price level changes, Dec 1997 – Jun 2000

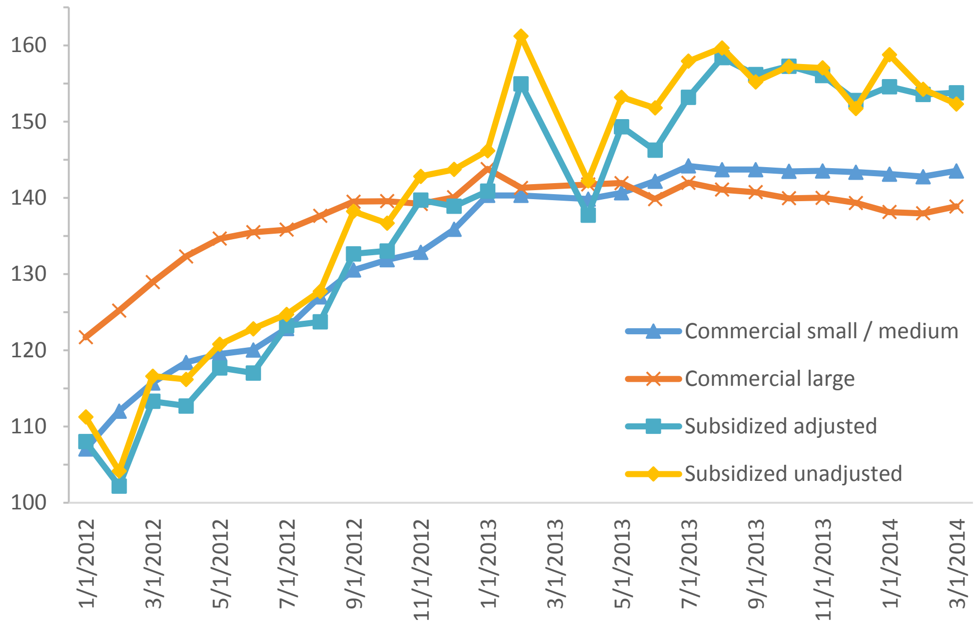
% Changes	Subsidized Adjusted	Subsidized Unadjusted	Small/ medium	Large
(12-1997)				
06-1998	-21.3	-16.3	-27.4	-27.7
12-1998	-0.4	-8.2	-6.8	-10.1
06-1999	9.7	10.4	-2.5	1.5
12-1999	-6.3	-6.6	-6.5	-5.0
06-2000	-8.4	-8.1	-10.5	-5.4

Although this analysis is far from being quantitatively rigorous, a few interesting observations can be made. The first one is that sector-specific policy decisions have the potential to push prices in one sector of the market significantly above that of other sectors. In this case, since the long-term supply halt only occurred in the subsidized market, its price levels were understandably pushed beyond those of the commercial sector. Normally, one would not expect such effects to be persistent, as consumers shift away from consumption in one sector of a market when its relative price increases. However, such mechanisms are likely not effective with regard to the subsidized housing market because of its internal price distortions and barriers to entry.

A second observation is that, given that the baseline effect of the other policy initiatives of the rescue package can be estimated by price level changes to the commercial sector, it is possible that with further, macro-level adjustments, the short-term price effect of an abrupt termination of land supply to the subsidized market can be evaluated. Even if there may be rather large errors associated with a quantitative policy shock estimate, it would still provide a valuable perspective if Hong Kong policy-makers are once again under circumstances unfortunate enough to contemplate such an option. If the estimated shock is large enough, for example, an argument could perhaps be made against the use of such measures in the future.

The second illustrative example comes from the subsidized market deregulation of 2012-2013. A series of policies that relaxed the eligibility criteria of secondary market buyers were announced on November 2012 and enacted on January 2013, the most significant being an increase of the maximum family income ceiling to $40,000 HKD per month from the original, decade-old $30,000 HKD limit. The move has since been criticized as being excessive, allowing high-income individuals to access government subsidies and crowd out less-wealthy families. The CEO of Centaline Properties, Yongqing Shi, wrote on his blog at the time that “I have calculated using income bracket data that some 450,000 extra families in Hong Kong with be able to enjoy subsidized prices with a $40,000 HKD limit,” concluding that “so many relatively high-income families having access to this kind of market is surely undesirable”(Shi, 2012).

Fig.31 Comparison of housing market indices in Hong Kong, Aug 1997 – Dec 1999 (Aug 1997 = 100)



As shown in Fig.31, Shi's criticism seems to be well-founded. The subsidized market experienced an enormous price hike of 11.6% between December 2012 and February 2013. During the same period, prices increased by 3.3% in the small and medium apartment commercial market and 0.9% in the commercial market for large apartments. Clearly, relaxing purchase eligibility standards had resulted in a short-term demand surge of subsidized apartments. However, there is also evidence that HOS secondary market prices have been pushed to higher long-term levels because of the new eligibility requirements. After a period of price volatility in Q1 and Q2 of 2013, the subsidized index has stabilized to levels 5-10% higher than the commercial indices.

This contrasts sharply with figures for early and mid-2012. In the months prior to January 2013, there were virtually no price level disparities between the subsidized and commercial small and medium markets. For 2012, the average 12-month index level of the subsidized market is 121.8 and the average index level of the small and medium apartment commercial market 122.8. For 2013, the respective index averages are 152.6 and 142.9 – a difference of 6.8%. These results suggest that there are fundamental distortions of demand-supply relationships caused by an over-relaxing of HOS eligibility requirements. Compared to short-term price shocks, such long-term distortions are without doubt much more problematic and challenging.

Beyond own-market demand shock effects, it is also possible that commercial sector prices, particularly those of small and medium apartments, are actually suppressed by the relaxation of eligibility requirements. During periods of housing market expansion, the value of HOS apartments increases, yet the general purchasing power of potential HOS apartment buyers, capped by monthly income requirements, stays relatively low. This disparity dis-incentivizes HOS apartment holders from selling their apartments, since few buyers are able to afford them at prices acceptable to sellers. While it is possible to sell apartments on the open market and refund the original subsidy, there might not be much demand on the open, commercial market for apartments at HOS quality levels. Procedures to make refunds and transform subsidized apartments into commercial ones can also be quite complex and time-consuming.

The new policies expand the eligible buyer base into a segment of the population with monthly income levels between $30,000 and $40,000 HKD. When ineligible to participate in HOS, individuals at such income levels could only shop among small and medium commercial apartments. However, once eligible for HOS subsidies, many of them enter the subsidized market where they become unequivocally high-income buyers. Sellers, more than happy to see such buyers, take advantage of the demand surge and begin to list apartments. In the process demand falls for commercial apartments, and downwards price pressure ensues. With inflationary effects in the subsidized market and deflationary effects in the commercial market, an artificial wedge is driven between price levels of the two markets.

Note that for both examples, the quality-adjusted subsidized price index does not perfectly follow the unadjusted index. While the two trend closely to each other, the adjustment effect of the hedonic index is clearly present and highly influential for numeric estimates based on subsidized market price levels. In the 6-month price change figures in Fig.30, estimates from the two indices differ significantly. The quality-adjusted subsidized index decreased by 0.4% between June and December 1998, while the unadjusted index decreased by 8.2% during the same period. With such inaccuracies being a legitimate concern, the quality-unadjusted subsidized index is not meaningful for any type of quantitative analysis. However, once quality-induced noise is eliminated, market behavior can simply be read from changes in the index. Using the hedonic adjustment approach outlined in this section, short and long-term policy-induced price effects, such as those discussed in the examples, can be accurately described, compared and studied in detail. These abilities make possible a wide range of policy-oriented research inquires.

This section discusses the creation of a quality-adjusted hedonic index for the Hong Kong subsidized market, offering two examples of policy effects studied with the aid of such an index. While there are no large market trend changes caused by quality adjustments, the hedonic index is nonetheless a powerful tool for policy analysis. The size, duration and external effects of policy-induced price fluctuations in the subsidized market can be much better understood by reducing noise caused by quality variation across months. Of course, the inquiry methods used in this section can be readily applied to other policy-induced changes in the subsidized market, both past and future. If policies such as eligibility requirements cause quality level shifts in secondary market supply, it would also be interesting to investigate the general quality composition of apartments being sold in different periods in the dataset.

4.5 Broader Implications of Elevation Gradient Effects

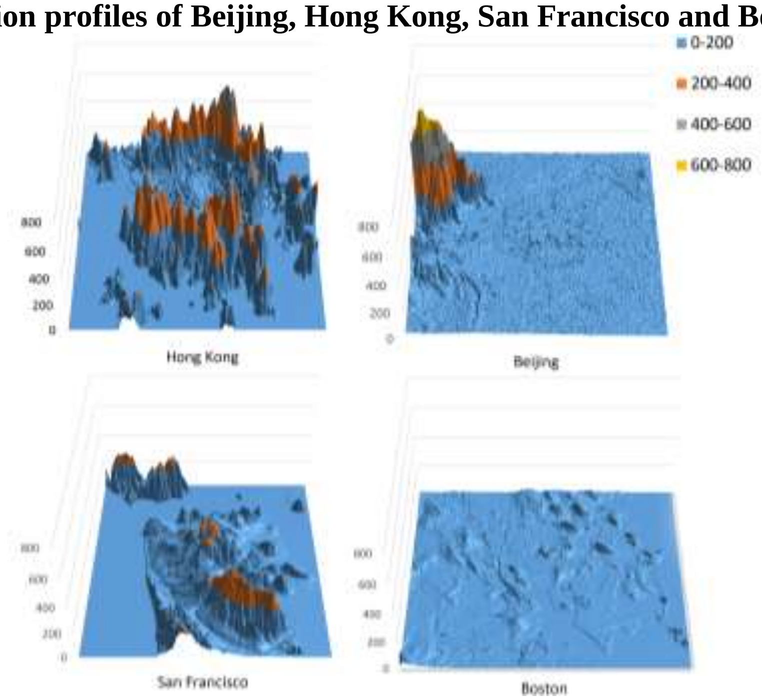
This section returns to the metro network elevation gradient effects observed three sections prior. As demonstrated in section 4.2, people value not only distance to public transit networks but also elevation gradients such as the level of incline between apartments and closest subway stations. The negative price effect of a greater distance to the subway station becomes stronger as the elevation difference increases, most prominently for moderate walking distances and larger elevation gradients.

