CUER Working Paper

Fu Yuming, Daniel P. McMillen, and Tsur Somerville
Land Assembly: Measuring Holdup
December 2002
CUER Working Paper 02-08
Land Assembly: Measuring Holdup

December 2002

Fu Yuming*, Daniel P. Mc Millen**, and Tsur Somerville***

* Department of Real Estate, School of Design and Environment, National University of Singapore, 4 Architecture Drive 117566, Singapore. Tel: (+65) 874-5161, Fax: (+65) 774-8684. Email: rstfuym@nus.edu.sg

** Department of Economics, University of Illinois at Chicago, 601 S. Morgan St. MC 144, Chicago, IL 60607, USA. Tel: (312) 413-2100. Email: mcmillen@uic.edu

*** Faculty of Commerce, University of British Columbia, 2053 Main Mall, Vancouver, BC, V6T1Z2, Canada. Tel: (604) 822-8343, Fax: (604) 822-8477. Email: tsur.somerville@commerce.ubc.ca
Abstract

One justification for the use of eminent domain for urban redevelopment is to mitigate the holdup costs associated with assembling parcels of urban land. Theoretical work by Eckhart (1985) and Strange (1995) shows that the holdup power held by the owner of the last lot to be assembled raises the price of that lot. In equilibrium, optimal development size will fall with the number of parcels that must be assembled. We use data on site transactions for lots that are subsequently redeveloped in Hong Kong to estimate the holdup premium for a lot. Using a semiparametric approach, we study how land prices vary with the occurrence and degree of land assembly in a sample of more than 700 transactions in Hong Kong. We find that the last lot in an assembled group of parcels receives a premium as does the sale of a group of assembled units. However, lots that will be later assembled sell at a discount, in order to offset the expected holdup costs, the eventual premium paid for the last lot. The discounts for land assembly suggested by the theoretical literature support the hypothesis that land prices will be convex over the range of lot sizes where assembly is an issue. We find that unit land prices (price per square foot of lot area) rise with site size up to 4,300 square feet, after which they fall. This finding is the first evidence of the long speculated region where land values are convex in lot size. However, because we control for land assembly, the likely explanation for the convex region is the higher construction costs for building on small parcels in Hong Kong’s dense urban environment.
I. Introduction

The State often takes a role in private urban development by using its power of eminent domain to assemble separate lots to form the development parcel. This use of state power is justified by arguing that a single landowner may hold up a planned development that is in the public interest by trying to extract rents from the developer. This is a particular variant of the type of holdup that concerns studies of the organizational boundaries of the firm. Both vertical integration in the production process, the presence of long term contracts, and repeated contracting between the two parties have been explained as solutions to the “holdup” problem.\(^1\) While concerns about land assembly are often raised, there is no literature to measure the actual scope of these costs. In this paper we use data from redeveloped parcels in the core areas of Hong Kong to provide the first empirical measure of land assembly costs. In doing so, we provide a new view on the costs of holdup in an industry where both vertical integration and the type of long-term contracts observed in other industries are not possible.

Our data comprises over 700 transactions of both assembled and single lots in Hong Kong that are redeveloped subsequent to the observed transaction. Controlling for the characteristics of the lots, their locations, and the sequence of assembly, we find clear evidence of holdup costs: the last lot in a group of assemble parcels enjoys a premium of 23 to 28 percent, assembled parcels charge a premium of 14 to 20 percent, and those parcels that will subsequently be combined to form an assembled lot are sold at a

\(^1\) The potential for one party to a transaction to “hold up”, extract rents, from another when their exchange requires the second party to invest in physical or human capital unique to the transaction, i.e. transaction-specific assets.
discount of 5 to 11 percent. The holdup premium is greatest when the size of the surplus is largest relative to the value of the last lot assembled.

One advantage of this data set is that it includes many relatively small lots at the heart of a major urban area. These small lots allow us to measure the relationship between lot value and size using just those parcels for which land assembly is likely to be an issue. The existing literature has used larger suburban residential tracts and overwhelmingly found that lot value is concave in lot size, i.e. unit land prices falling in lot size. Using a semiparametric specification that allows the marginal price of lot size to vary nonparametrically, we find the first evidence of a range of over which the price per square foot of land is rising with lot size. Between 421 sq. ft. and 4,300 sq. ft., the price per square foot of land is rising with lot size, so that land values are convex in lot size. Above 4,300 sq. ft. it falls. Our regressions control for holdup using a variety of continuous, dummy, and interaction variables. We believe the range of convex land values reflects larger discounts for small lots because of the additional costs of constructing buildings on small lots in a dense, developed urban area.

