## From Today's City to Tomorrow's City: An Empirical Investigation of Urban Land Assembly

Leah Brooks Trachtenberg School of Public Policy and Public Administration George Washington University

> Byron Lutz Research Division Federal Reserve Board of Governors

#### May 2015

Because urban areas are fundamentally constrained by the boundaries of land ownership, economic and technological shocks cannot modify the urban landscape without also changing the delineation of land. In this paper, we ask if urban land markets are capable of producing such changes in land delineation. Specifically, we test if there is enough land assembly – the legal joining together of two or more parcels of land – to put land to its highest value use. Failure to assemble land may reduce agglomerative benefits, produce urban blight, push growth to the city edge and away from the core; ultimately it may cause cities to forfeit economic growth. We develop a simple theoretical framework which provides a testable hypothesis: in the absence of market frictions, the price of land sold for assembly should not exceed the price of land sold for other uses. This hypothesis does not hold when frictions, such as holdouts and land use regulation, drive the market for land assembly. We test this hypothesis using a novel dataset that follows each of the 2.3 million parcels in Los Angeles County over a twelve year period and allows us to observe all instances of assembly. We address the potentially significant problem of selection into assembly by comparing, within very small neighborhoods, the price of land for parcels sold into assembly to price of land for parcels sold where the structure is immediately torn down. This method and extensions find that to-be-assembled land trades at a 15 to 40 percent premium. Thus, we find that urban land markets are subject to significant frictions that prevent assemblies and produce sclerotic urban redevelopment. Additional empirical results suggest that private market frictions, such as holdouts, play an important role in blocking assembly.

The analysis and conclusions set forth are those of the authors and do not indicate concurrence by the Board of Governors of the Federal Reserve. We are extremely grateful to staffers Bulmaro Borrero, Rex Hartline, and Alan Santos at the Los Angeles County Assessors office for providing data and guidance while we endlessly assembled the dataset for this paper. We are also appreciative to the Lincoln Institute for Land Policy, McGill University and the John Randolph Haynes and Dora Haynes Foundation, all of which contributed funds towards the construction of the dataset. We thank Gustavo Bobonis, Edward Coulson, Richard England, Lynn Fisher, Edward Glaeser, Robert Inman, Jeffrey Lin, Dan McMillen, Raven Molloy, Kathy O'Regan, Vasia Panousi, Jaren Pope, Hui Shan, William Strange, David Wildasin and participants at the 2010 and 2011 International American Real Estate and Urban Economics Association conferences, 2010 Federal Reserve System Regional Analysis conference, 2010 Lincoln Institute Conference (October), 2010 National Tax Association Annual Conference, 2011 ASSA meetings, 2011 Lincoln Institute Urban Economics and Public Finance Conference (May), 2011 NBER Summer Institute (Law and Economics & Economics of Real Estate and Local Public Finance), 2011 Planning Law and Property Rights (Alberta), the Philadelphia Federal Reserve, the Federal Trade Commission, the Urban Institute, the University of Michigan Economics Department, the University of Toronto Economics Department, and the Federal Reserve Board of Governors. We are very grateful to the planners and developers that shared insights with us and greatly improved the paper; they are listed in the bibliography. This paper would not exist but for the superior research assistance of Brian McGuire. Cities are composed of individual pieces of land called parcels. Just as atoms dictate the properties of matter, parcels dictate the size, shape and placement of the built infrastructure of which cities are composed. Changes in technology and economic conditions cannot induce fundamental changes to this built infrastructure without changes to parcel boundaries. Thus, the long-run evolution of cities – and the economic growth and innovation they generate – depends upon the ease of modifying parcel boundaries.

Changes to parcel boundaries within a city usually involve the assembly of multiple individually-owned parcels into one larger, singly-owned parcel – a process called land assembly. Because land assembly requires geographic contiguity, an individual parcel owner has a form of monopoly power when his neighbor wishes to assemble land (Merrill, 1986). Does this monopoly power create economically meaningful frictions? In this paper we address this question by testing whether the market produces enough assemblies to put land to its highest value use. In doing so, we estimate what is, to the best of our knowledge, the first overall premium to land assembly.

Theorists, economic historians, and practitioners all argue that market imperfections yield "too little" land assembly – and that this failure can hinder economic growth. Whether due to inefficiencies from asymmetric information (Strange, 1995), or positive externalities arising from assembly (Grossman and Hart, 1980; O'Flaherty, 1994), theorists argue that holdout behavior is likely to prevent profitable assemblies.<sup>1</sup> Economic historians contend that the ability to assemble ownership interests is often a crucial pre-requisite for economic growth, citing examples as diverse as how fragmented powers of eminent domain in pre-Revolutionary France inhibited profitable irrigation projects (Rosenthal, 1990), and how the land subdivision boom of the 1920s caused subsequent problems of land assembly and

<sup>&</sup>lt;sup>1</sup>Other theoretical work on land assembly includes Asami (1988), Eckart (1985), Grossman et al. (2010), and Menezes and Pitchford (2004). Cadigan et al. (2011) conduct experimental research on the potential private market inefficiencies in land assembly. Kominers and Weyl (2010), Plassmann and Tideman (2008) and Tanaka (2007) suggest mechanisms for solving the holdout problem.

stagnation in the 1930s (Field, 1992).<sup>2</sup> Practitioners, such as urban planners, concur with this dour assessment and call land assembly the "single biggest obstacle to central city redevelopment" (Nelson and Lang, 2007).

If land cannot be assembled in sufficient quantity, cities will fail to adjust to new economic realities. Most significantly, land assembly allows cities to become denser. Market frictions which inhibit assembly therefore cause land to be misallocated to sub-optimally dense uses.<sup>3</sup> Moreover, cities exist largely due to agglomerative externalities, which increase with density (Duranton and Puga, 2004; Rosenthal and Strange, 2004). Insufficient land assembly, by restraining density, reduces the magnitude of agglomerative effects. Similarly, it may lower economic growth by forcing some workers to reside in less productive cities (Hsieh and Moretti, 2015). Cities which fail to redevelop at the center due to problems with land assembly may instead expand at the edge, yielding congestion and attendant environmental ills (Miceli and Sirmans, 2007). Moreover, problems of land assembly may cause cities to turn to eminent domain: nearly seventy percent of legally contested eminent domain cases involve land which has been seized for use in a land assembly (Merrill, 1986).<sup>4</sup>

"Too little" land assembly also has important consequences for how economists think about cities. The canonical model of urban redevelopment posits that land is redeveloped when the present value of the capital situated on it is exceeded by the present value of redevelopment.<sup>5</sup> If land assembly is required for capital adjustment, but assembly is blocked

<sup>&</sup>lt;sup>2</sup>Bogart and Richardson (2009) contend that the British Parliament's willingness to "assemble" ownership interests in land following the Glorious Revolution of 1688 yielded dividends in economic growth. Rosen suggests that problems of land assembly inhibited redevelopment even after very large shocks, such as the Great Chicago Fire of 1871 (Rosen, 1986). Hornbeck and Keniston (2014)'s empirical work on Boston's 1872 Great Fire confirms this view. Relatedly, differences in systems of land demarcation across the U.S. states (Libecap and Lueck, 2011) and across former British colonies (Libecap et al., 2010) yield divergent economic outcomes. Finally, legal scholars such as Heller (1998) argue that problems of land assembly helped to inhibit redevelopment in Eastern Europe.

<sup>&</sup>lt;sup>3</sup>For instance, as cities grow in population, more capital intensive use of the land is optimal (Henderson, 1977). The inability to assemble land impedes this population driven evolution toward a denser city.

<sup>&</sup>lt;sup>4</sup>The recent *Kelo v. City of New London* case (545 U.S. 469 (2005)), in which the city of New London used its power of eminent domain to assemble land that it then provided to a private developer, brought these issues into focus.

<sup>&</sup>lt;sup>5</sup>Brueckner (1980b), Brueckner (1980a), Wheaton (1982) and Wheaton (1983); see Rosenthal (2008) for

by market frictions, capital decay alone is insufficient to generate redevelopment.

Finally, understanding if and why the market fails to produce land assembly also helps us to understand how the more general problem of fragmented ownership – the "tragedy of the anticommons" – may impede efficient outcomes (Heller, 1998; Fennell, 2004). Other problems of fragmented ownership include the resolution of sovereign debt consolidations (Pitchford and Wright, 2008), and the efficient use of multiple related patented discoveries (Heller and Eisenberg, 1998).

We begin with a simple theoretical framework that considers an area where the parcel boundaries, set at the time of initial development, are no longer optimal. Specifically, we assume that the per square foot price of a large parcel has come to exceed the price of a small parcel (motivated by the greater range of uses available for larger parcels). As a result, there is an economic incentive to assemble land. We also assume free entry into the market for assembly and that, correspondingly, developers earn zero profits and all surplus from assembly goes to the landowners who sell into an assembly. This framework yields the following testable assertion: in a market free of frictions, the price of land sold for assembly should not differ from the price of land sold for other purposes. In a competitive market, any differences in the land price per square foot should be arbitraged away. This test is similar in spirit to Glaeser et al. (2005).

To evaluate this contention, we require a method to value land in isolation from any capital which may be situated on it. We rely on the technique pioneered by Rosenthal and Helsley (1994) and refined by Dye and McMillen (2007) that identifies the price of land from parcels sold shortly before the structure is subsequently torn down. Because the structure is worthless to the new owner, the sale value represents only the land value. Similarly, the value of properties sold just before for assembly should recover the price of land for assembly properties.

empirical work.

To further address the endogenous selection of land into assembly, we compare sales just before teardown to sales just before assembly, conditional on small neighborhood (tract or block group) fixed effects that net out the main component of land value: location. We also perform falsification tests to assess whether land very near teardown properties is priced differently than land near assembly properties; we find no such evidence.

For our empirical work, we construct what is, to the best of our knowledge, the most appropriate existing dataset for this question. We have assembled a panel dataset that traces each of the 2.3 million parcels in Los Angeles County over a 13-year period from 1999 to 2011. The dataset is based on annual cross-sections containing all parcels in the county and a database, provided by the County Assessor, which identifies all instances of changes to parcel boundaries and links each parcel that ceases to exist with its descendant. Our dataset allows us to follow each individual piece of land in the county over this entire period.<sup>6</sup>

We find that assembly land sells at a 15 to 40 percent premium relative to non-assembly land. We show that this premium is not likely driven by the endogenous location of assembly by using a repeat sales specification in addition to the falsification check. We argue that our results are not driven by owners' subjective valuations of their properties (i.e., high reservation prices), by considering the premium by property use and by owner occupancy. Taken together, we interpret our results as evidence that the market for land assembly is subject to substantial frictions.

The frictions we document could result from imperfections in both the public and private spheres. Public frictions arise from the regulation of land by local governments, including zoning restrictions, development fees, and building codes (Glaeser et al., 2005). Private market failures stem from bargaining problems between the developer of the assembled land and the land sellers, causing problems such as holdouts. To establish whether private market

<sup>&</sup>lt;sup>6</sup>Relative to the country as a whole, land in Los Angeles is expensive and highly regulated. In this vein, our results should be interpreted as being most applicable to large coastal metropolitan areas, which tend to be expensive and highly regulated. See a full comparison on page 21.

frictions are an important part of the problem, we perform two additional tests motivated by theoretical conjectures about the relationship between parcel size and both assembly probability and the sales price of assembled parcels. We find suggestive evidence that private market imperfections are substantial.