One would expect that these effects not only apply to subway systems but also to public utilities of cities in general. Certain types of public transportation are no doubt especially susceptible to such effects, but there is no reason to believe that similar preferences do not exist for other amenities such as shopping and educational facilities. Furthermore, assuming the lack of alternative modes of transportation immune to elevation gradient effects, the model presented in this paper suggests that for average quality-level apartments, negative price effects of elevation gradients generally outweigh positive attributes associated with greater altitude. In other words, given that motor vehicle ownership is scarce, individuals seem to strongly prefer living both closer to and on similar elevation level as public utility hubs.

This conclusion has several interesting implications. Firstly, given that major cities are often built on the coastline or next to major waterways, elevation levels and variation of elevation typically increase with distance to the central, downtown area. This means that it is highly likely that there is a positive correlation between densities of cities in general and a measurement of their terrain variance, after controlling for a number of geographic traits. Even if the aforementioned assumption does not hold, the *variation* of population density gradients in cities should be linked to elevation variance: the more uneven a city's terrain is, the more likely it is to have extreme high-density and low-density living areas.

It is possible to create models to test these hypotheses. Using data from the Microsoft Bing Maps REST Services, highly accurate elevation profiles of virtually all cities in the world can be generated using a fixed-pitch matrix sampling method over rectangular pieces of land. Points could then be randomly sampled from these profiles and used to create variance estimations that track terrain unevenness. The same profiles can also be used to control for the amount of water within the centered area-of-influence of a given city, hence removing the obvious influence of relative land coverage on population density.[104] Since elevation estimates are specific to longitude/latitude combinations, they can also be directly matched against population density or income gradient profiles of different cities.

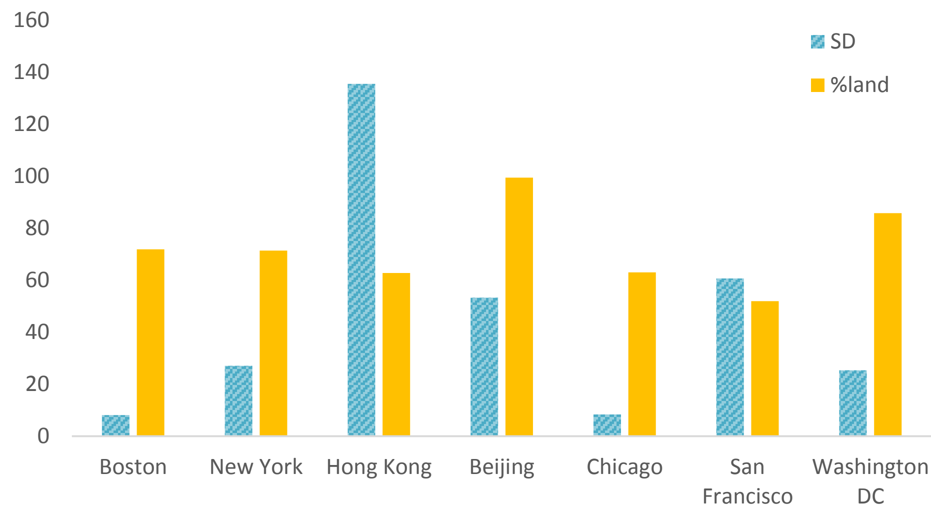
Fig.32 Elevation profiles of Beijing, Hong Kong, San Francisco and Boston (meters)



[104] This is achieved by counting all observations with elevation equal to the local water level (sea level for coastline cities), and then obtaining the ratio of observations above that value to the total number of observations. The results is a percentage estimate of total land coverage.

As an example, I have completed elevation profiles for Hong Kong, Beijing, San Francisco and Boston (Fig.32) and summarized the analysis results in Fig.33. Elevation sample points are selected over a square area that encapsulates the entirety of city limits or, for Hong Kong's case, the island and densely populated districts on the peninsula. For each city, a grid of 201*201, or 40401 elevation observations, are taken from a 10m or 90m-level resolution base dataset.[105] It must be noted that this method is not very accurate, since profiles may not match actual city boundaries and could omit regions that are in fact within a city's economic and policy influence. Circular or even asymmetric, boundary-defined profiles could perhaps be applied in future research to partially address this issue.

Fig.33 Terrain standard deviation and land coverage percentage estimates for sample US and Chinese cities



The SD measurement in Fig.33 is simply the standard deviation of a 10,000-point sample of land observations within the area of each city profile. A city with a greater number of particularly large or small elevation values will report a larger SD than a city where terrain observations are of uniform altitude. Also displayed are the relative land ratios of the seven cities, denoted in percentage terms. Mostly landlocked cities such as Washington DC and Beijing have land ratios close to 100%, whereas ports such as Boston, New York and San Francisco consist only of 50-70% land. Note that Hong Kong shows by far the greatest amount of terrain variance among the listed cities, reporting a standard deviation of 135.5, more than twice that of San Francisco (60.6) and about five times that of New York (27.0).

However, one could object that a single standard deviation estimation cannot capture in full the characteristics of terrain variation. Just as income segregation takes form at both the city level and the local, district level, there is a distinction between variance caused by large, city-scale, relatively uniform terrain formations and variance caused by numerous, small and steep hills. The distinction is best exemplified by the difference between Beijing and San Francisco in the dataset. Both cities report roughly similar standard deviation values, yet Beijing sits on an almost perfectly flat terrain with one major mountain range in the west part of the city. San Francisco, on the other hand, has dozens of small hills broken up by lakes and the ocean. One would expect the two types of terrain

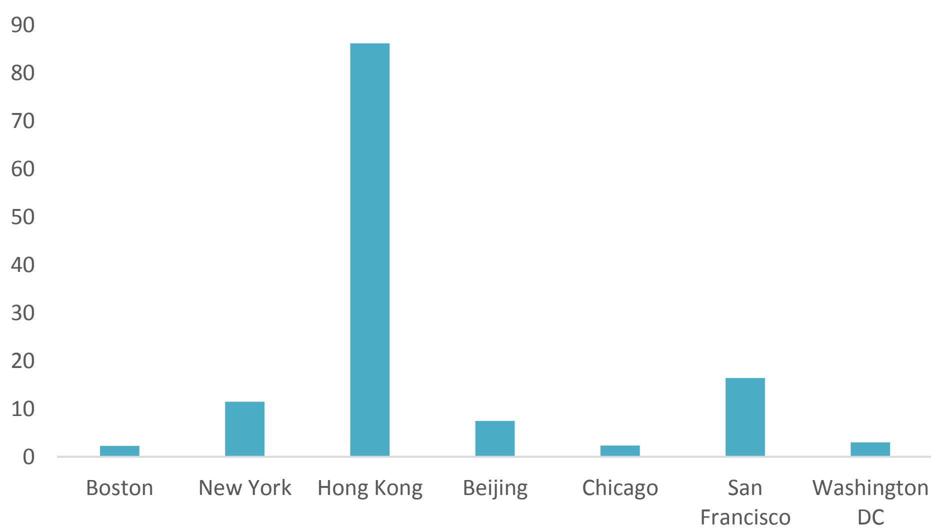
105 The Bing Maps data has 10-meter resolution for the US and 90-meter resolution for the rest of the world, excluding Polar Regions. The grid size is selected to not saturate either dataset, with distances between observations being close to 100m for the Beijing and Hong Kong profiles.

to not have the same impact on population density gradients, an effect that the standard deviation of elevation fails to account for.

To deal with this issue, a second measurement of elevation gradients is introduced. The "land variance indicator" shown in Fig.34 measures the general "unevenness" of a given area by summing over the absolute differences between all points in an elevation profile across latitude and longitude. In other words, the elevation difference between each point and its four direct neighbors is aggregated for all sample points in the profile, with each between-point difference only being counted once. The result is then normalized through dividing by the total number of observations for each city, 40,401 in this case.

The intuition behind this measurement is fairly straightforward. If the plane where the observations take place is perfectly flat, the variance indicator reports a value of 0. If each observation reports an elevation change of exactly 1m – the example being a perfectly uniform slope over a square area with a slope of 1m per unit distance over both latitude and longitude – the variance indicator reports a baseline value of 2. By definition, if the same set of elevation values is arranged in any other configuration, the indicator will always be greater than or equal to 2. The more "jagged" the terrain is, with large elevation values right next to small ones, the greater the indicator value is. Therefore, a city with numerous, small and relatively steep hills will report a much greater land variance indicator value than a city with a similar standard deviation value but only has a single tall mountain.