The remainder of the paper is organized as follows. Section II lays out the relevant theoretical and empirical work in both the industrial organization and urban economics literatures. Section III describes the data. In Section IV we lay out the empirical

---

2 The exception is Colwell and Munneke (1999) who use vacant lots in Chicago and find no evidence of convexity.

3 The basic urban economics relationship is that land values are the value of the lot with the profit maximizing structure on it net of the cost of building that structure, so that values will fall as construction costs rise.
specification and provide the estimation results. Finally, a conclusion with possible
extensions is offered in Section VI.

II. Theoretical and Empirical Background

This paper draws on a wide range of literature that addresses holdup in one form or
another. Holdup plays a prominent role in the theory of the firm literature. There is a
small theoretical literature that explicitly models the role of holdup in land assembly.
Finally, research most associated with Peter Colwell studies the relationship between lot
size and value. This paper relates to aspects of all of these areas, making an exhaustive
literature survey excessive. Instead, we will try to highlight those theoretical and
empirical finding of most relevance or consequence for the research presented here.

An extensive literature relies on holdup resulting from physical or human capital
specificity to explain both vertical integration and aspects of contracting. In their seminal
study of General Motor’s 1926 acquisition of Fisher Body, Klein, Crawford, and Alchain
(1978) argue that GM used vertical integration to overcome the problem of holdup that
resulted from asset specificity. Monteverde and Teece’s (1982) study of why
automotive manufacturers choose to subcontract some work but do other activities within
the firm is the first among many empirical papers to use holdup-motivated subcontracting

---

4 So much so that Holmstrom and Roberts (1998) argue that there has been too much focus on the role of
holdup and asset specificity in determining the boundaries of the firm.
5 Though both Coase (2000) and Freedland (2000) reject this particular case as an example of holdup
driving vertical integration.
to examine vertical integration. The transactions cost literature, which is most associated with Williamson (1975, 1979, and 1985), argues that uncertainty about the nature and range of future outcomes or other problems with the design of optimal contracts motivate firms to decide to do which activities are internal and which external to the firm. Williamson’s approach predicts that the activity will be inside the firm only if contracts cannot be designed to mitigate the holdup problem. Joskow (1987) provides clear empirical support to the role of contracts in addressing problems of holdup. He finds that electric utilities and coal providers agree on enter into longer terms for contracts as the degree of asset specificity (exposure to holdup) rises.

The contribution of our paper to this literature is that we explicitly identify the costs of holdup. With real estate development, developers typically acquire the site and then develop it. The only contractual alternative is to obtain a ground lease from the landowner, with the contractual issues being renewal options and the treatment of any structures put in place by the developer. In Hong Kong, the government retains the ownership of all sites, so that developers must acquire the ground lease to develop. Thus, there is no alternative for developers other than either acquisition or joint venture partnership with the current holder of the ground lease. In contrast to the existing literature, which focuses on outcomes and firm characteristics, we are explicitly able to identify holdup costs by identifying the premium on land that is subject to holdup.

---

6 Whereas most of this literature is cross-sectional, González-Díaz, Arruñada, and Fernández (2000) use panel data on subcontracting by construction firms to study the effects of firm activity specificity, output heterogeneity, and asset use on subcontracting.
The urban economics literature on holdup and land assembly consists of just two papers. Eckert (1985) and Strange (1995) both study land assembly in a bargaining-game model, where a developer simultaneously negotiates with separate landowners to assemble land. The equilibrium outcome of these games shows that small landowners ask for more per acre for their land than large owners because they have less to lose if their holdout jeopardizes the assembly project. The empirical implication of this outcome is that holdup leads the price per square foot of transacted land to be lower when the parcel accounts for a larger proportion of the entire developed site.  

In reality, land assembly often involves sequential bargaining instead of simultaneous bargaining. This timing allows the owner of the last assembled lot to extract a premium. However, small landowners may not be able to exploit their potential holdup power under sequential bargaining. The developer has the ability to change the characteristics of the planned development depending on the path of the game. For instance, a developer might construct a different, smaller project on the site that excludes the lots of landowners who tried to extract an excessive share of the redevelopment profit. Thus, the holdup power of any landowners depends on whether their sites are indispensable to the land assembly project.

Appraisers have long been interested in the relationship between lot area, geometry, and the per unit price of land. They use the term *plottage* to describe how land prices per

---

7 The price for an assembled unit will also depend on the nature of the assembly. Strange (1995) shows that the total cost of land assembly increases (thus the profit of the developer decreases) with the number of landowners involved, regardless of the physical size of the land.