## **1** Theoretical Framework

## 1.1 Surplus to Assembly

We begin our analysis with a simple theoretical framework that generates testable predictions about land assembly. Assume that at the time of initial development, time t - j, j > 0, land is developed into parcels of size a and 2a. The relative quantities of the different sized parcels are chosen optimally given market conditions at time t - j such that the per-square foot price of parcels size a and size 2a is equal. As time passes, those initial parcel definitions may become sub-optimal. In this vein, we assume that by time t, market conditions have evolved such that larger parcels size 2a command a premium relative to smaller parcels sized a. In other words, the relationship between land value and parcel size has become convex, such that V(2a) > 2V(a) where V(x) is the market price of a parcel of size x. Consistent with our empirical approach, which focuses on the value of land and not capital, we assume that any capital placed on the land at time t - j has depreciated to zero by time t. Figure 1a graphically presents the convexity of the land price function, V. The price per unit a of land increases with the size of the parcel. In contrast, the thin dashed line, ending at 2V(a), illustrates a linear land value function where the price per unit a of land is constant.

Land values tend to become convex when the optimal capital to land ratio increases. Convexity arises because the density implied by high capital to land ratios requires large lots. For example, taller buildings typically require larger footprints. Similarly, builders may require large lots in order for buildings to be of sufficient size to absorb fixed costs such as elevators.

The optimal capital to land ratio in a metropolitan area may increase for any number of reasons. Population growth tends to increase the optimal capital to land ratio (Henderson, 1977). Similarly, increased commute times in an urban area may push the optimal ratio up in the urban core. The optimal ratio may also rise when land use shifts geographically over time. For instance, land initially developed into small single family lots may eventually become more valuable for commercial purposes. Commercial uses typically imply a higher capital to land ratio, and thus larger lot sizes. Finally, technological shocks may also change this ratio. For example, Willis (1995) describes how the invention of fluorescent lighting changed the shape of office buildings and thus increased the optimal lot size for such buildings.

The convexity in the land value function implies that assembly generates a surplus relative to maintaining existing parcel boundaries. We define surplus value, s, as

$$s = V(2a) - \delta - 2V(a), \tag{1}$$

where  $\delta$  is the cost of assembly and captures factors such as conversion costs (e.g., demolition, grading to-be assembled parcels with different slopes, etc.) and "good-institution" transactions costs. Crucially,  $\delta$  only includes costs which would reasonably arise in a wellfunctioning land market free of frictions. We use the term "good-institution" to sharply distinguish transactions costs consistent with a well-functioning land market from "badinstitution" transactions costs better viewed as frictions. "Bad-institution" transactions costs might include holdouts. For example,  $\delta$  includes the cost of changing title to a property, but not delays to change in title caused by protesting neighbors.

Convexity in the land price function is a necessary condition for land assembly to occur (Shoup, 2008). The cost of assembly,  $\delta$ , can only be covered when the value of the assembled parcel, V(2a), exceeds the value of the unassembled parcels, 2V(a). The value of the

unassembled parcels, 2V(a), is the opportunity cost of assembly; it represents the economic returns to the unassembled land foregone in exchange for realizing the higher return to the assembled land.

In a frictionless world, arbitrage ensures that all surplus is realized and that assemblies continue until the market price of land has adjusted such that any surplus is eliminated. Specifically, as assemblies occur the supply of lots sized 2a expands and the price of these lots falls. Assembly ceases when the return to assembled and unassembled lots has equalized:  $V(2a) - \delta = 2V(a)$  and s = 0.7 Figure 1b—a supply and demand graph for parcels sized 2a—displays this process. For simplicity, we assume that the quantity of parcels sized a is arbitrarily large relative to the demand for parcels sized 2a. As a result, V(a) is fixed and the supply curve for parcels sized 2a is horizontal at the marginal cost of producing such parcels:  $2V(a) + \delta$ . In the absence of frictions, the market will assemble land until there are q' parcels sized 2a. When the market reaches equilibrium following the period of assembly activity, the land value function has flattened out from V to V' in Figure 1a and the price difference between parcels sized a and 2a is the cost of assembly,  $\delta$ .

If frictions are present, the surplus available from assembly may not be arbitraged away. For instance, holdouts may ask excessive prices for their parcels and thereby make projects infeasible for the developer (Eckart, 1985; Strange, 1995).<sup>8</sup> Similarly, individual landowners may attempt to increase their share of the surplus by being the final sale into an assembly. Such strategic delay may cause assemblies to fail (Menezes and Pitchford, 2004; Miceli and Segerson, 2007; Miceli and Sirmans, 2007). The public goods aspect of land assembly – the fact that assembly may increase the value of neighboring parcels not participating in the

<sup>&</sup>lt;sup>7</sup>It is also possible that the market reaches a corner solution such that all parcels have been assembled. As this is not relevant empirically, we do not further consider this case.

<sup>&</sup>lt;sup>8</sup>In Eckart (1985) and Strange (1995), holdouts arise primarily due to two factors. First, land owners may have imperfect information on the value of the completed assembly and may therefore demand prices in excess of the value of the completed project. Second, assembly is characterized by indivisibility. A developer must assemble contiguous parcels and therefore cannot simply purchase an alternative parcel if he encounters a holdout.

assembly – may also block arbitrage opportunities (Grossman and Hart, 1980; O'Flaherty, 1994). Finally, land use regulations may systematically block arbitrage opportunities. For example, regulation may bar a large building that would optimally occupy an assembled site. Returning to Figure 1b, if frictions block all assemblies, there will only be q parcels of size 2a (the initial endowment) and the surplus to assembly, s, will persist. In the absence of spillovers, the shaded region is the deadweight loss associated with the frictions.<sup>9</sup>

## **1.2** Testing for Frictions

We now lay out a strategy to assess whether or not land assembly is inhibited by frictions.<sup>10</sup> Our test estimates the magnitude of the surplus s accruing to successful assemblies in order to obtain a rough sense of the magnitude of the assembly market frictions. A large estimate of s is consistent with substantial frictions in the market for assembly. This approach is similar in spirit to the work of Glaeser et al. (2005) and Glaeser and Gyourko (2002) on the regulatory tax. Glaeser et. al. reason that in the absence of regulation the extensive value of land – the per-unit value of land with a house on it – will equal the intensive value of land – the value of a marginal increase in the area of a lot. If the extensive value exceeds the intensive value, landowners should optimally choose to subdivide their land and sell a portion of it. Our approach applies what is, in essence, the reverse of this logic: in the absence of market imperfections, if land is worth more combined than divided, owners will choose to combine it.

 $<sup>^{9}</sup>$ We only consider the effect of frictions on the *quantity* of assembled parcels. Frictions may also cause the final land area assembled to be smaller than it would otherwise be. Munch (1976) and McDonald (2007) model this possibility.

<sup>&</sup>lt;sup>10</sup>The extremely limited empirical evidence on land assembly does not directly addresses the central question of how successfully the land market produces assemblies. Cunningham (2010) and Fu and Somerville (2002) document that the final seller in a land assembly receives a premium. Harding et al. (2003) provide evidence of more general deviations from the competitive equilibrium in the market for housing. A distinct literature explores whether the price per square foot of land increases in the size of the plot and is indirectly related to assembly (Coulson, 1989; Colwell and Munneke, 1999, 1997; Colwell and Sirmans, 1993; Brownstone and DeVany, 1991).

Inferring the existence of frictions from the presence of surplus requires two assumptions. The first assumption is free entry into the market for development (or assembly). Developers earn zero profits; therefore the owners of the initial parcels size a capture any surplus s available from assembly. The value of an assembled parcel is V(2a) + K, where K is the amount of capital placed on the newly assembled parcel. If the developer earns zero profits, this post-assembly value must equal his costs. The developer's costs are capital, K, assembly costs,  $\delta$ , and the purchase price of the unassembled land,  $p_u$ . Thus,  $V(2a) + K = K + \delta + p_u$ which yields  $p_u = V(2a) - \delta$ . We can therefore estimate surplus as the difference between the sales price of to-be-assembled parcels,  $V(2a) - \delta$ , and the sales price of not assembled parcels, 2V(a):  $V(2a) - \delta - 2V(a)$ .<sup>11</sup>

The second assumption required to infer the existence of frictions from the presence of surplus is that the frictions in the urban land market operate purely as a supply constraint on assembly. This occurs if regulation prohibits assembly or if landowners cause assemblies to fail by asking prices that drive developer profits below zero. These supply restraints prevent arbitrage from entirely eliminating the surplus to assembly.

If the free entry and supply constraint assumptions fail to hold we will likely *understate* frictions in the market for assembly for at least two reasons. First, if there are barriers to entering the market for development, developers may capture a portion of the assembly surplus. The portion of the surplus accruing to the developer is reflected in the post-assembly sales price of the newly assembled parcel, not in the pre-assembly price we use to infer surplus. As a result, our estimate of surplus would be biased downward. Second, although the frictions described above almost certainly act as supply constraints on assembly, other types of frictions may exist as well. For instance, frictions such as regulatory costs (e.g., the time spent getting approval for a project) and strategic delay may increase the developer's costs. Such an increase will reduce our measured value of surplus by depressing the amount

<sup>&</sup>lt;sup>11</sup>We deliberately remain agnostic on how sellers of input parcels split the surplus from assembly.

a developer is willing to pay for assembled land. A given assembly may have no measurable surplus under our methodology, but large surplus in the absence of the friction-induced increase in developer costs. Thus, even a finding of no price premium for to-be-assembled parcels does not rule out the possibility that frictions influence the market for assembly.

We now discuss four potential conceptual objections to our test. (We leave the important issue of empirical bias arising positive selection into assembly and other endogeneity concerns to Section 2.) First, a strict claim of inefficiency requires that surplus exceeds zero for the marginal assembly, while our empirical procedure recovers the average surplus to assembly. In our simple framework this is not a concern because there is no heterogeneity and the average and marginal surplus equate. In reality, though, there may be heterogeneity in the costs of assembly (keeping in mind that "good-institution" transactions costs are not driven by market frictions). With heterogeneous assembly costs, average surplus can exceed zero even if the surplus to the marginal parcel is zero.<sup>12</sup>

This type of heterogeneity in assembly costs, however, is unlikely to explain our findings for two reasons. First, we find that average surplus equals 15 to 40 percent of the market value of a parcel in its unassembled state. Assembly costs – demolition, land grading, "goodinstitution" transaction costs, etc. – are negligible relative to the value of some of the most expensive urban real estate in the U.S.<sup>13</sup> For heterogeneous assembly costs to explain our estimates, such assembly costs would have to be extremely large relative to the value of the land in its unassembled state. Second, the distribution of assembly costs would have to be

<sup>&</sup>lt;sup>12</sup>Assume that there are contiguous pairs of parcels with unique assembly costs,  $\delta_i$ , that there are no frictions in the land market, and that the market is in equilibrium (and hence no assembly takes place because all positive surplus assemblies have already occurred). Then a shock increases V(2a). Order the surplus of each pair,  $s_i = V(2a) - \delta_i - 2V(a)$ , so that  $s_1 > s_2 > s_3$ .... Unless all parcels assemble, there must be an  $s_{i^*}$  such that  $s_{i^*-1} > s_{i^*} = 0 > s_{i^*+1}$ . The index  $i^*$  implicitly defines the number of assemblies produced by the shock. The marginal parcel  $i^*$  earns zero surplus, and all inframarginal parcels earn a surplus. We thank Robert Helsley for helpful discussion on this point.