Fig.34 Land variance indicator estimates for sample US and Chinese cities



From Fig.34, it is not difficult to observe how much of an outlier Hong Kong is compared to other, well-known US cities and Beijing. Not only does it have a robust amount of overall elevation variance, the terrain is constructed with an extremely high level of local variance. The island itself has a peak of 554m, and the peninsula has a mountain with a peak of nearly a kilometer in altitude situated less than 5 kilometers from the coast line.[106] Such levels of terrain steepness are certainly rare among major cities. In light of this observation, the large and significant elevation gradient effects associated with distances to the Hong Kong MTR network seem quite reasonable. On the

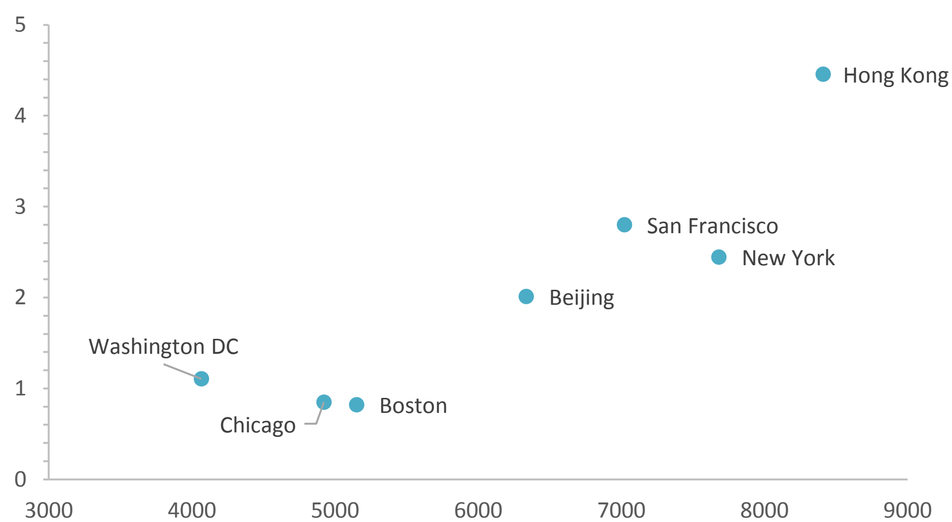
[106] The Taiping Mountain is the tallest peak on Hong Kong Island measuring in at 554m from sea level. The tallest mountain in the HKSAR is Damao, which is 957m from sea level.

other hand, one would not expect such elevation effects to be nearly as large for the majority of cities worldwide.

Note that the land variance indicator estimate is not independent of overall elevation variance. Cities with a higher standard deviation of elevation will, by definition, have a larger land variance indicator value. While this may not necessarily be undesirable from an investigative perspective, it may still be beneficial to use a variable that is more robust to the standard deviation of elevation samples. One example of such an option could be an approach similar to the "spatial ordering index" used in analyzing income distribution patterns (Dawkins, 2007). There is a large body of economic literature on the measurements of spatial disparities in general, which would no doubt be informative in the process of creating a good estimate for terrain unevenness.

For the seven cities in Fig.34, population density estimates are created with regard to the respective sampled areas.[107] Fig. 35 scatter plots the density estimates against a log transformation of the land variance indicator. As can be observed, cities with greater terrain variance have on average greater population densities. The three cities with the greatest population density – Hong Kong, San Francisco and New York City – are also the three cities with the largest land variance indicator values. Both being older, established east coast ports, New York City has both significantly greater local elevation variance and population density compared to Boston. Similarly, Beijing does not quite match Hong Kong in terms of population density. While the estimates used in this analysis are fairly rough, they do provide some support for the idea that there is a positive connection between population density and the terrain patterns of cities. Future work could involve the optimizing of this procedure, as well as applying similar methods of investigation to a larger body of global cities.

Fig.35 Scatter plot of estimated population density (ppl/km^2) over land variance indicator, with natural log transformation, of US and Chinese cities



107 Data for Chicago, Washington DC, San Francisco and Boston are city population estimates for 2011 divided by the estimated land area of the sampled region of Fig.34. Data for Beijing only includes the 2012 population of the six "urban" districts (Dongcheng, Xicheng, Chaoyang, Fengtai, Shijingshan and Haidian). Data for New York includes the population of Newark. Other counties of New York state and New Jersey within the sampled square are not accounted for, since their populations are most likely trivial compared to that of New York City. Data for the Hong Kong population is for the year 2011 and omits the Islands District and North District, both of which are not covered by the sample area. Sources: United States Census Bureau (USCB), 2012 Beijing Statistical Yearbook, NYC Department of City Planning, Hong Kong Census and Statistics Department (HKCenStat).

Given that Hong Kong's elevation variance can be quantitatively compared to that of other cities, it may even be possible to roughly gauge the amount of overall price influence associated with elevation gradient effects for a given housing market in any major city. Since the large effects observed in Hong Kong are the result of an extraordinarily great amount of elevation variance, one could speculate that there is a point where a certain city is "flat" enough that the effort to include elevation gradients becomes essentially unnecessary. For cities with greater levels of elevation variance, an argument could be made for using the type of adjustments that I have employed in this paper. For cities with less than that amount, the argument is weaker and elevation gradient effects in general can perhaps be safely disregarded.

The estimations would be crude at best, but it might be possible to at least provide a baseline for policy-makers and real estate developers charged with the regulation of housing markets and valuation of pieces of property. For example, one could argue that any major city with elevation variance levels comparable to or greater than those of Hong Kong ought to be modeled with elevation effects taken into account. However, the question as to which cities actually need such modelling techniques is one that is beyond the scope of this paper. Future research searching for evidence of elevation effects in cities with less terrain variance could provide greater insight into this issue. Such investigations can perhaps be as simple as adding a number of elevation adjustment variables to established research models involving housing market hedonics.

A third implication of elevation gradient effects is the connection between income segregation and elevation gradient effects. As is observed in cities such as San Francisco and Hong Kong, individuals with relatively large amounts of disposable income can choose to simply not use public transit and drive around the city. Since driving as a method of transportation is largely immune to elevation gradient effects, people who have already decided that they will be driving to work or school should always choose to live at higher altitudes and avoid the price premium of living at the same plane of elevation as bus stops and metro stations.

The effect works both ways. High-income individuals crowding the high-altitude housing market will most likely bid up prices for such pieces of property, further decreasing the willingness of low-income individuals to live at higher elevation levels. The result is an exacerbation of spatial income segregation for locations with greater local variance of elevation. In particular, places with steep hills and little "middle ground" can essentially be modelled as a two-good market, with high-income individuals and low-income individuals almost unanimously selecting different markets. If the difference in elevation is great enough, the two markets are distinct and, theoretically, spatial income segregation becomes absolute.

Translating into real-world terms, the finding suggest that for a number of cities or even within the same city, regions or districts with greater local elevation variance will fare worse in terms of spatial income distribution. Areas with high elevation levels are more likely to become "rich" communities that are priced out of reach for the majority, while low-elevation areas are more likely to have large low-income communities and, as an extension, issues with the lack of infrastructure and crime. This is perhaps particularly true for low-elevation area further away from the downtown area, where alternative modes of transportation generally incur high costs.

With detailed data on district-level income and spatial relationships, as well as further research, I am optimistic that all three of these issues can be sufficiently addressed. The issue of elevation gradient effects is perhaps a niche in the study of real estate systems, yet it should not be overlooked. The model of the Hong Kong subsidized market suggests that elevation effects on price are potentially no less robust than that of other, well-investigated factors such as distance and density gradient effects. A more in-depth look at these effects, either for individual cities or urban systems in general, is therefore at least somewhat worthwhile.

Section V:

5.1 Conclusion

This paper provides, to date, one of the most detailed examinations of a purchase-subsidy housing market in the urban economics tradition. Extensive geospatial analysis methods, including the use of ArcGIS, Bing Maps REST services and Google Maps API, are combined with a time dummy hedonic regression approach to create a highly robust model for the subsidized, semi-commercial Hong Kong housing market. Approximately 93% of the total variance in prices between transactions can be explained by the regression model, allowing for detailed investigations of time period-related interaction effects and public perception effects.

Noteworthy results include significant season, size, age, floor level and elevation-related effects on the selling price of a given apartment. Having the lucky number 8 in the address of an apartment leads to a statistically significant price premium, while the presence of unlucky numbers 4 and 13 has an even larger, negative impact on price. There are also price effects associated with the distance and time costs of commuting to the CBD, as well as significant effects related to distances to the coastline and airport. I am hopeful that the quantification of the influences of such factors in apartment selling prices will lead to optimized planning and better regulatory practices of housing agencies in Hong Kong.

The results suggest that for low total apartment density levels by Hong Kong standards, effects of higher concentrations of subsidized housing are ambiguous. For higher-than-average total density levels, higher proportions of HOS apartments have a clear positive price effect on subsidized apartment prices. The public rental unit market has clear, negative density-price effects at all except the lowest total density levels. This contrast suggests potential merits of home ownership over rental subsidization, although controls for the rates of subsidies on the two markets must be employed for more conclusive results. High concentrations, both relative and absolute, of apartments from the commercial Hong Kong housing sector generate significant positive price effects at all residential density levels.

There is strong evidence of elevation gradient effects in the relationship between subway system availability and selling prices of apartments. Negative price effects of greater walking distances become significantly more severe when elevation differences are involved. The effects are strongest for moderate walking distances and grow in severity with the level of average incline. These observations not only are in line with the intuition of individuals disliking uphill or downhill walks but also suggest the presence of similar effects for other major cities. However, Hong Kong might be particularly susceptible to elevation gradient effects, in part because of the extreme levels of terrain variance and in part because of a general reliance on public transit networks.

This paper discusses several implications of the existence of elevation gradient effects, including a positive relationship between terrain unevenness and population density as well as associations between local elevation variance and income segregation. The testing of such hypotheses could be an interesting addition to the urban economics literature. At the very least, the results presented here highlight the necessity of incorporating elevation data into models analyzing urban systems, housing or otherwise, for cities with large amounts of elevation variation.

The determining of exactly how much terrain unevenness requires the use of such measures may in itself be an interesting question.