8 See Colwell (1999) for a brief summary.
unit of area should rise with lot size because of the costs of land assembly for smaller lots. Whereas a convex land value function is the accepted wisdom, Colwell (1999) notes that earlier appraisal rules implicitly assume concavity when lot size is measured by depth. The empirical literature finds concavity. For example, Brownstone and De Vany (1991) and Colwell and Sirmans (1993) find that the price per acre is lower for large lots. Colwell and Sirmans provide a substantial citation list of earlier empirical and theoretical studies. Thorsnes and McMillen (1998) use a semiparametric estimator that allows the size-value relationship to vary nonparametrically. They too find that land prices are concave in size. Concavity suggests that subdivision costs dominate assembly costs.

A drawback of this literature is that the data sets typically include only raw land transactions in suburbs and at the edge of cities where subdivision is common. Thus, we would expect the costs of subdivision to give a concave land value function. The one previous study to use urban land values, Colwell and Munneke (1999), focuses on vacant land transactions in the Chicago metropolitan area. They find that the degree of concavity is lower within the urban center than in outlying areas. They find no evidence of a region in which units prices increase with lot size. Colwell and Munneke’s data set consists of large suburban parcels that are not likely candidates for assembly: the mean size is 282,000 square feet, only 30 percent are in the Chicago city limits, and the mean distance from the city center is 16.6 miles.9

---

9 In comparison our mean parcel size is 6,500 square feet with a mean distance of 3.4 miles (5.6 km) to the Hong Kong CBD.
III. Data

Our study focuses on the land value-parcel size relationship in a market where land assembly is commonly observed. The pricing of land in such a market is understudied due to the scarcity of land transaction data in such markets. Our data set covers the sales of sites for redevelopment in central urban areas of Hong Kong during the 1990s. All land in Hong Kong is owned by the government. Private development occurs through long-term land (ground) leases from the government. Lease conditions specify the allowable density, the plot ratio or the ratio of floor space to lot area, and permitted uses. A land premium based on the market value of the land lease is payable upon the granting of the lease. After the Sino-British joint declaration on Hong Kong’s handover in 1984, new land leases would mature before June 30, 2047 (50 years from the handover in July 1997). This policy of 50 years has been continued by the government of the Hong Kong Special Administrative Region. Land leases granted after 1984 require an additional annual rental payment equal to 3 percent of the property’s rateable value starting in July 1997. Typically, land leases can be extended for another 50 years upon expiry without an additional land premium.

The transactions in our dataset are sales of existing leases. These sites can be redeveloped. If lease terms, particularly the allowed density and use, are below those demanded in the market and found in newer leases, a developer can apply to the government for a modification of the lease conditions. The government would typically grant the change according to the prevailing lease conditions in the local area, charging a lease
modification premium equal to the difference in land value due to the modification. The
government effectively extracts part of the redevelopment option value as a kind of a
Henry George tax.

Our transaction data come from a database of block transactions with existing lease
conditions. Nearly all sites have includes structures. The database was obtained from a
commercial vendor (EPRC), who has collected all property transactions in Hong Kong
since 1992 (and some transactions in 1991). The company matches the transaction data
(price, date of agreement for sale and purchase, the identity of the property and the
identity of the buyer) with property-specific data from government files (e.g. existing
floor areas, plot ratios, permitted land uses and planning zones). Unfortunately, most
data on floor area is missing, and the buyer identity is the name of the acquiring
corporation rather than the owner of the corporation. Thus, we cannot link across
transactions. The block transaction records include comments summarizing any news
reports concerning the intentions of the buyers and sellers. In addition to existing lease
conditions, prices reflect any negotiated changes in terms, the redevelopment option, and
the value of the stream of rents from the existing structure until the development option
hits its trigger. These data give us 1,878 site transactions on Hong Kong Island and the
Kowloon peninsula, which is the core area of the Hong Kong SAR, for the period

Whereas we are interested in land prices, our data set is dominated lots with structures.
To extract land prices from these data we follow Helsley and Rosenthal (1994), who
show that for sites that are subsequently redeveloped, transactions prices are land prices net of demolition costs. Thus, if we can identify sites that redevelop, we can treat the transactions as a land price, controlling for the time from transaction until redevelopment.

We use two sources to identify those sites that develop. The first source is the list of approved general building plans (GBP). These plans must be filed with the Building Department and approved before any construction can occur in Hong Kong. We match the sites with plans filed after 1992 with those in the block transaction database to create a database of land sales for redevelopment. The GBP data list all addresses in the parcel that are approved for development. This procedure results in a set of 747 transactions, of which 412 are involved in an assembly. We also use the remarks about the transaction to identify another 52 transactions where there is an indication that a development is planned. For these latter sites we do not have GBP data because a development plan was never approved.\(^\text{10}\)

We use information on the timing of the GBP approved development plots to create a series of variables describing the assembly process. We identify whether a sale (1) is part of a land assembly; (2) consists of assembled lots; (3) is to be combined with other lots later; and (4) is the last acquisition of a land assembly. We also measure the total area of the final assembled site to which a sale belongs. In addition, we can calculate the time between a sale and the GBP approval date as a measure of time to redevelopment.