<sup>&</sup>lt;sup>13</sup>Rosenthal and Helsley (1994) document that demolition costs are only around  $1\frac{1}{2}$  percent of the value of the underlying land for single family homes in Vancouver (which, like LA, has relatively expensive land). Similarly, Dye and McMillen (2007) document that demolition costs are low enough in metro Chicago that they can effectively be ignored for purposes of using teardowns to establish the value of land (see pg. 48).

extremely disperse to generate an average surplus of 40 percent and a marginal surplus of zero.

Similarly, there may be variation across locations in the difference between the value of a parcel in its assembled state, V(2a), and its unassembled state, V(a), and hence in the surplus to assembly. However, in the absence of frictions, arbitrage should drive the assembly surplus to zero in all locations regardless of the initial surplus available. Were the outputs of assembly projects unique local properties, this might not be the case: Imagine a developer who assembles land to build a supermarket in an area that can support only one such store. However, our data show that over three-quarters of land assembled and developed is used for residential purposes. Such residential development is not the unique output that might give rise to this concern.

The second potential objection to our test is that it may take time for the market to complete enough assemblies to drive surplus to zero. During such a transition, surplus would be positive, even in the absence of frictions. However, the market for construction is quite deep in Los Angeles County. Many developers participate<sup>14</sup> and even large construction projects can typically be completed quickly. The Census Bureau reports that a small multifamily building worth less than \$3 million is completed in less than a year on average. Even very large buildings valued over \$10 million with over 100 units usually take less than two years to complete.<sup>15</sup> Thus, any transition period should be relatively short and is unlikely to explain a 15 to 40 percent surplus to assembly. While assemblies often take a significant amount of time to complete, the delay is typically caused by the very frictions, such as holdouts, whose impact we are attempting to quantify, not by the normal adjustment process of a well functioning land market.

The third possible objection to our surplus test is that many plots are constrained from

 $<sup>^{14}</sup>$ For example, the commercial developers association in Los Angeles has over 1000 members.

 $<sup>^{15} \</sup>rm http://www.census.gov/const/C30/length.html$ 

assembly. For instance, physical barriers such as steep slopes may prevent assembly. Public capital, such as roads, may separate parcels and prevent assembly. A parcel ready for redevelopment may be next to a parcel with new or valuable capital, making redevelopment as part of an assembly economically infeasible. These factors are reasonably viewed as materially different from factors such as zoning and holdouts which may also prevent assembly. It would be unreasonable to label the failure to assemble two parcels separated by a road a market "friction." However, arbitrage opportunities should cause the price of assembled and unassembled teardown parcels to converge – and surplus to go to zero – unless there is a corner solution where no feasible assemblies exist. At least in Los Angeles, it seems clear that ample assembly opportunities remain and that a corner solution has not been reached (Landis and Hood, 2005).

The fourth potential objection to our test is that potential sellers into an assembly may have reservation prices higher than the market price for idiosyncratic reasons. For example, retirees may prefer to remain in their long-term place of residence or parents may wish that their children continue to attend a particular school. The need to buy contiguous parcels may force a developer to transact with such owners. (In contrast, in most market settings a buyer would choose to transact with a seller whose reservation price is less than or equal to the market price.) Thus, high reservation prices may cause properties selling into an assembly to transact at above market prices even in the absence of market frictions (Miceli et al., 2008).<sup>16</sup>

We believe that reservation prices are unlikely to explain our surplus estimates for two reasons. First, while individuals with elevated reservation prices undoubtedly exist, the surplus to assembly should be arbitraged away as long as there are *some* potential assemblies

<sup>&</sup>lt;sup>16</sup>To illustrate, assume that the developer can always buy the first parcel of an assembly at market price and that the reservation price of the owner of the second parcel, R(a), exceeds the market price: R(a) > V(a). The true surplus from assembly becomes  $s = V(2a) - \delta - 2V(a) - (R(a) - V(a))$ . Reservation prices are unobservable and the measured surplus remains  $s = V(2a) - \delta - 2V(a)$ . The surplus estimate is therefore biased *upward* by the amount by which the reservation price exceeds the market price: R(a) - V(a).

where the owners have reservation prices equal to market prices (that is, as long as there are enough such potential assemblies that surplus is driven to zero before developers must begin transacting with high reservation price sellers). Second, our empirical work devotes considerable effort to determining the extent to which our surplus estimates reflect high reservation prices. We use a number of different approaches, all of which suggest our surplus estimates primarily reflect market frictions, not high reservation prices. Most notably, although owners of commercial property are likely less prone to having idiosyncratically high reservation prices (Ellickson, 1973), these properties display the same surplus to assembly as do residential properties.

#### **1.3** Sources of Land Assembly Frictions

In the presence of non-zero surplus, we propose two tests to examine the source of frictions in the market for land assembly. For the first of these tests, assume that at time t parcels of size  $\frac{a}{2}$  and a exist. Assembly technology allows for generating parcels of size 2a through any combination of parcels yielding an area of 2a. The convexity of the land value function suggests that  $V(a) > 2V(\frac{a}{2})$ . The assembly surplus is therefore higher for assemblies involving only parcels of size  $\frac{a}{2}$  because these smaller parcels have a lower opportunity cost than the larger parcels. As a result, in a market free of frictions, small parcels should be more likely to be assembled than larger parcels.<sup>17</sup>

However, theorists argue that small parcels may increase the likelihood of private market frictions. Owners of small parcels may be more likely to ask excessive prices because, as a lower share of the ultimate assembly, their asking price is less likely to scuttle the project (Eckart, 1985; Strange, 1995). Similarly, the greater number of parcel owners involved in

<sup>&</sup>lt;sup>17</sup>Assembly costs,  $\delta$ , may be a function of the number of parcels included in an assembly (e.g., real estate transaction fees would be expected to be a function of the number of parcels). In such a case, we implicitly assume that as the number of input parcels increases, holding the size of the end assembly fixed, the opportunity cost of the input parcels falls more than the increase in assembly costs (roughly  $\frac{\partial \delta(n)}{\partial n} < -\frac{\partial V(n)}{\partial n}$ ). This is a reasonable assumption given that assembly costs are small relative to the value of land.

an assembly, the greater the odds of strategic delay (Miceli and Sirmans, 2007). Both these factors may cause assemblies with positive surplus to fail.

Given this, our first test for the sources of surplus examines the influence that parcel size has on the probability of assembly. A finding that larger parcels are more likely to be assembled suggests that private market frictions inhibit assembly.

Our second test for the sources of surplus examines the relative sales prices of to-beassembled parcels by size. Evidence that small parcels command a significant per-square foot price premium over large ones, despite a lower opportunity cost of selling into assembly, would support the theoretical predictions that small parcels owners are unusually likely to ask excessive prices and engage in strategic delay. These two additional tests shed light on the source of the frictions, as both focus on frictions likely caused by private market failures.

# 2 Empirical Approach: Establishing the Existence of Surplus

In this section, we outline how we take the surplus test to the data. In the results section, we discuss robustness checks, including repeat sales methods, a falsification check, and a consideration of the role of reservation prices.

To estimate surplus, s, we must recover the land value of both assembled and unassembled parcels. We recover the land value of an unassembled parcel using the technique pioneered by Rosenthal and Helsley (1994) and refined by Dye and McMillen (2007). This technique recovers the value of land using sales of properties where the structure is torn down shortly after sale. The value of such a teardown sale reflects only the value of the underlying land, since the capital is discarded and demolition costs are generally very small relative to land values (see footnote 13 for information on costs).

We apply a similar logic to value the land used in assemblies. Most assemblies discard the

existing capital and place new capital on the assembled site to take advantage of the larger building area. Intuitively, if the capital on the initial parcels were retained, there would be no gain from assembly. As a result, we assume that to-be-assembled parcels are teardowns and also use their sales price as a measure of the value of the land.<sup>18</sup> We then estimate surplus as the difference between the value of assembly sales and the value of teardown sales.

We estimate this difference with the following specification

$$\log(\frac{\text{real sale price}}{\text{lot square footage}})_{i,g,t} = \alpha_0 + \alpha_1 \text{assembly}_i + \text{year-quarter}_t + \text{neighborhood}_g + \alpha_2 X_{i,t} + \alpha_3 \text{amenities}_i + \varepsilon_{i,g,t}, \qquad (2)$$

where  $\frac{real \ sale \ price}{lot \ square \ footage}_{i,g,t}$  is the per-square foot price of land for parcel *i* in neighborhood *g* at time *t* and  $assembly_i$  equals one for an assembly and zero for a teardown.

The estimation sample includes only teardown and assembly sales. Specifically, we select the sales sample in the following way. Each assembly has a start and end year (which may be the same, but frequently are not). We identify all parcels which are "input" parcels to assembly. We then identify the sales of these parcels that are no more than four years before the beginning of the assembly, no more than three years after the beginning of the assembly, and before the end of the assembly. Of these, we keep only the final sale. We use this window to keep only sales "close" to the time of assembly, which should be sales that reflect the value of land. For teardown sales, we keep only sales that are four or fewer years before the structure is torn down.<sup>19</sup>

With this sample,  $\alpha_1$  captures the surplus to assembly relative to redevelopment within the existing parcel boundaries measured by teardowns. If  $\alpha_1 = 0$ , there is no surplus to assembly, consistent with an absence of market frictions.

<sup>&</sup>lt;sup>18</sup>As demonstrated below, however, our results are robust to limiting the assembly sample to cases where we can observe that the pre-assembly capital is torn down.

<sup>&</sup>lt;sup>19</sup>Our results are robust to the use of alternative pre-assembly and pre-teardown windows.

Interpreting  $\alpha_1$  as the surplus to assembly requires that assemblies and teardowns are comparable in the unobserved determinants of price,  $\varepsilon_{i,g,t}$ . However, there are reasons to suspect that assembly may be correlated with unobservable determinants of price. Most significantly, developers choose parcels of land for assembly and may prefer relatively more valuable land. For instance, rising land values often dictate increasing the capital-to-land ratio and may therefore motivate assembly. For our estimation, we are concerned with whether positive selection into assembly exceeds positive selection into teardown.