The final contribution of this paper is a hedonic index of the Hong Kong subsidized housing market which contains all secondary-market transactions between August 1997 and July 2014. In its current form, the index controls for the vast majority of quality-induced price fluctuations of the market, opening up possibilities for quantitative examinations of topics such as policy-induced shock and long-term interactions among the various market sectors in Hong Kong. Two examples are selected from 1999 and 2013, respectively, to demonstrate the versatility of the index with regard to policy-side applications. The index is no doubt a powerful tool for local policy-makers, one with the potential to modernize regulatory approaches and aid the future development of housing subsidy programs in Hong Kong.

As of this writing, the opening of the first new HOS courts of the 2012 program reboot is close to a year away.[108] For a housing market surrounded by controversy and drama, it has, not until now, been subjected to scrutiny. In this regard, I have done my best to go beyond the qualitative and provide a data-backed, comprehensive examination of its characteristics and behaviors. I have also created tools to explore this market, including geospatial analysis packages and a hedonic subsidized housing price index. These tools are capable of furthering our understanding of issues such as tastes and preferences of individuals who access housing subsidies, incentives and disincentives such markets provide, and the role of the purchase-subsidy market among other housing sectors. Through answering these questions, I hope to accomplish a more balanced and nuanced evaluation of the subsidized market's merits and detriments.

Moving beyond the local, Hong Kong-oriented perspective, I have considered ways in which observations about this particular market raise questions of a broader nature and appeal. The existence of housing subsidies in various forms, their associated price spillovers and elevation gradient effects all play a role in property markets of most, if not all major cities worldwide. The methods involved in this paper, such as elevation adjustments and geospatial modelling, can therefore be readily applied to the investigating of other urban societies. To the extent which basic individual preferences converge worldwide and similarities abound what we view as "urban," much of the findings here could, at the very least, serve as a baseline for the analyses of comparable aspects of other housing markets.

[108] The first new HOS reboot apartment group, Mei Pak court, is scheduled for completion in early 2016. Source: http://warmhomehk.com/NewCourt.aspx

References

Autor, D. H., Palmer, C. J., & Pathak, P. A. (2014). Housing Market Spillovers: Evidence from the End of Rent Control in Cambridge, Massachusetts. *Journal of Political Economy, 122*(3).

Beijing Municipal Bureau of Statistics. (2013). *Beijing Statistical Yearbook 2013*. Beijing: Retrieved from http://www.bjstats.gov.cn/nj/main/2013-tjnj/index.htm.

Cebula, R. J. (2009). The Hedonic Pricing Model Applied to the Housing Market of the City of Savannah and Its Savannah Historic Landmark District. *The Review of Regional Studies, 39*(1), 13.

Chen, L., Zhao, S., Romanis, P., & Lim, P. P. (2004). *Exploring Hedonic Methods for Constructing a House Price Index*. Canberra: Australian Bureau of Statistics.

Chow, W. (2011). *Hedonic Price Index: An Illustration with Residential Property Prices*. HKMA: Economic Analysis and Business Facilitation Unit.

Dawkins, C. (2007). Space and the measurement of income segregation. *Journal of Regional Science, 47*(2), 17.

Diewert, E. (2003). Hedonic Regressions. A Consumer Theory Approach. In R. C. Feenstra & M. D. Shapiro (Eds.), *Scanner Data and Price Indexes* (pp. 317-348). Chicago: University of Chicago Press.

Er, D., & Li, J. (2008). *Home Ownership Courts of Hong Kong: An Exploration* Hong Kong: Ming Bao Publishing

Goodman, A. C. (1978). Hedonic Prices, Price Indices and Housing Markets. *Journal of Urban Economics, 5*(1), 14.

Haan, J. d. (2003). *Hedonic Regression: The Time Dummy Index As a Special Case of the Imputation Tˆrnqvist Index* Paper presented at the International Working Group on Price Indices, Paris.

Haizhen, W., Shenghua, J., & Xiaoyu, G. (2005). Hedonic price analysis of urban housing: An empirical research on Hangzhou, China. *Journal of Zhejiang University, 6*(8), 7.

Harrison, D., & Rubinfeld, D. L. (1978). Hedonic Housing Prices and the Demand for Clean Air. *Journal of Environmental Economics and Management, 5*(1), 21.

Haurin, D. R., Parcel, T. L., & Haurin, R. J. (2002). Does Homeownership AffectChild Outcomes? *Real Estate Economics, 30*(4), 31.

Hayes, K. J., & Taylor, L. L. (1996). Neighborhood school characteristics: What signals quality to ho-mebuyers. *Federal Reserve Bank of Dallas Economic Review, 4th Quarter*, 7.

Hill, R. (2011). Hedonic Price Indexes for Housing. *OECD Statistics Working Papers*.

Hong Kong Census and Statistics Department. (2011). *2011 Hong Kong Population Census*. CenStat Retrieved from http://www.census2011.gov.hk/en/index.html.

Hong Kong Housing Authority. (2014). *Housing Statistics 2014*. Hong Kong: Hong Kong Housing Authority.

Hong Kong Legislative Council. (2005). *Summary Statistics: Public Housing in Hong Kong*. Hong Kong Hong Kong Legislative Council.

Hong Kong Ratings & Valuations Department. (2013). Technical Notes: Hong Kong Property Price Indicies from http://www.rvd.gov.hk/doc/en/statistics/15_technotes.pdf

Huang, H. (1999). *Scandels of Public Housing in Hong Kong: A Reporter's In-depth Investigation*. Hong Kong: Jin Yi Bu Publishing.

Liu, G. (2001). *Housing Statistics in Hong Kong: An Overview*. Hong Kong: Hong Kong City University.

Liu, G. (2003). *25 Years of HOS: Changes and Developments* Paper presented at the Housing Development Forum 2003, Shenzhen.

Lok, S. (2000). *A Case Study of Economic Ecology: The Hong Kong Economy's Plunge into a Deep Recession in 1998.* Lingnan University.

Malpezzi, S. (2003). Hedonic Pricing Models: A Selective and Applied Review. In A. OSullivan & K. Gibb (Eds.), *Housing Economics: Essays in Honor of Duncan Maclennan*. Malden, MA: Blackwell.

McMillen, D. (2004). Airport Expansions and Property Values: The Case of Chicago OHare Airport. *Journal of Urban Economics, 55*.

Melser, D. (2005). The Hedonic Regression Time-Dummy Method and the Monotonicity Axioms. *Journal of Business & Economic Statistics, 23*(4), 7.

New York City Department of City Planning. (2014). Current Population Estimates. *Population.* from http://www.nyc.gov/html/dcp/html/census/popcur.shtml

Ni, J., & Decker, C. (2009). The Impact of Homeownership on Criminal Activity: Empirical Evidence from United States' County Level Data. *Economics & business Journal: Inquiries & Perspectives, 2*(1), 21.

Ottensmann, J. R., Payton, S., & Man, J. (2008). Urban Location and Housing Prices within a Hedonic Model. *Journal of Regional Analysis and Policy, 38*(1), 16.

Pakes, A. (2003). *A Reconsideration of Hedonic Price Indexes with an Application to PC's*. Harvard University.

Rohe, W. M., & Stewart, L. S. (1996). Homeownership and Neighborhood Stability. *Housing Policy Debate, 7*(1), 46.

Rosen, S. (1974). Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition. *Journal of Political Economy, 82*(1), 21.

Shi, Y. (2012). The New HOS Income Eligibility Limit is Too High (Ju Wu Ru Xi Shang Xian Si Wan Guo Gao).

Shiller, R. J. (2008). *Derivatives Markets for Home Prices*. Cowles Foundation Discussion Paper, (1630).

Sliver, M., & Heravi, S. (2006). The Difference Between Hedonic Imputation Indexes and Time Dummy Hedonic Indexes. *IMF Working Papers*.

Smart, A. (2006). *The Shek Kip Mei Myth: Squatters, Fires and Colonial Rule in Hong Kong, 1950-1963*. Hong Kong: Hong Kong University Press.

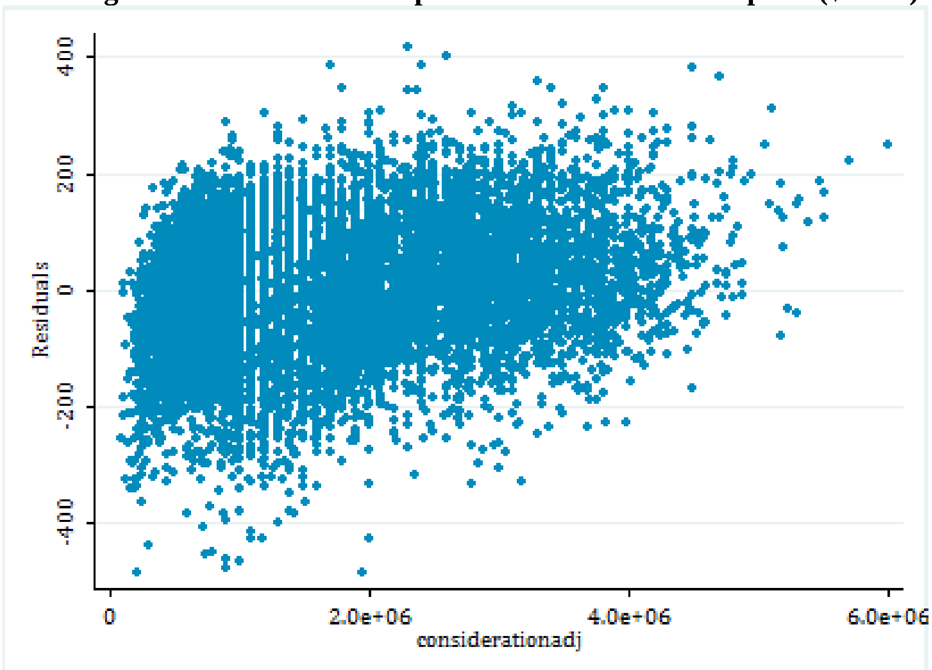
U.S. Census Bureau Population Division. (2012). *Table 1. Annual Estimates of the Population of Metropolitan and Micropolitan Statistical Areas: April 1, 2010 to July 1, 2011*. Retrieved from: http://www.census.gov/popest/data/metro/totals/2011/tables/CBSA-EST2011-01.xls

Zhao, L. (2005). *Where will Hong Kong's Housing Policy Go?* Hong Kong: Hong Kong University Department of Urban Planning and Environment Management.