\(^{10}\) Since we use time to development in our regressions, these observations are eventually excluded from the regressions.
Geocoding the data allows us to include the distance from each site to significant locations such as the central business district (CBD) and mass-transit stations. It also provides us with geographic coordinates to measure the distance between observations, which is a critical component of our nonparametric estimation procedure. We index prices to extract the temporal pattern of real estate prices in Hong Kong. Using the quarterly property price indices published by the Rating and Valuation Department for each of four different property types – residential, office, commercial, and factory – we deflate the land sale prices so that they reflect the variation around the general trend for the particular class of real estate.\textsuperscript{11} As a result, we only make limited use of time dummies in the regressions.

We have a total of 799 land sales in the final sample. They are dated between 1991 and 1998. Most sales occurred between 1992 and 1995 (some of the later block transactions are not yet reflected in our GBP database and thus not included in the final sample). The descriptive statistics for these data are shown in Table 1. On average, they are located about 5.7 km from the CBD. About 63 percent of these sales are associated with a land assembly, and 42 percent of them are for sites that sold more than once during the sample period. 39 percent of the sales are of assembled lots and 33 percent would be assembled later. In all but 58 cases, we are able to identify the size of the final assembled site for redevelopment.

\footnote{\textsuperscript{11}This is an appraisal based survey, but one derived from a very large volume of transactions of both whole buildings and residential and non-residential strata (condominium) units.}
### Table 1
#### Summary Statistics

<table>
<thead>
<tr>
<th>Variable</th>
<th>Obs</th>
<th>Mean</th>
<th>Std. Dev</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lot Price (HK 000000)</td>
<td>799</td>
<td>88.11</td>
<td>172.23</td>
<td>2</td>
<td>2450</td>
</tr>
<tr>
<td>Indexed Lot Price (HK 000000)</td>
<td>799</td>
<td>42.95</td>
<td>78.37</td>
<td>0.92</td>
<td>1056</td>
</tr>
<tr>
<td>Indexed Lot Price/ Sq. Ft. Lot Area - HK</td>
<td>799</td>
<td>11171</td>
<td>9700</td>
<td>263.8</td>
<td>81880</td>
</tr>
<tr>
<td>Lot Area (sq.ft.)</td>
<td>799</td>
<td>6510</td>
<td>11785</td>
<td>421</td>
<td>100211</td>
</tr>
<tr>
<td>Distance to Central (km)</td>
<td>793</td>
<td>5.66</td>
<td>3.4493</td>
<td>0.152</td>
<td>34.23</td>
</tr>
<tr>
<td>Dummy - Primarily Residential Building</td>
<td>799</td>
<td>0.5181</td>
<td>0.5</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Dummy - Primarily Industrial Building</td>
<td>799</td>
<td>0.1114</td>
<td>0.3148</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Dummy - Primarily Office Building</td>
<td>799</td>
<td>0.3479</td>
<td>0.4766</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Dummy - Primarily Commercial/Other Building</td>
<td>799</td>
<td>0.0225</td>
<td>0.1485</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Maximum Allowable Developable Sq.Ft.</td>
<td>799</td>
<td>38861</td>
<td>74745</td>
<td>2719</td>
<td>1316295</td>
</tr>
<tr>
<td>Size of Assembled Parcel (sq.ft.)</td>
<td>741</td>
<td>8265</td>
<td>12186</td>
<td>421</td>
<td>162246</td>
</tr>
<tr>
<td>Dummy - Last Lot Assembled Into Parcel</td>
<td>798</td>
<td>0.2932</td>
<td>0.4555</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Dummy - Transaction of Assembled Lots</td>
<td>798</td>
<td>0.3872</td>
<td>0.4874</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Dummy - Transaction Lot to be Included in Assembled Lots</td>
<td>799</td>
<td>0.3292</td>
<td>0.4702</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Years From Transaction to Accepted Development Plan</td>
<td>729</td>
<td>1.692</td>
<td>1.4711</td>
<td>0</td>
<td>8.167</td>
</tr>
<tr>
<td>Transacting Lot Share of Assembled Parcel Size</td>
<td>741</td>
<td>0.7053</td>
<td>0.3352</td>
<td>0.0179</td>
<td>1</td>
</tr>
</tbody>
</table>
IV. Empirical Analysis