To illustrate the bias from potential positive selection into assembly a bit more formally, consider neighborhoods H and L which are initially identical and have linear land price functions such that there is no land assembly:  $V^H(2a) = 2V^H(a) = 2V^L(a) = V^L(2a)$ . Then a positive demand shock hits area H, producing two changes. First, parcels of a given size in H are now more valuable than identically sized parcels in area L. Second, the increased demand in H creates the need for larger lot sizes and, hence, the land value function becomes convex and there is excess return to land assembly:  $V^H(2a) > 2V^H(a) > 2V^L(a) = V^L(2a)$ . The true economic surplus to assembly is  $s = V^H(2a) - \delta - 2^H V(a)$ . In the context of our empirical approach, it is possible that we observe teardowns in area L, and assemblies in area H. If so, the empirical estimate of surplus,  $\hat{s}$ , is biased:  $\hat{s} = V^H(2a) - \delta - 2^L V(a)$ . The magnitude of the bias is  $\hat{s} - s = 2(V^H(a) - V^L(a))$ . We view this positive selection into assembly relative to teardown as likely the most important potential source of correlation between assembly and  $\varepsilon_{i,g,t}$ .

A parcel's pre-existing zoning is one specific example of a basis for selection into assembly. For instance, assembly may be more likely on permissively zoned parcels of land because such zoning permits the type of large, dense projects that make assembly valuable. If more lenient zoning boosts the value of both assembled and unassembled parcels, and assemblies are more likely than teardowns to occur in leniently zoned areas, then our estimates could be biased upward by positive selection into assembly.<sup>20</sup> On the other hand, zoning may make high value land less likely to assemble. If high value land tends to be more stringently zoned, assembly may be more likely than on less stringently zoned low value land. High land value may be high value for reasons having nothing to do with the likelihood of assembly, and this could introduce downward bias. Thus, zoning may introduce positive or negative bias.

We tackle these endogeneity concerns by observing that the value of land, virtually by definition, is a function of location. We therefore include a very fine set of geographic fixed effects, either census tract indicators or census block group indicators,  $neighborhood_g$ .<sup>21</sup> The comparison of land price between assemblies and teardowns therefore comes only within very small areas. There are 2,054 census tracts and 6,346 block groups in Los Angeles County. The median tract contains 985 parcels, and the median block group 290.<sup>22</sup>

Of course, some elements of location vary even within small geographic areas. For instance, access to a highway may differ within a neighborhood. We therefore control for distance to the nearest major highway, nearest major road, and the shortest distance to urban and commuter rail with the *amenities<sub>i</sub>* vector. To control for the market-wide evolution in price over time we include a full set of indicators for each quarter in our sample,

<sup>&</sup>lt;sup>20</sup>Positive selection bias from lenient zoning will only occur if two conditions hold. First, lenient zoning must be present on assembled sites but not teardown sites. Second, lenient zoning must increase the value of *both* assembled and unassembled parcels. This need not be the case. For instance, lenient zoning may increase the value of a large assembled parcel because it allows for building a more dense, more valuable building, but not increase the value of a small parcel because you cannot build a large condo building on a row house lot even if it is zoned for it. In this case, the teardown control group provides an appropriate value for an assembled parcel in its counterfactual, unobservable unassembled state.

 $<sup>^{21}</sup>$ If this strategy is valid, the average variation in land price should decline as we consider smaller geographies. Using our teardown sample to isolate the value of land, we find that the standard deviation of log sales price is 28 percent of the mean at the county level. In contrast, when we take the mean and standard deviation of log price by tract and then average across all tracts, we find that the tract standard deviation is 18 of the mean. At the block group level, this figure falls to 16.

<sup>&</sup>lt;sup>22</sup>Our test relies on the co-existence, in geographically comparable areas, of both teardown and assembly parcels. Appendix Figure 1 shows census tracts in Los Angeles County, and marks areas that ever (1999 to 2010) have both at least one assembly and one teardown; areas that ever have only teardowns; areas that have only assemblies; and areas that have neither. Roughly three-quarters of all tracts have an assembly over the period of study; this figure is 80 percent for tracts with teardowns. Roughly 65 percent of all tracts have at least one assembly and at least one teardown over the period of study. The map shows that these redeveloped areas are located widely across the county.

year-quarter<sub>t</sub> (i.e., indicators for 1999Q1, 1999Q2, etc, are included) and we allow the coefficients on these indicators to vary by the residential or non-residential status of the input parcel.<sup>23</sup> To control for parcel attributes, we include a vector of parcel characteristics,  $X_{i,t}$ , which includes a cubic in lot size and indicators for the pre-assembly or pre-teardown use of the parcel (e.g. single family). "Use" is the current use of the parcel, and may in practice be distinct from the parcel's zoned use.<sup>24</sup> Finally, we report standard errors clustered at the city level; Los Angeles County has 88 cities.<sup>25</sup>

# **3** Land Assembly Definition and Institutions

This section presents our empirical definition of land assembly and discusses the institutions for assembly in Los Angeles County.

To define land assembly, we first define a "parcel change group." A parcel is part of a parcel change group if it or its ancestor or descendant parcel(s) ever changes. However, parcels combine and disaggregate in a number of different ways, so membership in a parcel

<sup>&</sup>lt;sup>23</sup>Pooling different property types together in equation (2) is appropriate because we are valuing land, not structures. Furthermore, it is quite common for the input parcel use to differ from the assembled parcel use (e.g., commercial properties are assembled and a condo building placed on the new parcel).

<sup>&</sup>lt;sup>24</sup>The use classifications we use are single family, non-condo multi-family, condo, vacant and other. Parcels can have a use classification which differs from their zone code. For instance, a commercially zoned property is almost always allowed to host a residential dwelling. Such a property will have use and zoning codes which differ. Unfortunately, we do not have reliable data on zoning for our entire sample.

<sup>&</sup>lt;sup>25</sup>Our approach departs from the teardown literature in one important regard. The teardown literature controls for selection into redevelopment using a Heckman two-stage procedure. Accordingly, coefficient estimates yield marginal effects for the untruncated, latent dependent variable. When the goal is to recover the value of all land, whether redeveloped or not, this is an appropriate approach. Because our aim differs, we estimate equation (2) using OLS. Two points bear emphasis. First, the OLS coefficients recover marginal effects for the truncated, observed dependent variable. These effects are precisely those required for the primary test. Specifically, the OLS estimate of  $\alpha_1$  answers the question: among redeveloped parcels (i.e., in the observed portion of the distribution), is there excess value to assembly? The arbitrage argument underlying the test is only valid for parcels actually undergoing redevelopment. Arbitrage should not eliminate the surplus to assembly for parcels whose condition makes redevelopment suboptimal (e.g., those containing capital too valuable for redevelopment). Second, although the conditional expectation function is non-linear, OLS provides a well-defined minimum mean squared error linear approximation. Angrist and Pischke (2009, page 102-3) and Cameron and Travedi (2005, page 542) discuss using OLS to fit the conditional expectation function is non-linear.

group does not alone indicate assembly. The list below presents three examples of possible parcel changes.

- A.  $3 \rightarrow 1 \rightarrow 7$
- B.  $1 \rightarrow 5$
- C.  $8 \rightarrow 1 \rightarrow 2$

We define a change group to have land assembly if any part of the parcel change group goes from n > 1 properties to one. In the example above, this includes cases A and C, but excludes case B. Case A is an instance of net disassembly (3 parcels into 7 parcels), though it contains assembly as an intermediate step (3 parcels into 1 parcel). We define case A as a land assembly because in most such instances redevelopment cannot occur without the assembly step. For instance, a developer may wish to convert multiple detached single family home parcels into denser townhouse style homes. The lots first need to be combined before they can be properly subdivided into the smaller townhouse lots.<sup>26</sup>

When is the type of land assembly we have defined required? Regulations for land assembly are the province of cities, or the county for unincorporated areas. In general, cities do not require developers to assemble parcels, even when a new structure spans more than one parcel. Importantly, legal assembly has no implications for zoning. The prior zoning on the land as previously delimited continues to apply. Legal assembly is the jurisdiction of the county; zoning regulations are the province of individual municipalities.

There are two circumstances under which formal assembly is required.<sup>27</sup> The first is when the new land use will be condominiums. Each unit in a condominium must have a separate parcel, as each unit may have a unique legal owner. Therefore, any land combined for

<sup>&</sup>lt;sup>26</sup>Though this definition of land assembly could include parcels assembled via eminent domain, our empirical analysis of prices uses only arm's length sales. Furthermore, the municipal officials with whom we spoke suggested that eminent domain is used very sparingly.

<sup>&</sup>lt;sup>27</sup>Information in this section comes primarily from an interview with Wolfgang Krause, Chief Planner, City of Glendale, May 2010. To the best of our knowledge, these two circumstances do not vary across regulatory jurisdictions in the county.

condos must be assembled. The second exception from the laissez-faire policy is a function of the use of the property. Suppose that a city's zoning requires two parking spaces for each multi-family unit, and that the developer has purchased two parcels upon which to build a new multi-family development. If the developer builds the parking on one parcel and the structure on the other, he is required to legally assemble the parcels. Cities make such a requirement to ensure that all future sales keep parcels in compliance with zoning regulations.

Outside of these two exceptions, a developer may purchase adjoining land with the intent of building new structures but not go through the formal process of legal assembly.<sup>28</sup> It is the legal assembly which we observe in the data. For the purposes of our estimation, this type of underreporting likely biases our estimates of the prevalence of frictions downward, by misclassifying some assembled parcels—the treated group—as teardown parcels—the control group.<sup>29</sup>

While there are substantial benefits to legal assembly, the costs in terms of administrative burden and fees are extremely low. Once an owner acquires adjacent parcels, if he wishes to legally combine them, he need only fill out a form from the County Assessor's office. There is no charge. Importantly, the legal land assembly process does not trigger a reassessment under California's Proposition 13. Proposition 13 limits the increase in a property's assessed value to two percent per year, with the assessed value resetting to the market value at sale. A developer may face an increased assessment due to property purchase, but not due to the legal act of assembly.<sup>30</sup> Thus, legal assembly is virtually costless. Turning to the benefits of assembly, interviews with practitioners suggest that legal assembly is very likely for projects

 $<sup>^{28} \</sup>mathrm{In}$  such cases, with multiple owners, owners usually write legally binding easement agreements across properties.

<sup>&</sup>lt;sup>29</sup>However, the bias from misclassification could be positive if (a) more valuable parcels were selected into legal assembly, and less valuable parcels were selected into informal assembly and then misclassified as teardowns and (b) such misclassified assemblies account for a substantial share of teardowns.

<sup>&</sup>lt;sup>30</sup>This information comes from the Special Investigations Section, Los Angeles County Assessor.

requiring financing, since assembly lessens the financier's cost in the case of default (interview citations appear after the bibliography). In addition, selling an assembled parcel, rather than multiple unassembled parcels, reduces paperwork and uncertainty in future transactions. Thus, our interviews suggest that the combination of substantial benefits and low costs results in developers choosing legal assembly over informal assembly in most cases.

To give a sense of the prevalence of assembly, Figure 2 presents the total number of assemblies in Los Angeles County divided by the total number of permits issued for residential construction by year from 1999 to 2010. We expect that this number actually underestimates the importance of land assembly. While one permit allows for construction of one unit, one assembly may (and frequently does) result in more than one unit. For instance, three parcels of land may be assembled to construct a multi-story 30-unit condo building. Thus, while three parcels are assembled, permits are given for 30 units. It is also possible that this number overstates the impact of assembly, since permits include only residential construction and our data include all land for assembly. However, the majority of assembly land (like all land) is used for residential purposes. Even given this likelihood of underestimation, assemblies account for roughly 7 to 20 percent of residential permits per year from 1999 to 2008 and account for a substantially larger share in 2009 and 2010.