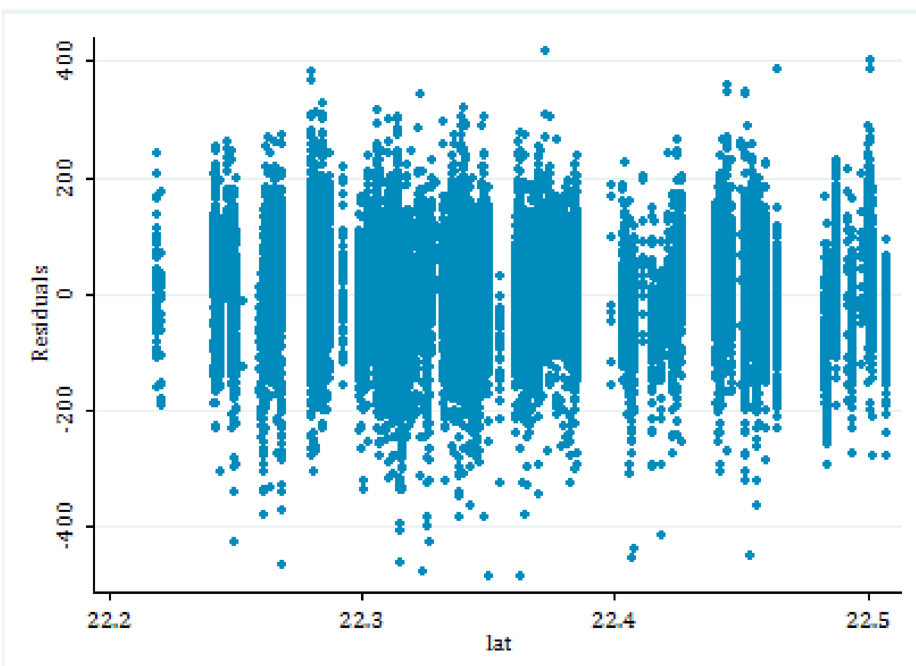
Zou, Y. (2013). *Contradictions in China's affordable housing policy: Goals vs. structure.* Temple University, Philadelphia

Appendix

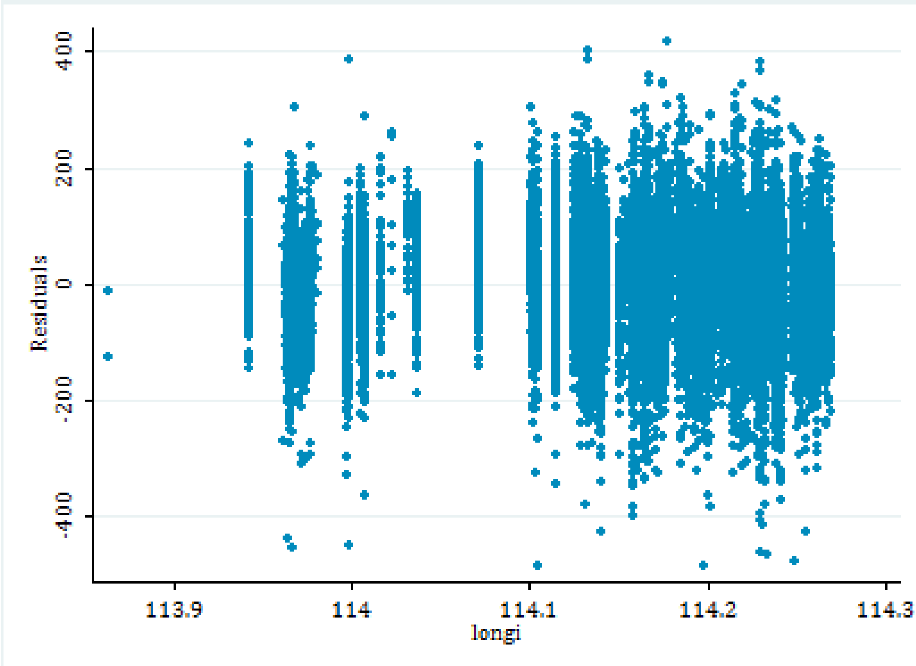
1. Regression residual scatter plot over total transaction price ($HKD)



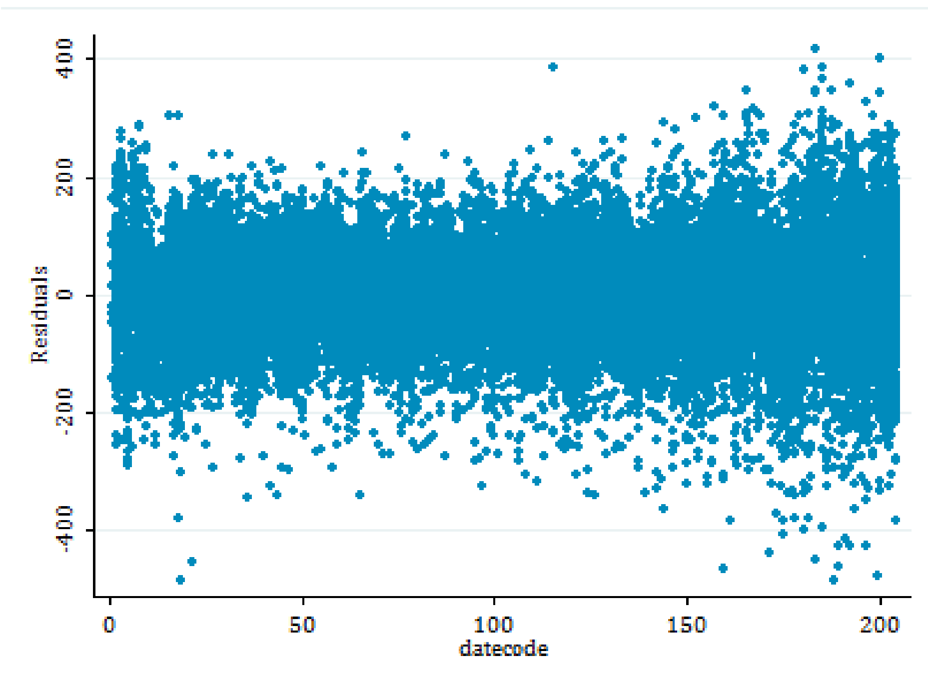
2. Regression residual scatter plot over latitude (decimal)



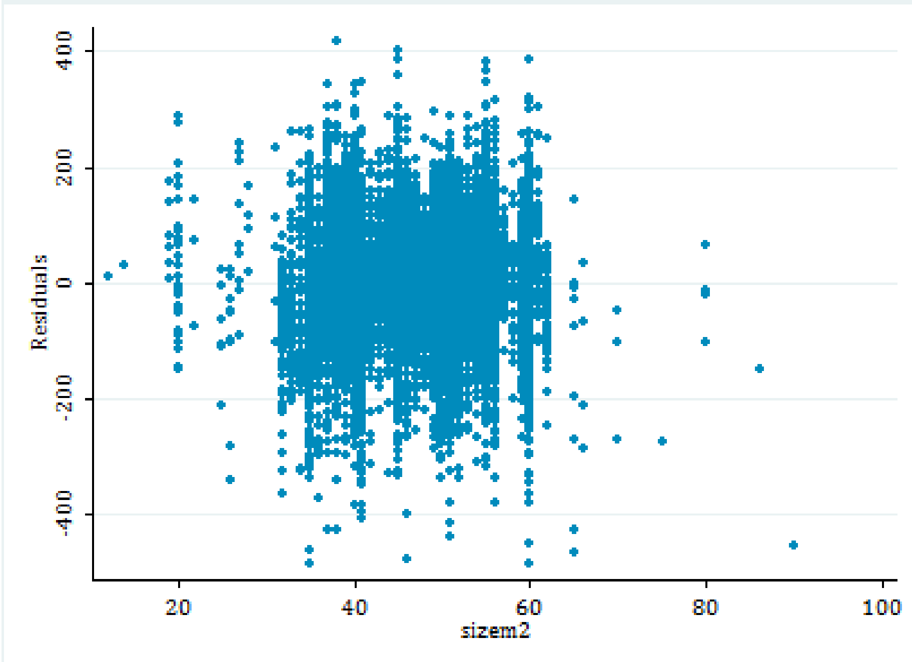
3. Regression residual scatter plot over longitude (decimal)