A. Specification

Our empirical tests are structured to estimate the premiums and discounts associated with land assembly. The models of land assembly both assume a simultaneous game, though the data suggest a sequential pattern. The behavior of landowners whose lots are the first ones assembled should resemble that of the large landowners in Strange’s (1995) model. We develop the following testable hypotheses from the theoretical literature, making this change. First, the land assembly premium will be highest for the last lot assembled. Second, when land assembly is expected, those sites acquired prior to the final assembly – those held by large landowners in Strange’s model – have a strong incentive to lower their reservation price, leaving the total ask price unchanged for lots to be assembled. Third, an assembled parcel should trade at a premium because the holdup costs must be embodied in its price. Fourth, the premium will decrease with the size of the last lot relative to the size of the assembled parcel, mimicking Strange’s result that it is the small landowner who is able to extract the premium.

As noted above, whereas we would like to have direct measures of land value, most of the transactions in our data sets include structures. Unlike Helsley and Rosenthal (1994), not all of our sites transacted immediately after sale. While 45 percent of the transactions have a development plan approved within one year, the lag is three years or longer for 14 percent of the transactions. To control the discounted flow of rents received on the
property from the transaction date until the development option is exercised, we use the time from the transaction until GBP approval date to measure the length of time over which rents are received. The GBP approval date is a good measure of the option exercise time. Revenues also depend on the size of the existing structure. Unfortunately, we only know the size of the existing structure for 97 observations. Instead, we use the maximum allowable developable area.

Our base regression includes dummy variable for each of 38 districts. District-level fixed effects help to control for differences among neighborhoods and the effects of omitted variables that are correlated over space. Though the simple fixed-effects specification is adequate for districts with few sales, it fails to account for spatial effects within districts. Nonparametric estimators such as locally weighted regression (e.g., Fu and Somerville, 2001; McMillen and McDonald, 1997; Meese and Wallace, 1991) have proved useful for analyzing spatial models. However, these estimators are best suited for larger data sets because the variance of the estimates is high with even a small number of explanatory variables.

Semiparametric estimators combine the flexibility of nonparametric regression with the efficiency of standard parametric models. Examples of semiparametric models of house prices include Anglin and Gencay (1996), Pace (1995), Stock (1991), and Thorsnes and McMillen (1998). The semiparametric model is written

\[
\ln y_i = X_i \beta + f(z_i) + u_i
\]  

(1)
where $y_i$ is the price per square foot. Compared with a fully nonparametric model, the
parametric part of the equation, $X_i \beta$, increases the efficiency of the estimates. The
nonparametric part of the model, $f(z_i)$, is constrained only to be a smooth, continuous
function.

We use two semiparametric model specifications. As an alternative to district dummy
variables, the first specification models spatial effects nonparametrically by letting $z$
represent the geographic coordinates of each sale. In contrast to the fixed-effects
specification, this model accounts for spatial effects within districts. The second
semiparametric model follows Thorsnes and McMillen (1998)'s analysis of the effect of
lot size on sales prices. This specification retains the district-level fixed effects
specification, but allows lot size to enter the model nonparametrically. The flexible
semiparametric specification does not impose monotonicity on unit land prices, allowing
for regions where land values are convex in lot size and others where they are concave.

Following Robinson (1988), the semiparametric estimation procedure proceeds as
follows. First, a nonparametric procedure is used to estimate $E(\ln y_i | z_i)$ and $E(X_i | z_i)$. The
residuals are $e_y$ and $e_x$. Second, $e_y$ is regressed on $e_x$ to estimate $\beta$. The usual standard
errors from this regression are consistent estimates of the standard errors of $\hat{\beta}$.

We use Cleveland and Devlin’s (1988) nonparametric estimator, *locally weighted
regression*, to calculate the conditional expectations. The estimator has been used in
several urban applications, including Fu and Somerville (2001), Meese and Wallace
(1991), and McMillen and McDonald (1997). As an example, consider the calculation of \( E(\ln y_i | z_i) \). In the first specification, \( z \) includes two variables, the latitude and longitude of each sale. In the second specification, \( z \) includes the single variable lot size. Cleveland and Devlin approximate the general function \( E(\ln y_i | z_i) \) with a simple linear regression, with more weight given to observations that are closer to tract \( i \). In the first specification, “closer” has a simple definition as geographic distance. In the second specification, more weight is given to sales with similar lot sizes.