This evidence from Los Angeles County is not obviously generalizable to the US as a whole. Prices in the county are substantially higher, recent prices changes are somewhat larger (see Appendix Figure 2 for details), and regulation is more stringent.<sup>31</sup> Given this, our results are likely most applicable to large coastal metropolitan areas, which tend to be both expensive and highly regulated. However, these metros account for a substantial share of the U.S. population and economic activity, and are critical centers of innovation (e.g. Glaeser (2011)). Indeed, 1 in 30 US residents lives in Los Angeles County. Further, if

<sup>&</sup>lt;sup>31</sup>Los Angeles is between the 75th and 90th percentiles of regulatory intensity using the Wharton Land Use Regulation Index (Gyourko et al., 2008).

current trends in land use regulation continue, the rest of the country will eventually look more like today's Los Angeles than vice versa. Thus, the County's experience can also be interpreted as a harbinger.

## 4 Data

Our project relies on multiple sources of data. We summarize the data here, and refer interested readers to our lengthy Data Appendix for full details on all data inputs and data construction details. The three key components of our data are the annual property-level data for Los Angeles County, sales data for properties, and census neighborhood measures.

Our annual property data consist of three key parts: (i) twelve annual cross-sectional observations of the 2.3 million parcels in the County of Los Angeles, (ii) an administrative dataset from the County Assessor listing all parcels that change, and the identification numbers of the parcel(s) to which they change, and (iii) electronic maps with geographic information on all properties.

The annual cross-sections are the heart of the dataset. In each year from 1999 to 2011 (except for 2003, which we have been unable to obtain) we observe attributes about each individual piece of property in the 88 cities and the large unincorporated area of Los Angeles County. We observe too many attributes to list here, but briefly the data include attributes about the property itself (e.g., size, location, and current use) and attributes about the building on the property (e.g., building size). Thus, this part of the dataset includes somewhat more than 24 million observations with many descriptive variables.

The second part of the data is a file that allows us to take the 12 cross-sections and make them into a true panel by linking property identification numbers over time. Though most properties retain a constant identification number throughout the sample, some properties split or merge. Our dataset of all property identification number changes allows us to follow each initial piece of land to its current, perhaps aggregated or disaggregated, form. While this task is conceptually simple, in practice it has been exceedingly difficult, and the bulk of our data assembly has been devoted to making sure that we have built these linkages correctly.

The third and final part of the annual property data is electronic maps of all parcels. These maps, which we have from 2006 onward, allow us to pinpoint the exact location of each individual property and calculate distances from one property to another, or from a given property to key urban amenities, such as light rail stops or freeways. These maps also allow us to assign each property to a unique census block group.

We combine this panel of all properties in the county with all property transactions by property identifier. Specifically, we observe the last three sales on each property as of 2006, and sales in the last two years each year from 2009 to 2011. This leaves a small gap of sales in 2006. We limit the sample of transactions to include only arms' length transactions and make other small adjustments as described in the Data Appendix.

We measure neighborhood economic and demographic factors with data from the 1990 and 2000 Decennial Censuses at the block group level. To use the 1990 block group data, we use GIS mapping to make a correspondence from 1990 to 2000 census block groups.

Finally, we define a property as a teardown if the structure's age changes in our panel. Specifically, we require that the replacement structure be newer than the old structure, that the new structure is built after 1998, and that the old structure was built before 1990.<sup>32</sup>

We identify approximately 40,000 parcels that are inputs into an assembly and 14,000 parcels with structures that are town down. Of these properties, there are 2,700 sales of assemblies and 6,800 sales of teardowns that are "close" to the time of assembly or teardown (see our definition of "close" in Section 2).<sup>33</sup>

<sup>&</sup>lt;sup>32</sup>We observe structure age at only a single point in time each year (in July). We therefore omit teardown sales in year t when the new structure appears in year t because we cannot determine if the structure was built before or after the sale. Thus, our sample surely omits some teardowns that should be included.

<sup>&</sup>lt;sup>33</sup>Not surprisingly, teardowns are more likely to have a sale in our estimation sample than are assemblies.

## 5 Results

This section presents results from our test for the existence of surplus. It also contains robustness checks which aim to demonstrate that the estimated surplus reflects neither the endogeneity of assembly location nor high reservation prices.

## 5.1 Existence of Surplus

We motivate our comparison of assembly to teardown parcels by presenting their mean characteristics in columns (1) and (2) of Table 1. As a reminder, the sample contains only sales of assembly and teardown parcels that occur shortly before assembly or teardown. In order to provide context, column (5) displays means for non-assembly and non-teardown parcels. Overall, assemblies (column (1)) and teardowns (column (2)) are located in similar types of neighborhoods. Although there are statistically significant differences between the groups – column (3) – the magnitude of the differences are economically small (e.g., a black neighborhood share of 5.3 percent for assemblies versus 6.0 percent for teardowns). Moreover, regression-adjusted mean differences that condition out census tract fixed effects, presented in column (4), generally show extremely small and imprecise differences.<sup>34</sup> This similarity is advantageous for our primary test given its requirement that assemblies and teardowns be comparable in the unobserved determinants of price.

A glaring exception to the comparability is the average parcel's use before the assembly or teardown occurs: teardown parcels are much more likely to be single family than are assembly parcels. We take great care below to address this difference. All specifications include a control for use type and we allow the coefficients on the year by quarter of sale

Assemblies often involve the purchase of parcels over a long period. In contrast, teardowns are a single, discrete event. Thus, restricting to sales "close" in time to assembly or teardown filters out relatively more assemblies than teardowns.

<sup>&</sup>lt;sup>34</sup>Demographic characteristics, such as the poverty rate, are from the census and are measured at the block group level.

controls to differ by residential and non-residential status. We also present specifications which restrict the sample to only single family parcels or only non-residential parcels to make teardowns and assemblies as comparable as possible.

Table 2 presents the results for our test of surplus – absent frictions, the price of teardowns and assembly parcels should equate – implemented by estimating equation (2) on the sample of teardown and assembly sales. The first column of this table shows the results from a regression without any controls. This raw comparison shows that, on average, parcels sell into assembly for an insignificant one percent lower price than parcels sold for teardown.

In Column (2), we include tract fixed effects. The coefficient estimate indicates that being sold into an assembly is associated with a roughly 40 percent price premium relative to being sold for redevelopment without a change in parcel boundary. The extremely large magnitude of the assembly surplus suggests substantial frictions in the market for assembly.<sup>35</sup> However, this large estimate is not inconsistent with a simple, plausible example. Suppose that unassembled, two parcels can house two single family homes (one each) and that assembled, the parcels can house a six story condo building; such a situation would not be uncommon in parts of Los Angeles. Further, assume that the houses sell for \$500,000 each and condos sell for \$300,000 each. Assume that each house's capital value is \$300,000 and each condo's capital value is \$200,000. In this case, the premium to assembly is 50 percent:  $(6^{*}(\$300,000 - \$200,000) - 2^{*}(\$500,000-\$300,000)) / 2^{*}(\$500,000-\$300,000)).$ 

Columns (3) through (6) present alternative specifications. Column (3) addresses within neighborhood variation by adding neighborhood demographic controls at the block group level and controls for local amenities. Column (4) further addresses intra-neighborhood variation by replacing the tract fixed effects with very finely grained block group fixed effects. Column (5) controls for the evolution in price specific to each tract by including tract-year

<sup>&</sup>lt;sup>35</sup>We use the test developed in Oster (2013) to assess the sensitivity of our estimate to unobservables based on coefficient movement and  $R^2$ s across columns (1) and (2). Compared to the estimation with no covariates in column (1), we find a relatively tight bound on our estimate in column (2) of (0.42, 0.64).

terms. This is an extremely saturated and demanding specification: identification of the assembly premium comes solely from comparing assemblies to teardowns within census tracts in the same year.<sup>36</sup> Regardless of specification, the magnitude of the surplus estimate varies little, ranging from around 35 to 40 percent.

It is possible that assemblies occur with the aim of future, rather than immediate, redevelopment. If so, the sales price of the to-be-assembled parcels may reflect the expected return to any existing capital over the period before redevelopment (McMillen and O'Sullivan, 2011). Such a scenario would bias our surplus estimates upward. To address this possibility, in column (6) we restrict our sample to only assemblies where the existing capital is immediately torn down. (Due to the structure of our data, we under-identify these parcels. See Section 3 of the Data Appendix for details.) The magnitude of the surplus estimate is virtually the same as the unrestricted sample.

To address the concern that assembled parcels are systematically less likely to be single family than are teardowns, columns (7) and (8) restrict the sample to single family parcels. The restriction makes the teardown and assembly samples comparable in terms of preassembly use. The surplus estimates remain similar to those produced by the full sample.<sup>37</sup>

These surplus results are credible only insofar as we control for the endogenous location of assembly, and show that the results are not driven by owners' high subjective valuation of their homes. We deal with each of these issues in turn, starting with a falsification exercise and a repeat sales analysis that address the issue of endogenous location.

Intuitively, our falsification check asks whether land very near an assembly is more valuable than land very near a teardown. Such a difference would suggest that teardown areas

<sup>&</sup>lt;sup>36</sup>With the inclusion of the tract-year fixed effects in Column (5), the identification of the assembly premium comes solely from tract-years with both an assembly and a teardown. There are 172 such tract-years. Nonetheless, the estimate of the assembly premium with the inclusion of tract-year fixed-effects is reasonably robust to alternative specifications (available from the authors upon request).

<sup>&</sup>lt;sup>37</sup>Our results also depend on observing both teardown and assembly parcels. The relative share of these parcels is stable after 2003; in 2002 and before, we observe low levels of teardowns. Results in Table 2 are robust to omitting teardowns and assemblies before 2003.

differ systemically from assembly areas and that teardowns may not be an appropriate control group for assemblies. Stated differently, we estimate whether there is surplus to assembly for the non-assembly neighbors of assemblies. A positive estimate of assembly surplus for non-assembled parcels would suggest the estimates on Table 2 are spurious. We focus on the years just before the assembly to avoid any direct influence of the assembly on the value of neighboring parcels. Such spillovers would contaminate the falsification check. (We explore the possibility of spillovers to assembly in Section 5.3.)

Specifically, we estimate

$$\log(\frac{\text{real sale price}}{\text{lot square footage}})_{i,g,t} = \lambda_0 + \lambda_1 1\{\text{pre-assembly very close neighbor}\}_{i,t} + \text{year-quarter}_t + \text{neighborhood}_g + \lambda_2 X_{i,t} + \lambda_3 K_{i,t} + \epsilon_{i,g,t},$$
(3)

where  $K_{i,t}$  is a vector measuring the value of capital on parcel i at time t and

 $1\{pre\text{-assembly very close neighbor}\}_i$  is an indicator variable equaling one if parcel *i* is a very close neighbor of an assembled parcel and *t* is one to three years prior to the start of the assembly. The sample for this estimation is the very close neighbors of teardowns and assemblies three years before assembly or teardown.