4. Regression residual scatter plot over transaction date (Aug 1997 = 1)



5. Regression residual scatter plot over size of apartment (m^2)



6. List of variables and variable names used in regression model

feb	Dummy for transaction being in February
mar	Dummy for transaction being in March
apr	Dummy for transaction being in April
may	Dummy for transaction being in May
jun	Dummy for transaction being in June
jul	Dummy for transaction being in July
aug	Dummy for transaction being in August
sep	Dummy for transaction being in September
oct	Dummy for transaction being in October
nov	Dummy for transaction being in November
dec	Dummy for transaction being in December
yd1998	Year Dummy 1998
yd1999	Year Dummy 1999
yd2000	Year Dummy 2000
yd2001	Year Dummy 2001
yd2002	Year Dummy 2002

yd2003	Year Dummy 2003
yd2004	Year Dummy 2004
yd2005	Year Dummy 2005
yd2006	Year Dummy 2006
yd2007	Year Dummy 2007
yd2008	Year Dummy 2
yd2009	Year Dummy 2009
yd2010	Year Dummy 2010
yd2011	Year Dummy 2011
yd2012	Year Dummy 2012
yd2013	Year Dummy 2013
yd2014	Year Dummy 2014
age	Age of apartment in year (with higher order terms in regression)
lgsize	Size of apartment in m^2, with natural log transformation
floorM	14-26 floors in height
floorH	>27 floors in height
discountrate	Amount which apartment is subsidized
unluckynum	Apartment with number 4 or 13 in address
luckynum	Apartment with number 8 in address
lgdtonosec	Distance to nearest major road, defined as highways or four lanes and above, with natural log transformation
lgdtocentralcar	Logged driving distance to central Hong Kong by automobile
INTperioddtocentral	Time period of transaction in months interacted with Logged driving distance to central (with higher order terms)
lgttocentralcar	Logged driving time to central Hong Kong by automobile
INTperiodttocentcar	Time period of transaction in months interacted with Logged driving time to central (with higher order terms)
lgttocentralpublic	Logged public transit time to central Hong Kong
INTperiodttocentpub	Time period of transaction in months interacted with Logged time by public transit (with higher order terms)
lgincome	Logged income of district where apartment is located in
INTincomeborder[109]	Distance from apartment to district border interacted with logged income level of district where apartment is located in

[109] The inclusion of a district border distance-income interaction is meant to introduce some degree of nuance into district-level income figures. The link could be explained by the way modern Hong Kong districts are established: they do not serve many practical purposes but are often drawn with borders in places with least population or human activity.

lgelev	Logged elevation level of apartment, meters
lgcoast	Logged distance to coast, meters
CountDSS	Number of DSS middle schools in a 3km radius of the apartment
INTperiodDSS	Time period of transaction interacted with number of DSS middle schools in a 3km radius of apartment (with higher order terms)
lgWalkElemt	Logged walking distance to nearest elementary school, meters
INTperiodelementary	Time period of transaction in months interacted with logged walking distance to nearest elementary school (with higher order terms)
SlopeElementary	Average slope between apartment and nearest elementary school, degrees
INTslopewalkelem	Average slope between apartment and nearest elementary school interacted with logged walking distance to nearest elementary school
lgMTRonehalf	Logged walking distance, in best fit equivalent estimation, to nearest metro station, meters
WalkMTRangle	Average slope between apartment and nearest metro station, degrees
INTMTRcompslope	Average slope between apartment and nearest metro station interacted with logged walking distance to nearest metro station
dtoairportkm	Distance to nearest airport, kilometers
courtsize	Size of apartment group, units
dcodeX	Dummy for being in district X, with 15 districts in total
ncode190	Dummy for being in court “Kornhill”
ncode94	Dummy for being in court “Tun Yuk Court”
ncode33	Dummy for being in court “Yu Shing Court”
intcoastelev[110]	Interaction between logged distance to coastline (meters) and logged elevation (meters)
intcoastfloorH	Interaction between logged distance to coastline (meters) and indicator for high floor levels

This means that in general, regions with greater amounts of commercial activity, and hence higher average income and living costs, will be located closer to the center of districts. In the regression, this term is positive (coefficient = 12.4) and significant at the 99.9% level (P≈0.0009).

110 The following five variables (intcoastelev, intcoastfloorH, intcoastfloorM, inteleveflooorH, intelevfloorM) are not part of the basic time dummy hedonic regression. They are only used in the analysis for coastline distance effects for different elevation and floor levels.

intcoastfloorM	Interaction between logged distance to coastline (meters) and indicator for medium floor levels
intelevfloorH	Interaction between logged elevation (meters) and indicator for high floor levels
intelevfloorM	Interaction between logged elevation (meters) and indicator for medium floor levels

7. Regression output, original regression model

Source	SS	df	MS
Model	2.8845e+09	88	32778673
Residual	217163980	37656	5767.04855
Total	3.1017e+09	37744	82176.9607

Number of obs	= 37745
F(88, 37656)	= 5683.79
Prob > F	= 0.0000
R-squared	= 0.9300
Adj R-squared	= 0.9298
Root MSE	= 75.941

priceadjtwothird	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
feb	6.267646	2.064175	3.04	0.002	2.221808	10.31349
mar	18.60077	1.964897	9.47	0.000	14.74952	22.45202
apr	26.46751	2.035945	13.00	0.000	22.477	30.45802
may	33.90122	1.986659	17.06	0.000	30.00732	37.79513
jun	44.47692	2.09937	21.19	0.000	40.36209	48.59174
jul	52.17508	2.179748	23.94	0.000	47.90272	56.44745
aug	59.74368	2.293789	26.05	0.000	55.24779	64.23957
sep	66.81224	2.399469	27.84	0.000	62.10922	71.51526
oct	71.30201	2.496199	28.56	0.000	66.4094	76.19463
nov	74.33912	2.651315	28.04	0.000	69.14247	79.53577
dec	83.07476	2.796911	29.70	0.000	77.59274	88.55678
yd1998	-88.83367	5.583359	-15.91	0.000	-99.77721	-77.89014
yd1999	63.56388	8.138634	7.81	0.000	47.61194	79.51583
yd2000	156.5645	10.91939	14.34	0.000	135.1622	177.9668
yd2001	221.8377	13.25151	16.74	0.000	195.8644	247.811
yd2002	293.3	15.18029	19.32	0.000	263.5462	323.0538
yd2003	348.1088	16.85266	20.66	0.000	315.0771	381.1405
yd2004	508.5509	18.4371	27.58	0.000	472.4137	544.6882
yd2005	644.8428	19.93569	32.35	0.000	605.7683	683.9172
yd2006	733.6841	21.43503	34.23	0.000	691.6708	775.6973
yd2007	828.8928	22.88499	36.22	0.000	784.0376	873.748
yd2008	944.3168	24.39514	38.71	0.000	896.5017	992.1319
yd2009	992.4911	25.84873	38.40	0.000	941.8269	1043.155
yd2010	1149.702	27.36571	42.01	0.000	1096.065	1203.34
yd2011	1310.752	29.01756	45.17	0.000	1253.877	1367.627
yd2012	1419.257	30.78148	46.11	0.000	1358.924	1479.589
yd2013	1609.623	32.97017	48.82	0.000	1545	1674.245
yd2014	1629.452	35.46912	45.94	0.000	1559.932	1698.972
age	98.76466	11.9553	8.26	0.000	75.33195	122.1974
sqrage	-24.47612	2.402262	-10.19	0.000	-29.18462	-19.76763
age3	2.387324	.2366214	10.09	0.000	1.92354	2.851108
age4	-.1155994	.0121579	-9.51	0.000	-.1394292	-.0917696
age5	.0026948	.0003117	8.65	0.000	.0020839	.0033057
age6	-.000024	3.14e-06	-7.64	0.000	-.0000302	-.0000178
lgsize	142.9555	2.60638	54.85	0.000	137.8469	148.0641
floorm	36.77543	.9196493	39.99	0.000	34.97289	38.57796
floorh	50.67992	1.013061	50.03	0.000	48.69429	52.66555
discountrate	-3.788236	.0602543	-62.87	0.000	-3.906336	-3.670136
unluckynum	-16.95358	2.134466	-7.94	0.000	-21.13719	-12.76997
luckynum	8.629191	1.282952	6.73	0.000	6.11457	11.14381
lgdtoNoSec	-.3966299	.3860973	-1.03	0.304	-1.153391	.3601312
lgdtocentralcar	-179.4157	17.19418	-10.43	0.000	-213.1167	-145.7146
INTperioddtocentral	4.595493	.3403459	13.50	0.000	3.928406	5.26258
INTperioddtocentral2	-.0184446	.0015539	-11.87	0.000	-.0214903	-.0153988