Following Cleveland and Devlin (1988) and previous urban research, we use the tri-cube function to define the weight given to observation \( j \) when constructing an estimate of \( E(\ln y_i | z_i) \) for observation \( i \). This weight is written as:

\[
\phi_{ij} = I(\rho_{ij} < \omega_i) \left( 1 - \left( \frac{\rho_{ij}}{\omega_i} \right)^3 \right)^3
\]

where \( \rho_{ij} \) is the distance between tracts \( i \) and \( j \), and \( I(\rho_{ij} < \omega_i) \) is an indicator variable that equals one when the distance between tracts \( i \) and \( j \) is less than \( \omega_i \). \( \rho_{ij} \) represents geographic distance for the first specification, and is the absolute value of the difference in lot sizes for the second model. The estimated value for \( \ln \, y \) at observation \( i \) is then simply the prediction from a weighted least squares regression of \( \ln \, y \) on \( z \), using \( \phi_{ij} \) for the weights. Only observations with \( \rho_{ij} < \omega_i \) are used in estimation.

The number of observations used for each regression is the \textit{window size}. The window size plays a role similar to the bandwidth in kernel regression. The window size stays fixed, leaving \( \omega_i \) to vary across the target points for the regressions. Small window sizes
lead to small bias but large variances, whereas large window sizes trade variance for bias. The choice of window sizes is much more important than the choice of functions for the weights in equation (2). We use the method of cross-validation to choose the window size (Pagan and Ullah, 1999). The cross-validated estimates are obtained by omitting observation $i$ and re-estimating the entire model. The residual for observation $i$ is $e_i = \ln y_i - X_i \hat{\beta}_i - \hat{f}_i(z_i)$, where $\hat{\beta}_i$ and $\hat{f}_i$ are the estimates with observation $i$ deleted.

The optimal window size is the value that minimizes $n^{-1} \sum_{i=1}^{n} e_i^2$. The cross-validation procedure leads to a window size of 40 observation for the first semiparametric specification and 350 observations for the specification in which lot size is the sole variable modeled nonparametrically.

**B. Empirical Results**

We present our empirical results in Table 2.\textsuperscript{12} The dependent variable in all cases is the natural logarithm of the indexed transaction price per square foot of lot area. The first regression is an OLS regression with neighborhood fixed effects.\textsuperscript{13} The second is a semiparametric specification where the spatial effects enter nonparametrically. Regression (3) is semiparametric, but it is the effect of lot area on the log of indexed transaction price per square foot of lot area that enters nonparametrically. As such, we do

\textsuperscript{12} To obtain coefficient results of similar size, we divide all lot area measures by 10000 before running the regressions. This also mitigates against potential problems in computation from using values with too many zeros after the decimal point.