We define very close neighbors as parcels located on the same "map book page" as assembled parcels. Map book pages are geographical units defined by the County Assessor's office. They are quite small, containing an average of only 26 parcels, and typically encompass a single square block. Figure 3 displays a representative example.

The coefficient  $\lambda_1$  therefore captures the "surplus" to being a very close assembly neighbor relative to being a very close teardown neighbor. If the estimated premium to assembly reflects bias from very localized unobserved amenities—those whose effect on property values is more geographically limited than a block group—these amenities should be reflected in the price of the assembled parcels' very close neighbors. Thus, a positive and precise  $\lambda_1$  indicates that the Table 2 results may be biased. A zero estimate of  $\lambda_1$  suggests no bias, and a negative  $\lambda_1$  suggests potential negative selection in assembly relative to teardown.<sup>38</sup>

Table 3 presents the results. Column (1) suggests that there is virtually no difference in price between very close assembly neighbors pre-assembly and very close teardown neighbors pre-teardown. The 95 percent confidence interval is bounded by -4.2 percent and 3.2 percent. Unlike the prices for teardowns and assemblies, the market value of neighboring parcels likely includes the value of capital. Column (2) therefore controls, in a very flexible manner, for the parcel's capital.<sup>39</sup> The discount falls a bit and remains imprecise. Using block group fixed-effects produces an even smaller point estimate—column (3). Including tract-year fixed effect in column (4) produces a small and significant negative effect. Restricting the sample to single family parcels produces an extremely small, imprecise estimate—column (5).

Overall, these results suggest that there is little bias involved in comparing assemblies to teardowns and they strongly suggest that there is no upward bias in the comparison: The upper 95th percent confidence interval for the estimate in column (3) is a mere 0.016.

In order to further address the endogenous location of assembly, we now turn to a repeat

<sup>&</sup>lt;sup>38</sup>This falsification test is cleanest when the decision to assemble or tear down certain parcels within a mapbook page is determined by idiosyncratic factors unrelated to unobservable determinants of parcel value. For example, imagine two very similar neighboring houses. The owner of one house has a child in 10th grade who wishes to graduate at his current school and the other neighbor has no children. These owners may have very different subjective costs of moving. Unlike this example, it is possible that parcels, even within a map book page, may be selected into assembly or teardown. For example, positive selection into assembly could occur if developers avoided parcels with unobservable negative characteristics. Negative selection into assembly could occur if houses on inferior parcels were less well maintained and were therefore cheaper to scrap in a redevelopment. Any such positive or negative selection into either assembly or teardown would be problematic for the falsification check. However, if the selection is the same for both assemblies and teardowns, the selection would difference out.

<sup>&</sup>lt;sup>39</sup>We control for the presence of capital in a relatively non-parametric manner by including the full set of interaction terms between indicator variables for the decile of the vintage of the existing capital (year built) and indicator variables for the decile of the quantity of the capital (square feet of structure per square foot of land). This approach fails to account for both differences in depreciation across properties and differences in the initial quality of the capital. We therefore also control for the capital to land ratio (assessed improvements/assessed land value), and the value of capital per square foot of land (assessed improvements/lot size). We allow all of the coefficients on the capital variables to vary by residential and non-residential pre-assembly use.

sales approach, which we estimate using the following specification

$$\log(\frac{\text{real sale price}}{\text{lot square footage}})_{i,t} = \beta_0 + \beta_1 1\{\text{redevelop}\} * 1\{\text{ever assembly}\}_{i,t} + \beta_2 1\{\text{redevelop}\} + \beta_4 K_{i,t} + \beta_5 K_{i,t} * 1\{\text{redevelop}\} + \text{parcel}_i + \text{year-quarter}_t + \varepsilon_{i,t}.$$

$$(4)$$

We restrict the sample to teardown and assembly parcels which have, in addition to the teardown or assembly sale, a previously occurring sale (after 1985). The indicator 1{redevelop} equals one when the observation is the sale immediately before either an assembly or teardown. The indicator 1{ever assembly} is one if the parcel is ever sold into an assembly in our sample; it is not included as an uninteracted, main effect because it is collinear with the parcel fixed effect. Thus, 1{redevelop} \*1{ever assembly}<sub>*i*,*t*</sub> equals one if this sale is the final sale before assembly, and  $\beta_1$  is the coefficient of interest.

This repeat sales estimation can be viewed as a double difference: relative to the same parcel over time, and relative to teardowns, is there a premium for being the last sale before an assembly? The principle advantage of the approach is that the parcel fixed effect,  $parcel_i$ , controls for any time-invariant feature of the parcel such as greenery, freeway access, proximity to commercial strips, etc. However, if assemblies are anticipated, sales before the final assembly sale may partially reflect the value of the parcel as used in an assembly.<sup>40</sup> In this case, the repeat sales estimates is downward biased.

While the "redevelopment" sales values should not reflect the value of capital, the earlier sales values likely do (i.e., the earlier sale buyers likely intended to use the capital located on the parcels). We therefore include controls in equation (4) for capital,  $K_{i,t}$ , and capital on the final sale  $(K_{i,t} * 1\{\text{redevelop}\}).^{41}$ 

<sup>&</sup>lt;sup>40</sup>Specifically, the sales value of the parcel should equal the flow of discounted rents over the years the parcel is expected to be used "as is" plus the discounted flow of rents from the years the parcel is expected to be in use as part of an assembly.

<sup>&</sup>lt;sup>41</sup>Given the much smaller sample size, we use a more parsimonious set of capital controls than for Table

Table 4 presents the repeat sales results. From the smaller numbers in the observation rows, it is clear that the data requirements of the approach yield a substantially thinner sample than that used for the initial test in Table 2. This is particularly true for the number of unique assemblies and teardowns. The repeat sales approach thus places significant demands upon the data. Column (1) presents a specification which lacks the capital controls and column (2) adds in these controls. Both specifications yield an assembly surplus of about 15 percent. Unlike the initial test, these estimates are sensitive to controlling for the geographicspecific evolution in prices. Controlling for city-year effects in column (3) boosts the surplus estimate to 21 percent.<sup>42</sup> While smaller than the surplus estimate produced by our principal approach, 21 percent remains a sizable surplus to assembly. Moreover, as we note above, these estimates may be biased downward if assembly is anticipated when the parcel sells before the final assembly sale.<sup>43</sup>

We conclude our analysis of surplus by considering whether our estimates of surplus are driven by high subjective valuations. Subjective valuations should be particularly high for single family homes (Ellickson, 1973). Therefore, in order to assess the importance of reservation prices in determining surplus, we examine the difference in surplus between single family and non-single family parcels. Column (1) of Table 5 presents results from our surplus estimation (Table 2) using a sample that excludes single family homes. The surplus estimate of 33 percent is extremely similar to the analogous estimate of 36 percent for single family homes (column (7) of Table 1). Similarly, column (2) further restricts the sample to nonresidential parcels which we expect would have lower subjective valuations, on average, than single family homes. The result is again quite similar to the estimate for single family homes.

<sup>3&#</sup>x27;s falsification check. See the note on Table 4 for a complete list of the capital control variables.

 $<sup>^{42}</sup>$ The tract-year terms used elsewhere in the paper cannot be employed here given the thinness of the estimation sample.

 $<sup>^{43}</sup>$ If we re-estimate the specifications in Table 2 with this sample, results are similar to those in Table 2. Therefore, differences between the results in this table and Table 2 are a function of the estimation strategy, not the sample.

Finally, single family owner-occupants would likely have higher subjective valuations than absentee owners. In column (3) we restrict the sample to single family homes and allow the assembly premium to vary by owner status. If reservation prices were the primary factor driving the surplus estimate, owner-occupied homes should have a higher surplus. The results suggest the opposite, but are quite imprecise. In sum, the patterns of surplus by parcel use and ownership status are not consistent with surplus consisting primarily of high reservation prices.

## 5.2 Causes of Assembly Surplus

Broadly, we construe the results from Tables 2 through 5 as jointly offering reasonably strong support for the existence of an economically significant premium to assembled land. Given this, we now turn to a consideration of the sources of this surplus. Surplus can be driven by both public and private market frictions. Public frictions that yield surplus are regulations on the use of land. Private market frictions include problems of holdouts and other bargaining issues. We use two tests to assess whether private market failures contribute to the surplus.<sup>44</sup> The results point toward an important role for private market failures in creating assembly surplus. We cannot rule out the possibility, though, that public market failures may also play an important role.

Starting with the first test: in a well functioning market, smaller parcels should be more likely to assemble than larger parcels due to the convexity of the land value function in areas undergoing land assembly. Alternatively, in a poorly functioning market, small parcels are particularly prone to market frictions such as holding out and strategic delay.

<sup>&</sup>lt;sup>44</sup>A natural third test would examine the land assembly premium as a function of the time of sale, where later selling parcels would garner a larger premium. Unfortunately, the structure of our data precludes such an approach. Market participants told us that sales into assembly are frequently "contingent sales" so that no sale of any potentially assembled property occurs until the developer has gathered a sufficient amount of land. In such a case, all properties that sell into the assembly officially transact on the same date, even though the actual bargaining date may differ substantially. We observe only the official transaction date, which is frequently the same for parcels selling into a given assembly.

We therefore examine the influence of parcel size on assembly probability. To do so, we use the cross-section of all 2 million county parcels from the first year of the sample (1999).<sup>45</sup> The estimating equation is

$$assembly_{i,g} = \gamma_0 + \gamma_1 \log(lot \ square \ footage_i) + neighborhood_g + \gamma_2 amenities_i + \gamma_3 K_{i,t} + \varepsilon_{i,g},$$
(5)

where  $assembly_{i,g}$  equals one if the parcel is involved in an assembly over the following 12 years of the sample and zero otherwise, and *lot square footage*<sub>i</sub> is the size of the parcel.  $\gamma_1$  captures the marginal effect of lot size on the probability of assembly and is the coefficient of interest.

The coefficient  $\gamma_1$  is properly identified when lot size is uncorrelated with unobserved determinants of assembly,  $\varepsilon_{i,g}$ . There are at least three possible reasons to be concerned that this may not be the case. First, the likelihood of assembly may vary by location, and location may be correlated with parcel size. For example, parcels near the urban center may be ripe for assembly and redevelopment due to traffic congestion. These parcels may be relatively small, but it is their lower commute time, not their size, driving the assembly decision. As before, we net out such neighborhood-specific factors with tract or block group fixed effects, so that all our comparisons are made within a given, limited geographic area.

The second threat to the credibility of the identifying assumption is the possibility that large plots tend to have more valuable capital per square foot of land than smaller parcels. All else equal, the more valuable the capital on a parcel, the less likely the capital will be scrapped to allow for redevelopment and the less likely is assembly. We therefore control for the presence of capital on each parcel in a very flexible manner (see the note to Table 6).

A third reason lot size may be correlated with the error term is because not all lots are

<sup>&</sup>lt;sup>45</sup>Summary statistics for this sample are in Appendix Table 1.

"assemble-able," and "assemble-ability" may be correlated with lot size. In general, larger parcels should be more likely to be unable to assemble because they have no contiguous neighbors. However, this possibility would bias our estimation toward small parcels being more likely to assemble, and hence make it less likely we will conclude that the market for assembly is inhibited by frictions.