lgttocentralcar	32.43666	22.95859	1.41	0.158	-12.56281	77.43612
INTperiodttocentral	-.3139625	.5898407	-0.53	0.595	-1.470066	.8421412
INTperiodttocentral2	-.0326846	.0048038	-6.80	0.000	-.0421003	-.023269
INTperiodttocentral3	.0001257	.0000136	9.24	0.000	.000099	.0001523
lgttocentralpublic	47.94877	10.36984	4.62	0.000	27.6236	68.27393
INTperiodcentralpublic	-5.703425	.3495382	-16.32	0.000	-6.38853	-5.018321
INTperiodcentralpublic2	.0401846	.0037658	10.67	0.000	.0328035	.0475657
INTperiodcentralpublic3	-.0000733	.0000118	-6.21	0.000	-.0000964	-.0000502
lgdtoborder	-127.0069	34.89896	-3.64	0.000	-195.4098	-58.60396
lgincome	127.6456	26.73498	4.77	0.000	75.24429	180.0468
INTincomeborder	12.38592	3.572977	3.47	0.001	5.382784	19.38905
lgelev	2.320344	.9005951	2.58	0.010	.555153	4.085534
lgcoast	-3.990305	1.119347	-3.56	0.000	-6.184254	-1.796356
countdss	-.3845319	.2288844	-1.68	0.093	-.8331515	.0640878
INTperiodDSS	-.0247872	.0083196	-2.98	0.003	-.0410937	-.0084806
INTperiodDSS2	.0002422	.0000892	2.71	0.007	.0000673	.000417
INTperiodDSS3	-8.52e-07	2.79e-07	-3.05	0.002	-1.40e-06	-3.05e-07
lgwalkelementary	24.91483	1.661573	14.99	0.000	21.6581	28.17156
INTperiodelementary	-.2513562	.0343377	-7.32	0.000	-.3186589	-.1840534
INTperiodelementary2	.0012663	.0001588	7.97	0.000	.0009551	.0015775
slopeelementary	39.1732	1.425423	27.48	0.000	36.37933	41.96707
INTslopewalkelem	-7.430824	.2551369	-29.12	0.000	-7.930899	-6.930749
lgmtronehalf	-14.18257	.9296417	-15.26	0.000	-16.00469	-12.36045
walkmtrangle	-14.43456	1.852028	-7.79	0.000	-18.06458	-10.80453
INTmtrcompslope	1.668664	.2786628	5.99	0.000	1.122477	2.21485
dtoairportadj	-.0003271	.0001273	-2.57	0.010	-.0005767	-.0000775
courtsize	-.0071154	.0004487	-15.86	0.000	-.0079947	-.006236
dcode2	-102.5085	9.600835	-10.68	0.000	-121.3264	-83.69063
dcode3	-5.366907	9.560381	-0.56	0.575	-24.10551	13.3717
dcode4	58.52261	6.272558	9.33	0.000	46.22823	70.817
dcode5	116.0198	5.950018	19.50	0.000	104.3576	127.682
dcode6	54.30747	7.987772	6.80	0.000	38.65122	69.96372
dcode7	-16.19138	3.247488	-4.99	0.000	-22.55654	-9.826212
dcode8	47.70415	4.026691	11.85	0.000	39.81173	55.59658
dcode9	142.2875	7.666229	18.56	0.000	127.2615	157.3135
dcode10	82.99183	3.720582	22.31	0.000	75.69939	90.28427
dcode11	89.78031	6.114951	14.68	0.000	77.79484	101.7658
dcode12	-10.49501	8.431582	-1.24	0.213	-27.02114	6.03112
dcode13	146.79	6.049083	24.27	0.000	134.9336	158.6464
dcode14	147.4052	4.913077	30.00	0.000	137.7755	157.035
dcode15	-81.69436	9.097531	-8.98	0.000	-99.52577	-63.86296
ncode190	-134.6933	6.440079	-20.91	0.000	-147.316	-122.0706
ncode94	192.3898	5.016556	38.35	0.000	182.5572	202.2224
ncode33	234.0675	5.750387	40.70	0.000	222.7966	245.3385
_cons	-484.1725	266.8052	-1.81	0.070	-1007.118	38.77291

8. Regression output, model with elevation/coastline distance interaction effects

Source	SS	df	MS
Model	2.8854e+09	93	31026022.1
Residual	216267151	37651	5743.99488
Total	3.1017e+09	37744	82176.9607

Number of obs	= 37745
F(93, 37651)	= 5401.47
Prob > F	= 0.0000
R-squared	= 0.9303
Adj R-squared	= 0.9301
Root MSE	= 75.789

priceadjtwothird	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
feb	6.292134	2.060191	3.05	0.002	2.254104	10.33016
mar	18.34334	1.961177	9.35	0.000	14.49938	22.1873
apr	26.23186	2.032138	12.91	0.000	22.24881	30.2149
may	33.73313	1.982872	17.01	0.000	29.84665	37.61961
jun	44.09935	2.09571	21.04	0.000	39.9917	48.207
jul	51.79824	2.175946	23.80	0.000	47.53333	56.06315
aug	59.20152	2.290058	25.85	0.000	54.71294	63.69009
sep	66.15327	2.395601	27.61	0.000	61.45783	70.84872
oct	70.48474	2.492715	28.28	0.000	65.59895	75.37053
nov	73.43895	2.647627	27.74	0.000	68.24953	78.62837
dec	82.34444	2.792383	29.49	0.000	76.87129	87.81758
yd1998	-89.75825	5.57298	-16.11	0.000	-100.6814	-78.83506
yd1999	61.87977	8.124517	7.62	0.000	45.9555	77.80404
yd2000	153.8699	10.9011	14.12	0.000	132.5034	175.2363
yd2001	217.7052	13.23119	16.45	0.000	191.7717	243.6386
yd2002	288.278	15.15813	19.02	0.000	258.5676	317.9883
yd2003	341.7242	16.83083	20.30	0.000	308.7353	374.7131
yd2004	500.4287	18.41763	27.17	0.000	464.3296	536.5277
yd2005	635.7957	19.91616	31.92	0.000	596.7595	674.8319
yd2006	723.5017	21.41669	33.78	0.000	681.5244	765.479
yd2007	818.0055	22.86556	35.77	0.000	773.1884	862.8226
yd2008	932.4985	24.37647	38.25	0.000	884.72	980.2771
yd2009	979.9529	25.83001	37.94	0.000	929.3254	1030.58
yd2010	1136.524	27.34606	41.56	0.000	1082.925	1190.123
yd2011	1296.791	28.99748	44.72	0.000	1239.956	1353.627
yd2012	1404.754	30.75935	45.67	0.000	1344.465	1465.043
yd2013	1594.499	32.94448	48.40	0.000	1529.927	1659.071
yd2014	1613.26	35.44044	45.52	0.000	1543.796	1682.725
age	104.4817	11.94181	8.75	0.000	81.07539	127.8879
sqrage	-25.73002	2.399943	-10.72	0.000	-30.43397	-21.02607
age3	2.511631	.2363959	10.62	0.000	2.048289	2.974974
age4	-.1217986	.0121455	-10.03	0.000	-.1456042	-.0979931
age5	.002847	.0003113	9.14	0.000	.0022367	.0034572
age6	-.0000255	3.14e-06	-8.12	0.000	-.0000316	-.0000193
lgsize	141.6122	2.603531	54.39	0.000	136.5092	146.7152
floorm	52.46393	7.15292	7.33	0.000	38.44401	66.48385
floorh	59.72064	7.736108	7.72	0.000	44.55766	74.88362
discountrate	-3.788374	.0601637	-62.97	0.000	-3.906296	-3.670452
unluckynum	-13.27319	2.158238	-6.15	0.000	-17.50339	-9.042986
luckynum	9.909796	1.285939	7.71	0.000	7.389322	12.43027
lgdtoNoSec	-.4639383	.385586	-1.20	0.229	-1.219697	.2918207
lgdtocentralcar	-174.342	17.16619	-10.16	0.000	-207.9882	-140.6958
INTperioddtocentral	4.437695	.3400457	13.05	0.000	3.771196	5.104193
INTperioddtocentral2	-.017596	.0015533	-11.33	0.000	-.0206404	-.0145515
lgttocentralcar	39.73947	22.92292	1.73	0.083	-5.190077	84.66901
INTperiodttocentral	-.3306998	.5889467	-0.56	0.574	-1.485051	.8236517
INTperiodttocentral2	-.0316286	.0047963	-6.59	0.000	-.0410294	-.0222277
INTperiodttocentral3	.0001203	.0000136	8.86	0.000	.0000937	.000147
lgttocentralpublic	38.81608	10.37956	3.74	0.000	18.47185	59.1603
INTperiodcentralpublic	-5.513022	.3492566	-15.79	0.000	-6.197575	-4.82847
INTperiodcentralpublic2	.038509	.0037613	10.24	0.000	.0311368	.0458812
INTperiodcentralpublic3	-.0000687	.0000118	-5.84	0.000	-.0000918	-.0000456
lgdtoborder	-37.96516	35.88022	-1.06	0.290	-108.2914	32.36105
lgincome	184.1586	27.25781	6.76	0.000	130.7326	237.5847
INTincomeborder	2.995382	3.67943	0.81	0.416	-4.216401	10.20716

lgelev	50.37444	4.729537	10.65	0.000	41.10442	59.64446
lgcoast	22.95524	2.559196	8.97	0.000	17.93915	27.97133
countdss	-.3476987	.2285048	-1.52	0.128	-.7955743	.100177
INTperiodDSS	-.027378	.0083085	-3.30	0.001	-.0436628	-.0110932
INTperiodDSS2	.0002766	.0000891	3.11	0.002	.000102	.0004513
INTperiodDSS3	-9.45e-07	2.79e-07	-3.39	0.001	-1.49e-06	-3.98e-07
lgwalkelementary	24.82038	1.658492	14.97	0.000	21.56969	28.07107
INTperiodelementary	-.2612973	.0342801	-7.62	0.000	-.3284872	-.1941075
INTperiodelementary2	.0012821	.0001585	8.09	0.000	.0009715	.0015928
slopeelementary	38.64083	1.423716	27.14	0.000	35.85031	41.43135
INTslopewalkelem	-7.255296	.2551606	-28.43	0.000	-7.755418	-6.755174
lgmtronehalf	-14.80303	.9297606	-15.92	0.000	-16.62538	-12.98067
walkmtrangle	-14.92192	1.850017	-8.07	0.000	-18.548	-11.29583
INTmtrcompslope	1.734552	.2783482	6.23	0.000	1.188982	2.280122
dtoairportadj	-.0003353	.0001271	-2.64	0.008	-.0005845	-.0000862
courtsize	-.0080344	.0004563	-17.61	0.000	-.0089289	-.00714
dcode2	-101.6631	9.59429	-10.60	0.000	-120.4682	-82.85806
dcode3	-4.662498	9.541883	-0.49	0.625	-23.36485	14.03985
dcode4	54.68764	6.273125	8.72	0.000	42.39215	66.98314
dcode5	113.512	5.943838	19.10	0.000	101.8619	125.162
dcode6	47.95905	7.998197	6.00	0.000	32.28237	63.63573
dcode7	-14.02027	3.257695	-4.30	0.000	-20.40544	-7.635097
dcode8	46.63947	4.028778	11.58	0.000	38.74295	54.53598
dcode9	137.9871	7.660815	18.01	0.000	122.9717	153.0025
dcode10	78.54776	3.737166	21.02	0.000	71.22282	85.87271
dcode11	89.05953	6.112577	14.57	0.000	77.07871	101.0403
dcode12	-13.74353	8.426908	-1.63	0.103	-30.2605	2.773435
dcode13	147.2157	6.041743	24.37	0.000	135.3737	159.0577
dcode14	147.2461	4.903683	30.03	0.000	137.6348	156.8575
dcode15	-90.79479	9.121279	-9.95	0.000	-108.6727	-72.91683
ncode190	-130.5002	6.443347	-20.25	0.000	-143.1294	-117.8711
ncode94	189.8152	5.022256	37.79	0.000	179.9714	199.659
ncode33	228.8399	5.765466	39.69	0.000	217.5395	240.1404
intcoastelev	-7.015394	.6407216	-10.95	0.000	-8.271226	-5.759563
intcoastfloorH	-3.716676	.944821	-3.93	0.000	-5.56855	-1.864801
intcoastfloorM	-3.405274	.8594254	-3.96	0.000	-5.089771	-1.720777
intelevfloorH	4.93582	1.223925	4.03	0.000	2.536894	7.334746
intelevfloorM	2.233557	1.171197	1.91	0.057	-.0620204	4.529134
_cons	-1202.476	274.7787	-4.38	0.000	-1741.049	-663.9022