\textsuperscript{13} We have 39 identified neighborhoods on Hong Kong Island and in Kowloon. However, we aggregate all neighborhoods with fewer than 10 transactions giving us 24 neighborhood dummies.
### Table 2 - Regression Results
Dependent Variable: Ln(Indexed Lot Price/ Sq. Ft. Lot Area)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Reg (1)</th>
<th>Reg (2)</th>
<th>Reg (3)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Hold-up &amp; Lot Area Variables</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lot Area (00000 sq.ft.)</td>
<td>-0.1627 ***</td>
<td>-0.1708 ***</td>
<td>See note (3)</td>
</tr>
<tr>
<td>(0.0399)</td>
<td>(0.0396)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Size of Assembled Parcel (00000 sq.ft.)</td>
<td>-0.0428 *</td>
<td>-0.0322</td>
<td>-0.0641 **</td>
</tr>
<tr>
<td>(0.0252)</td>
<td>(0.0240)</td>
<td>(0.0263)</td>
<td></td>
</tr>
<tr>
<td>Transacting Lot Share of Assembled Parcel Size (size_sh)</td>
<td>0.0719</td>
<td>0.1312</td>
<td>-0.0411</td>
</tr>
<tr>
<td>(0.0878)</td>
<td>(0.0851)</td>
<td>(0.0959)</td>
<td></td>
</tr>
<tr>
<td>Dummy - Last Lot Assembled Into Parcel (lst_assm)</td>
<td>0.2822 ***</td>
<td>0.2345 ***</td>
<td>0.2849 ***</td>
</tr>
<tr>
<td>(0.0897)</td>
<td>(0.0839)</td>
<td>(0.0922)</td>
<td></td>
</tr>
<tr>
<td>Dummy - Transaction of Assembled Lots</td>
<td>0.1960 ***</td>
<td>0.1745 ***</td>
<td>0.1465 ***</td>
</tr>
<tr>
<td>(0.0411)</td>
<td>(0.0386)</td>
<td>(0.0488)</td>
<td></td>
</tr>
<tr>
<td>Dummy - Transaction Lot to be Included in Assembled Lots</td>
<td>-0.1148 **</td>
<td>-0.0558</td>
<td>-0.0879 *</td>
</tr>
<tr>
<td>(0.0469)</td>
<td>(0.0457)</td>
<td>(0.0482)</td>
<td></td>
</tr>
<tr>
<td>Years From Transaction to Accepted Development Plan (time_dev)</td>
<td>-0.0236</td>
<td>-0.0244 *</td>
<td>-0.0226</td>
</tr>
<tr>
<td>(0.0140)</td>
<td>(0.0133)</td>
<td>(0.0144)</td>
<td></td>
</tr>
<tr>
<td>Interaction - Lst_asm * Lot Area (00000 sq. ft)</td>
<td>0.1084</td>
<td>0.1490 **</td>
<td>-0.0028</td>
</tr>
<tr>
<td>(0.0689)</td>
<td>(0.0619)</td>
<td>(0.0717)</td>
<td></td>
</tr>
<tr>
<td>Interaction - Lst_asm * time_dev</td>
<td>-0.0263</td>
<td>-0.0397 *</td>
<td>-0.0249</td>
</tr>
<tr>
<td>(0.0246)</td>
<td>(0.0230)</td>
<td>(0.0253)</td>
<td></td>
</tr>
<tr>
<td>Interaction - Lst_asm * size_sh</td>
<td>-0.2740 **</td>
<td>-0.2488 **</td>
<td>-0.2380 **</td>
</tr>
<tr>
<td>(0.1174)</td>
<td>(0.1086)</td>
<td>(0.1208)</td>
<td></td>
</tr>
<tr>
<td><strong>Control Variables</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dummy - Transaction in 1992</td>
<td>-0.2256 ***</td>
<td>-0.1739 ***</td>
<td>-0.2136 ***</td>
</tr>
<tr>
<td>(0.0392)</td>
<td>(0.0372)</td>
<td>(0.0402)</td>
<td></td>
</tr>
<tr>
<td>Dummy - Transaction in 1993</td>
<td>-0.1340 ***</td>
<td>-0.0979 ***</td>
<td>-0.1369 ***</td>
</tr>
<tr>
<td>(0.0434)</td>
<td>(0.0401)</td>
<td>(0.0447)</td>
<td></td>
</tr>
<tr>
<td>Distance to Central (km)</td>
<td>-0.0247</td>
<td>-0.1573 **</td>
<td>-0.0383 ***</td>
</tr>
<tr>
<td>(0.0239)</td>
<td>(0.0798)</td>
<td>(0.0125)</td>
<td></td>
</tr>
<tr>
<td>Dummy - Primarily Residential Building</td>
<td>-0.7197 ***</td>
<td>-0.7512 ***</td>
<td>-0.7557 ***</td>
</tr>
<tr>
<td>(0.0507)</td>
<td>(0.0610)</td>
<td>(0.0514)</td>
<td></td>
</tr>
<tr>
<td>Dummy - Primarily Industrial Building</td>
<td>-0.5411 ***</td>
<td>-0.7563 ***</td>
<td>-0.6091 ***</td>
</tr>
<tr>
<td>(0.1050)</td>
<td>(0.1043)</td>
<td>(0.0997)</td>
<td></td>
</tr>
<tr>
<td>Dummy - Primarily Commercial/Other Building</td>
<td>-0.8403 ***</td>
<td>-0.6846 ***</td>
<td>-0.8684 ***</td>
</tr>
<tr>
<td>(0.1591)</td>
<td>(0.1517)</td>
<td>(0.1654)</td>
<td></td>
</tr>
<tr>
<td>Maximum Allowable Developable Sq.Ft.</td>
<td>0.0261 ***</td>
<td>0.0226 ***</td>
<td>0.0248 ***</td>
</tr>
<tr>
<td>(0.0036)</td>
<td>(0.0034)</td>
<td>(0.0032)</td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>-4.2081 ***</td>
<td>0.0025</td>
<td>0.0036</td>
</tr>
<tr>
<td>(0.1534)</td>
<td>(0.0157)</td>
<td>(0.0164)</td>
<td></td>
</tr>
<tr>
<td><strong>District Fixed Effects</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Specification</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Variable for non-parametric treatment</td>
<td>Parametric</td>
<td>Semi-Parametric location</td>
<td>Semi-Parametric lot size</td>
</tr>
<tr>
<td>Adj R-Squared</td>
<td>0.7791</td>
<td>0.3033</td>
<td>0.6871</td>
</tr>
</tbody>
</table>

Notes:
1. Standard errors are in parentheses. * indicates stat. sig. @ 10% level, ** @ 5% level, *** @ 1% level. All regressions have 717 obs.
2. Coefficient on lot area in regression is estimated non-parametrically and the plots of derived lot prices are shown in Figures 1-3
3. "**" indicates statistics sig. @ 10% level, "***" @ 5% level, "****" @ 1% level.
not present a coefficient for lot area. Later we plot the variation in these coefficients. As in regression (1), regression (3) includes neighborhood fixed effects.