Column (1) of Table 6 indicates that a 10 percent increase in the size of a parcel increases the probability of ever being assembled by 0.1 percent. This is a large effect, equal to around 10 percent of the sample mean probability of ever assembling (see the "Share Ever Assembled" row). The remaining columns use block group fixed effects—columns (2) and (4)—and include the capital controls—columns (3) and (4). These permutations have little effect on the lot size coefficient.<sup>46</sup>

We believe that these results are unlikely to be driven exclusively by a correlation between reservation prices and lot size, or by land use regulations. Conditional on the right-hand side controls, which include the pre-assembly use of the parcel, there is no obvious reason why reservation prices should vary by lot size. Similarly, land use regulations are not likely to create an incentive to use large parcels in an assembly.<sup>47</sup> Thus, we interpret this evidence as suggesting that private market frictions are prevalent in the market for land assembly.

Our second test for the causes of assembly surplus relies on a similar logic to that of the previous test. In the absence of frictions, the convexity of the land value function requires that small parcels should not receive a premium relative to large parcels when they sell into an assembly. Alternatively, when market frictions are present, small parcels may command

<sup>&</sup>lt;sup>46</sup>If we exclude the least dense quartile of parcels in terms of the population density of their block group from the sample, ensuring that our results are driven by outcomes in a dense urban setting, results are little changed.

<sup>&</sup>lt;sup>47</sup>In fact, land use regulation may provide an incentive to assemble small parcels. Consider minimum setback requirements. Under a fixed setback requirement, assembling parcels relaxes the regulatory constraint as the adjoining borders of the input parcels become interior to the assembled parcel and do not require a setback (Shoup, 2008). Hence, a higher percentage of the assembled land can have capital situated on it relative to the unassembled land. The gain along this dimension is greater for joining together several small parcels than it is for joining together a few large parcels.

a premium. We implement this test by estimating

$$\log(\frac{\text{real sale price}}{\text{lot square footage}})_{i,g,t} = \theta_0 + \theta_1 \log(\text{lot square footage}_i) + \text{year-quarter}_t + \text{assembly group}_g + \theta_2 \text{amenities}_i + \varepsilon_{i,g,t}, \tag{6}$$

where assembly  $group_g$  is a fixed effect for each assembly group — a set of contiguous parcels that are assembled together. The coefficient  $\theta_1$  captures the marginal effect of lot size on sales price and is the coefficient of interest. The sample is limited to assembly sales.

The assembly  $group_g$  term ensures the influence of lot size on sales price is measured solely within groups of parcels assembled together. The identifying assumption required for equation (6) — that lot size is uncorrelated with the unobserved determinants of price is therefore extremely plausible. To-be-assembled parcels are sold only for their land value. Within an assembly group parcels are contiguous and thus likely have very similar location value.

Table 7 presents the results. Column (1) shows that a 10 percent increase in parcel size reduces the sales price of a to-be-assembled parcel by roughly 8 percent. Column (2) includes the distance to amenities controls, which vary somewhat within a parcel group, and produces nearly identical results. There is little reason to think that reservation prices vary with lot size within a parcel group, conditional on use. Thus, the substantial premium to small parcels supports the hypothesis that the owners of these parcels tend to hold out and demand higher than average prices for their land – behavior which works to reduce the number of successful assemblies.<sup>48</sup> While it is possible that land use regulation differs within assembly groups in a manner correlated with lot size, we view this evidence as most

<sup>&</sup>lt;sup>48</sup>A similar logic yields an expectation that parcels geographically central to the assembly should receive a larger premium. Because we do not observe parcel maps until 2006, we are unable to implement such a test exactly. However, we can proxy for parcel centrality with the geographic center of the parcel (measured with substantial noise in the early years of our sample), and ask whether parcels "central" to the assembly receive a larger premium. We do not find any such evidence. Given the crudeness of our measure, we believe the role of centrality remains an open question.

consistent with private market failures.

## 5.3 Welfare Implications

Given that our evidence suggests a substantial deadweight loss associated with the inhibition of land assembly, it is natural to ask whether this inhibition causes an overall welfare loss. We explore this issue at length in an appendix and summarize our conclusion – that a welfare loss is likely but not certain – here. If an increase in land assembly decreases welfare, it must be that the negative external effects of assembly outweigh the reduction in direct deadweight loss (the shaded area in Figure 1b). We divide the external effects of land assembly into local and non-local (that is, metropolitan area wide) effects. To estimate the local external effects of land assembly, we estimate a modified version of equation (3), where we replace  $1{pre-assembly very close neighbor}_i$  with an indicator for being on the same map book page as an assembly after the assembly, and we restrict the sample to non-assembly and non-teardown sales. Intuitively, this specification estimates whether parcels very close to the assembly experience a price change after the assembly. In short, we observe no capitalization of assembly into nearby properties, suggesting no local externalities to land assembly.

Therefore, for additional land assembly to decrease welfare, metropolitan area externalities to assembly must be quite negative. We believe this is unlikely, but cannot rule out this possibility. Furthermore, assessing the welfare consequences of frictions is inherently difficult in an urban setting. Most urban models assume market failures that generate agglomeration and congestion. In such a second-best setting, removing a negative externality, such as land assembly frictions, need not improve overall welfare (e.g. Helpman, 2007; Kanemoto, 2013). Overall, though, we believe the combination of direct evidence of deadweight loss and the absence of evidence for negative local spillovers points toward the possibility of a large welfare loss from land assembly frictions.
# 6 Conclusion

The results of this paper provide robust confirmation of the hypothesis that the market for land assembly is inhibited by frictions. We find that land selling into assembly garners a 15 to 40 percent premium. Because we find that smaller parcels are less likely to be assembled and fetch higher per square foot prices within a given assembly, we contend that at least part of the overall frictions we document are due to private market failures. Though the evidence is not conclusive, it is possible that the frictions inhibiting land assembly are associated with large and negative welfare effects.

While the problems we document are salient for the developed world, they are likely to be just as salient, if not more, for the developing world, where urban growth often confronts problems of assembly. Seshadri (2011) holds that problems in land assembly inhibit the growth of special economic zones in India, and media coverage of holdouts in Chinese land assembly is abundant (for examples, see French (2007) and Molly (2012)).

If the market for private land assembly operates as poorly as suggested by this work, it is natural to look to government for remediation. The most recent large-scale government action to assemble land is known as "urban renewal," a process in the United States, Canada and in some parts of Europe in the 1960s and 70s. Urban renewal used the government's power of eminent domain to assemble small parcels of land, predominantly in the urban core. Urban renewal, however, ended amid charges of developer cronyism and racism and has largely been judged harshly by historians (Cord, 1974). While our findings suggest that the land assembly portion of urban renewal would likely not have been achieved by the private market, our results from the private market have little to offer in understanding the potential value of government-driven assembly.

Alternatively, there may be a middle ground between the private exchange of land and direct government intervention. Under the practice of "land readjustment," redevelopment districts are formed and land assembly occurs if a majority of land owners in the district vote for it. The owners are given a stake in the new development as compensation. The practice has been successfully used in a number of countries, including Germany, South Korea, and Israel (Hong, 2007). Land readjustment may provide a mechanism for overcoming the private market failures inherent in land assembly, while mitigating the costs associated with direct government intervention. Another possible middle ground approach is graduated density zoning, which awards the right to build at greater density to larger lots. The "density bonus" increases the value of large lots and encourages assembly. Although it does not eliminate holdouts, it may generate a fear of being left out and therefore reduce the prevalence of holdouts (Shoup, 2008).

More broadly, our evidence argues that "tragedies of the anti-commons," in which problems of fragmented ownership impede socially optimal outcomes, are of economically significant magnitude. While we illustrate these problems in the context of land assembly, they apply in many contexts with fragmented ownership, such as spectrum rights or intellectual property.

# References

- Angrist, Joshua and Pischke, Jorn-Steffen, 2009. Mostly Harmless Econometrics: An Empiricist's Companion. Princeton: Princeton University Press.
- Asami, Y., 1988. "A game-theoretic approach to the division of profits from economic land development." Regional Science and Urban Economics 18: 233–246.
- Bogart, Dan and Richardson, Gary, 2009. "Making property productive: reorganizing rights to real and equitable estates in Britain, 1660-1830." European Review of Economic History 13: 3–30.
- Brownstone, David and DeVany, Arthur, 1991. "Zoning, Returns to Scale, and the Value of Undeveloped Land." *Review of Economics and Statistics* 73(4): 699–704.
- Brueckner, Jan K., 1980a. "Residential Succession and Land-Use Dynamics in a Vintage Model of Urban Housing." *Regional Science and Urban Economics* 10: 225–240.
- Brueckner, Jan K., 1980b. "A Vintage Model of Urban Growth." *Journal of Urban Economics* 8: 389–402.
- Cadigan, John, Schmitt, Pamela, Shupp, Robert, and Swope, Kurtis, 2011. "The Holdout Problem and Urban Sprawl: Experimental Evidence." *Journal of Urban Economics* 69(1): 72–81.
- Cameron, A. Colin and Travedi, Pravin K., 2005. Microeconometrics: Methods and Applications. Cambridge: Cambridge University Press.
- Colwell, Peter and Sirmans, C. F., 1993. "A Comment on Zoning, Returns to Scale and the Value of Undeveloped Land." *Review of Economics and Statistics* 75(4): 783–786.
- Colwell, Peter F. and Munneke, Henry J., 1997. "The structure of urban land prices." *Journal of Urban Economics* 14: 321–336.
- Colwell, Peter F. and Munneke, Henry J., 1999. "Land Prices and Land Assembly in the CBD." Journal of Real Estate Finance and Economics 18(2): 163–180.
- Cord, Steven, 1974. "Urban Renewal: Boon or Boondoggle?" American Journal of Economics and Sociology 33(2): 184–186.
- Coulson, Edward, 1989. "The Empirical Content of the Linearity-as-Repackaging Hypothesis." Journal of Urban Economics 25: 295–309.
- Cunningham, Christopher, 2010. "Estimating the holdout problem in land assembly." Federal Reserve Bank of Atlanta Working Paper.
- Duranton, Gilles and Puga, Diego, 2004. "Micro-Foundations of Urban Agglomeration Economies." In J. V Henderson and J.E Thisse, (Eds.) "Handbook of Regional and Urban Economics," Elsevier, B.V.
- Dye, Richard F. and McMillen, Daniel P., 2007. "Teardowns and land values in the Chicago metropolitan area." *Journal of Urban Economics* 61: 45–63.