9. Regression output, model without observations that experienced metro system expansions during Aug 1997 – July 2014

Source	SS	df	MS
Model	2.5528e+09	88	29009464.1
Residual	183774312	30439	6037.46219
Total	2.7366e+09	30527	89645.4663

Number of obs	= 30528
F(88, 30439)	= 4804.91
Prob > F	= 0.0000
R-squared	= 0.9328
Adj R-squared	= 0.9327
Root MSE	= 77.701

priceadjtwothird	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
feb	5.036695	2.341309	2.15	0.031	.4476305	9.625759
mar	17.94827	2.218611	8.09	0.000	13.5997	22.29684
apr	29.46624	2.298151	12.82	0.000	24.96177	33.97071
may	36.58028	2.265643	16.15	0.000	32.13952	41.02103
jun	49.55153	2.391417	20.72	0.000	44.86426	54.23881
jul	57.59847	2.481195	23.21	0.000	52.73522	62.46171
aug	66.13634	2.623746	25.21	0.000	60.99369	71.279
sep	73.72215	2.75817	26.73	0.000	68.31602	79.12828
oct	78.27406	2.868939	27.28	0.000	72.65082	83.8973
nov	83.07389	3.046089	27.27	0.000	77.10342	89.04435
dec	91.84481	3.213883	28.58	0.000	85.54547	98.14416

yd1998	-85.81224	7.779012	-11.03	0.000	-101.0594	-70.56505
yd1999	79.56875	10.78312	7.38	0.000	58.43339	100.7041
yd2000	173.7437	14.07907	12.34	0.000	146.1482	201.3393
yd2001	245.1106	17.03435	14.39	0.000	211.7225	278.4986
yd2002	319.0334	19.40205	16.44	0.000	281.0046	357.0622
yd2003	379.9671	21.38608	17.77	0.000	338.0495	421.8848
yd2004	554.9014	23.1677	23.95	0.000	509.4917	600.311
yd2005	696.0894	24.8406	28.02	0.000	647.4008	744.778
yd2006	790.6888	26.49889	29.84	0.000	738.7499	842.6278
yd2007	895.7382	28.1407	31.83	0.000	840.5812	950.8952
yd2008	1021.947	29.84465	34.24	0.000	963.4498	1080.443
yd2009	1080.124	31.5153	34.27	0.000	1018.353	1141.896
yd2010	1249.124	33.27527	37.54	0.000	1183.903	1314.345
yd2011	1422.055	35.18394	40.42	0.000	1353.093	1491.017
yd2012	1544.779	37.2615	41.46	0.000	1471.745	1617.813
yd2013	1748.721	39.72456	44.02	0.000	1670.859	1826.583
yd2014	1781.539	42.48593	41.93	0.000	1698.265	1864.813
age	137.1057	13.33486	10.28	0.000	110.9688	163.2425
sqrage	-31.0066	2.665916	-11.63	0.000	-36.23191	-25.78129
age3	2.92607	.2610172	11.21	0.000	2.414466	3.437675
age4	-.1393753	.013327	-10.46	0.000	-.1654969	-.1132538
age5	.0032384	.0003396	9.54	0.000	.0025728	.0039039
age6	-.0000291	3.40e-06	-8.55	0.000	-.0000357	-.0000224
lgsize	142.5362	2.970779	47.98	0.000	136.7134	148.3591
floorm	39.11499	1.04422	37.46	0.000	37.06827	41.16171
floorh	54.49405	1.157991	47.06	0.000	52.22433	56.76376
discountrate	-4.506586	.0714687	-63.06	0.000	-4.646668	-4.366505
unluckynum	-25.85843	2.719161	-9.51	0.000	-31.1881	-20.52876
luckynum	8.601567	1.419511	6.06	0.000	5.819266	11.38387
lgdtoNoSec	-1.043144	.4618881	-2.26	0.024	-1.948464	-.1378245
lgdtocentralcar	-124.7752	19.3483	-6.45	0.000	-162.6987	-86.85178
INTperioddtocentral	3.862117	.378425	10.21	0.000	3.120388	4.603845
INTperioddtocentral2	-.0158978	.0016869	-9.42	0.000	-.0192043	-.0125914
lgttocentralcar	-18.00662	26.14225	-0.69	0.491	-69.24652	33.23328
INTperiodttocentral	-.4929462	.6555743	-0.75	0.452	-1.777899	.7920069
INTperiodttocentral2	-.0220519	.0054126	-4.07	0.000	-.0326609	-.011443
INTperiodttocentral3	.0000857	.0000155	5.54	0.000	.0000554	.0001161
lgttocentralpublic	47.35692	13.38076	3.54	0.000	21.13006	73.58378
INTperiodcentralpublic	-4.863018	.4236612	-11.48	0.000	-5.693412	-4.032624
INTperiodcentralpublic2	.0295148	.0043903	6.72	0.000	.0209096	.03812
INTperiodcentralpublic3	-.0000448	.0000134	-3.34	0.001	-.000071	-.0000185
lgdtoborder	6.650974	40.81666	0.16	0.871	-73.35139	86.65333
lgincome	160.5581	31.70807	5.06	0.000	98.40895	222.7073
INTincomeborder	-1.241668	4.175171	-0.30	0.766	-9.425178	6.941842
lgelev	7.100151	1.102289	6.44	0.000	4.939618	9.260684
lgcoast	-.0695761	1.315623	-0.05	0.958	-2.648251	2.509099
countdss	-.2114836	.2869211	-0.74	0.461	-.773861	.3508938
INTperiodDSS	-.0337092	.0099375	-3.39	0.001	-.0531871	-.0142312
INTperiodDSS2	.0002696	.000103	2.62	0.009	.0000678	.0004714
INTperiodDSS3	-8.61e-07	3.15e-07	-2.73	0.006	-1.48e-06	-2.43e-07
lgwalkelementary	38.63743	2.395735	16.13	0.000	33.94169	43.33317
INTperiodelementary	-.4279532	.0439495	-9.74	0.000	-.5140961	-.3418103
INTperiodelementary2	.0018224	.0001909	9.55	0.000	.0014483	.0021965
slopeelementary	40.20097	1.626182	24.72	0.000	37.01358	43.38835
INTslopewalkelem	-7.609063	.2910148	-26.15	0.000	-8.179464	-7.038662
lgmtronehalf	-25.33829	1.338212	-18.93	0.000	-27.96124	-22.71534
walkmtrangle	-23.92774	2.15782	-11.09	0.000	-28.15716	-19.69832
INTmtrcompslope	2.921758	.3243638	9.01	0.000	2.285991	3.557524
dtoairportadj	.000084	.000196	0.43	0.668	-.0003003	.0004683
courtsize	-.0069711	.0005111	-13.64	0.000	-.0079728	-.0059693
dcode2	-127.588	12.36721	-10.32	0.000	-151.8282	-103.3477
dcode3	-1.639931	9.927741	-0.17	0.869	-21.09872	17.81886
dcode4	23.45508	8.264876	2.84	0.005	7.25558	39.65459
dcode5	88.30628	7.877235	11.21	0.000	72.86657	103.746

dcode6	22.02418	9.783514	2.25	0.024	2.848079	41.20027
dcode7	-23.27953	3.874884	-6.01	0.000	-30.87446	-15.68459
dcode8	36.78261	4.966717	7.41	0.000	27.04763	46.51758
dcode9	113.7143	9.75144	11.66	0.000	94.60102	132.8275
dcode10	94.03473	4.138834	22.72	0.000	85.92245	102.147
dcode11	63.79151	7.333915	8.70	0.000	49.41673	78.16629
dcode12	-40.23401	11.28516	-3.57	0.000	-62.35339	-18.11462
dcode13	122.3391	7.925812	15.44	0.000	106.8041	137.874
dcode14	146.5835	5.631012	26.03	0.000	135.5465	157.6206
dcode15	-131.5825	11.65129	-11.29	0.000	-154.4195	-108.7455
ncode190	-115.9639	6.688652	-17.34	0.000	-129.074	-102.8539
ncode94	170.3244	5.55623	30.65	0.000	159.434	181.2149
ncode33	265.5384	8.040934	33.02	0.000	249.7778	281.2989
_cons	-903.2113	313.4548	-2.88	0.004	-1517.596	-288.8268