Our results are remarkably robust. With the exception of the effect of lot area in regression (3), the point estimates are quite similar across the regressions, and qualitatively nearly identical. The last lot assembled into a parcel, the lot most likely to experience holdup, enjoys a premium of 23 to 28 percent in the price per square foot. This premium is embodied in the price of parcels sold in their final assembled configuration: they experience a premium of 5 to 20 percent. Lots that will later be combined into a final assembled parcel transact at a discount of 5 to 11 percent.

The estimated coefficients on the control variables are also quite robust. Relative to sites with office properties, those with residential, industrial, and commercial properties sell at a discount. The transaction price is strictly rising in the maximum developable area, which also proxies for the size of the existing structure. It falls with the time to development. The latter is most likely describing the lower value of the development option. Lot price falls with distance to the CBD. All are coefficients reflecting within neighborhood variation because of the fixed effects.

The relationship between lot size and the assembled parcel size is mixed. Controlling for the size of the transacting lot, the unit price falls with the size of the final assembled parcel. At the same time, the conditional effect of the ratio of the lot size relative to the total parcel size is not statistically significant. One problem is that there are alternative
effects at work. The larger the size of the total assembled parcel, the larger the rents that can be extracted, which should boost the premium for an assembled parcel. On the other hand, it may also mean that the developer has more options to pursue in the event that final lot is not assembled. To try to sort out these different effects, we interact the dummy for last lot assembled into the assembled parcel with lot size, time to development, and the ratio of lot size to assembled parcel size respectively. We find that the assembly premium is time invariant, and the premium declines as the size of the holdup premium grows relative to the size of the assembled lot.

We find that the relationship between log lot price per square foot and lot area is not monotonic. The estimated coefficient is negative when lot area enters in a linear parametric fashion in regressions (1) and (2). This result is consistent with the findings in the literature that land prices are concave in lot area. However, in regression (3) we allow for this estimated coefficient to vary with lot size. As described above, for each observation we generate a separate point estimate of the marginal effect of lot size on the log of the indexed unit land weighted by closeness in size. The neighborhood fixed effects controls for the general pattern of spatial variation in size. We plot the regression estimate of the price of land per square foot derived from these coefficients for mean values of the included variables. Figure 1 shows the relationship for the full range of lot size. There is a small region where land prices are rising, before falling. The relationship appears to be linear within each region, though the slopes for the two regions are not identical. To observe the effect without the long-tail, Figure 2 shows the plot for lot sizes up to 19,000 square feet. Here the two regions are seen quite clearly. Within each
Figure 1
Semi Parametric Estimation
Price/SqFt by Lot Size
Figure 2
Semi Parametric Estimation
Price/SqFt by Lot Size
Figure 3
Semi Parametric Estimation
Price/SqFt by Lot Size
section the relationship is monotonic, with the exception of the price drop for lots between 1,500 and 2,000 square feet. The distribution of lot sizes in our data is not uniform, with substantial density for small lot sizes and a long right hand tail. In Figure 3, we plot the relationship with the rank order by size instead of area on the X axis. In this configuration there is a large region where unit land prices are rising.

V. Conclusion

In this paper we show quite clearly the costs of holdup. The inability to write contracts between the Hong Kong government and developers over development force developers to acquire ground leases in order to redevelop lots. This acquisition process gives us the opportunity to observe directly the effects of holdup, both in the premiums charged for the last units assembled and the discounts for earlier units since in equilibrium developer profits should be invariant to the constituent parts of the developed parcels. In doing this analysis, we also find the first evidence of the long assumed convexity between lot value and size. Interestingly, our results show that convexity holds even after controlling for land assembly. We speculate that this result reflects the higher construction costs for development on small lots.

A direction for further research is to determine whether these holdup costs have real effects or are just transfers between land-owners and developers. Keeping total acquisition costs unchanged and using mean lot sizes for the two categories and for the
size of final assembled parcels, the parcels acquired previously must be obtained at a
discount of 13 percent in order for the last parcel to receive a premium of 25 percent. It
is possible, then, that there are areas where no redevelopment occurs because, given rents
for existing structures and construction technology, no lot owners are willing to grant the
discount necessary for the first lots acquired in an assembly. The possible implication of
holdup is that the urban form will be affected by the historic pattern of land holding and
lot definitions.
References