- Eckart, Wolfgang, 1985. "On The land assembly problem." *Journal of Urban Economics* 18: 364–378.
- Ellickson, Robert C., 1973. "Alternatives to Zoning: Covenants, Nuisance Rules, and Fines as Land Use Controls." The University of Chicago Law Review 40(4): 681–781.
- Fennell, Lee Anne, 2004. "Common Interest Tragedies." Northwestern University Law Review 98(3): 907–990.
- Field, Alexander James, 1992. "Uncontrolled Land Development and the duration of the Depression in the United States." *Journal of Economic History* 52(4): 785–805.
- French, Howard W., 2007. "In China, Fight Over Development Creates a Star." New York Times.
- Fu, Daniel P. McMillen, Yuming and Somerville, Tsur, 2002. "Land Assembly: Measuring Holdup." University of British Columbia Centre for Urban Economics Working Paper 02-08.
- Glaeser, Edward, 2011. The Triumph of the City. New York: The Penguin Press.
- Glaeser, Edward and Gyourko, Joseph, 2002. "The Impact of Zoning on Housing Affordability." NBER Working Paper No. 8835.
- Glaeser, Edward, Gyourko, Joseph, and Saks, Raven, 2005. "Why is Manhattan so Expensive? Regulation and the Rise in Housing Prices." Journal of Law and Economics XLVIII: 331–369.
- Grossman, Sanford J. and Hart, Oliver, 1980. "Takeover Bids, The Free-Rider Problem, and the Theory of the Corporation." The Bell Journal of Economics 11(1): 42–64.
- Grossman, Zachary, Pincus, Jonathan, and Shapiro, Perry, 2010. "A Second-Best Mechanism for Land Assembly." Working Paper.
- Gyourko, Joseph, Saiz, Albert, and Summers, Anita, 2008. "A New Measure of the Local Regulatory Environment for Housing Markets: The Wharton Residential Land Use Regulatory Index." *Urban Studies* 45(3): 693–729.
- Harding, John P., Rosenthal, Stuart, and Sirmans, C. F., 2003. "Estimating Bargaining Power in the Market for Existing Homes." *Review of Economics and Statistics* 85(1): 178–188.
- Heller, Michael A., 1998. "The Tragedy of the Anticommons: Property in the Transition from Marx to Markets." *Harvard Law Review* 111(3): 621–688.
- Heller, Michael A. and Eisenberg, Rebecca S., 1998. "Can Patents Deter Innovation? The Anticommons in Biomedical Research." Science 280: 698–701.
- Helpman, Elhanan, 2007. "The Size of Regions." In David Pines, Efraim Sadka, and Itzhak Zilcha, (Eds.) "Topics in Public Economics: Theoretical and Applied Analysis," Cambridge University Press.
- Henderson, J. Vernon, 1977. Economic Theory and the Cities. New York: Academic Press.

- Hong, Yu-Hung, 2007. "Assembling Land for Urban Redevelopment: Issues and Opportunities." In Yu-Hung Hong and Barrie Needham, (Eds.) "Analyzing Land Readjustment: Economics, Law and Collective Action," Cambridge, MA: Lincoln Institute of Land Policy.
- Hornbeck, Richard and Keniston, Daniel, 2014. "Creative Destruction: Barriers to Urban Growth and the Great Boston Fire of 1872." Havard University Working Paper.
- Hsieh, Chang-Tai and Moretti, Enrico, 2015. "Why Do Cities Matter? Local Growth and Aggregate Growth." NBER Working Paper No. 21154.
- Kanemoto, Yoshitsugu, 2013. "Second-best costbenefit analysis in monopolistic competition models of urban agglomeration." *Journal of Urban Economics* 76: 83–92.
- Kominers, Scott Duke and Weyl, E. Glen, 2010. "Concordance Among Holdouts." Working Paper April.
- Landis, John D. and Hood, Heather, 2005. "Ther Future of Infill Housing in California: Opportunities, Potential, Feasibility and Demand." Tech. rep., Institute of Urban and Regional Development, University of California, Berkeley.
- Libecap, Gary D. and Lueck, Dean, 2011. "The Demarcation of Land and the Role of Coordinating Property Institutions." *Journal of Public Economics* forthcoming.
- Libecap, Gary D., Lueck, Dean, and O'Grady, Trevor, 2010. "Large Scale Institutional Changes: Land Demarcation Within the British Empire." NBER Working Paper No. 15820.
- McDonald, John F., 2007. "What is Public Use? Eminent Domain and the Kelo Decision." Cornell Real Estate Review 5(1): 1–17.
- McMillen, Dan and O'Sullivan, Arthur, 2011. "Option Value and the Price of Teardown Properties." Mimeo.
- Menezes, Flavio and Pitchford, Rohan, 2004. "The land assembly problem revisited." *Regional Science and Urban Economics* 34: 155–162.
- Merrill, Thomas W., 1986. "The Economics of Public Use." Cornell Law Review 72(61).
- Miceli, Thomas and Segerson, Kathleen, 2007. "A bargaining model of holdouts and takings." American Law and Economics Review 9(1): 160–174.
- Miceli, Thomas J., Sergerson, Kathleen, and Sirmans, C. F., 2008. "Tax Motivated Takings." National Tax Journal LXI(4): 579–591.
- Miceli, Thomas J and Sirmans, C. F., 2007. "The holdout problem, urban sprawl, and eminent Domain." Journal of Housing Economics 16: 209–219.
- Molly, Mark, 2012. "Motorway construction thwarted by stubborn elderly couple's house." Metro UK.

- Munch, Patricia, 1976. "An Economic Analysis of Eminent Domain." *Journal of Political Economy* 84(3): 473–497.
- Nelson, Arthur and Lang, Robert, 2007. "The Next 100 Million." Planning 73(1): 4–6.
- O'Flaherty, Brendan, 1994. "Land Assembly and Urban Renewal." Regional Science and Urban Economics 24: 284–300.
- Oster, Emily, 2013. "Unobservable Selection and Coefficient Stability: Theory and Validation." NBER Working Paper No. 19054.
- Pitchford, Rohan and Wright, Mark L. J., 2008. "Holdouts in Soverign Debt Restructuring: A Theory of Negotiation in a Weak Contractual Environment." Working Paper.
- Plassmann, Florenz and Tideman, T. Nicolaus, 2008. "Efficient Urban Renewal Without Takings: Two Solutions to the Land Assembly Problem." Working Paper.
- Rosen, Christine Meisner, 1986. The Limits of Power: Great Fires and the Process of City Growth in America. Cambridge: Cambridge University Press.
- Rosenthal, Jean-Laurent, 1990. "The Development of Irrigation in Provence, 1700-1860: The French Revolution and Economic Growth." *Journal of Economic History* 50(3): 615–638.
- Rosenthal, S.S. and Strange, W.C., 2004. "Evidence on the Nature and Sources of Agglomeration Economies." In J. V Henderson and J.E Thisse, (Eds.) "Handbook of Regional and Urban Economics," Elsevier, B.V.
- Rosenthal, Stuart, 2008. "Old Homes, Externalities, and Poor Neighborhoods. A Model of Urban Decline and Renewal." *Journal of Urban Economics* 63(3): 816–840.
- Rosenthal, Stuart and Helsley, Robert W., 1994. "Redevelopment and the Urban Land Price Gradient." *Journal of Urban Economics* 35: 182–200.
- Seshadri, Triyakshana, 2011. "An Analysis of the Feasibility of Private Land Assembly for Special Economic Zones in India." Urban Studies 49(10): 2285–2300.
- Shoup, Donald, 2008. "Graduated Density Zoning." Journal of Planning Education and Research 28(2): 161–179.
- Strange, William C., 1995. "Information, Holdouts and Land Assembly." Journal of Urban Economics 38: 317–332.
- Tanaka, Tomomi, 2007. "Resource Allocation with Spatial Externalities: Experiments on Land Consolidation." The B.E. Journal of Economic Analysis and Policy 7(1): article no. 7.
- Wheaton, William, 1982. "Urban Spatial Development with Durable but Replaceable Capital." Journal of Urban Economics 12: 53–67.
- Wheaton, William, 1983. "Theories of Urban Growth and Metropolitan Spatial Development." In J.V. Henderson, (Ed.) "Research in Urban Economics, Vol. 3," Greenwich, Conn.: JAI Press.

Willis, Carol, 1995. Form Follows Finance. New York: Princeton Architectural Press.

#### Interviews

Krause, Wolfgang. Principal Planner, City of Glendale. By phone May 18, 2010.
Garcia, Jose. Planner, City of Maywood. By phone July 2, 1010.
Ashworth, Paul, Director of Planning and Development, City of Santa Fe Springs. In person, November 30, 2010.
Morrell, Wayne M., Principal Planner, City of Santa Fe Springs. In person, November 30, 2010.
Fuentes, Al, Redevelopment Consultant, City of Santa Fe Springs. In person, November 30, 2010.
Pastucha, Simon, Planner, Los Angeles Urban Design Studio. In person, December 1, 2010.
Olson, Todd, President, Community Development, The Olson Company. In person, December 2, 2010.





Note: Figure 1a illustrates convexity in land prices. Were land prices to increase linearly, they would follow the dashed line ending at 2V(a). When prices are convex, as shown with the thick curve, V(2a) > 2V(a). The thin curve shows how the convex curve flattens out as developers complete assemblies and the market moves toward equilibrium. In equilibrium, we expect the price of assembled and unassembled parcels to differ only by the "good institution" cost of assembly,  $\delta$ .





Sources: Permit data from Census Bureau, assemblies calculated from authors' dataset. See Data Appendix for details.

Notes: While the Census Bureau counts one permit for each new unit, one assembly (in our terms) may result in more than one unit. This means that our measure in the chart may understate the importance of assembly. However, the Census Bureau counts residential permits only, while our measure of assemblies includes assemblies that result in non-residential construction. In this sense, our measure overstates the importance of assembly. Because we do not have a cross-section of properties from 2003, we attribute all assemblies in 2002 and 2003 to 2002; for the purposes of this chart, we split the assemblies in 2002 evenly between 2002 and 2003. Though our assembly information continues through 2011, this chart ends in 2010 when the permit data end.

Figure 3: Map Book Pages are Small Areas



Source: Los Angeles County Assessor website.

Note: This picture shows a "map book page." The Assessor organizes all land in the County into map books. Each very small neighborhood is referenced by its page in such a map book.

|                       | Primary Estimation Sample: Assembly and Teardown Sales |          |            |  |                             |
|-----------------------|--|----------|------------|--|-----------------------------|
|                       | Assembly   | Teardown | Difference | Difference<br>conditional on<br>tract FE | & Non-<br>Teardown<br>Sales |
|                       | (1)  | (2)      | (3)        | (4)                                      | (5)                         |
| Ln(Price per square   | 3.897  | 4.104    | -0.207     | 0.643                                    | 4.069                       |
| foot)                 | [1.790]  | [1.273]  | (0.031)*** | (0.067)***                               | [1.251]                     |
| Single family         | 0.204  | 0.736    | -0.532     | -0.335                                   | 0.634                       |
|                       | [0.403]  | [0.441]  | (0.009)*** | (0.019)                                  | [0.482]                     |
| Poverty rate          | 0.106  | 0.114    | -0.008     | 0.001                                    | 0.133                       |
|                       | [0.111]  | [0.118]  | (0.003)**  | (0.002)                                  | [0.114]                     |
| Neighborhood share    | 0.052  | 0.060    | -0.008     | 0.004                                    | 0.084                       |
| Black                 | [0.094]  | [0.121]  | (0.003)**  | (0.002)                                  | [0.148]                     |
| Neighborhood Share    | 0.232  | 0.242    | -0.010     | -0.003                                   | 0.338                       |
| Hispanic              | [0.253]  | [0.266]  | (0.006)+   | (0.003)                                  | [0.272]                     |
| Neighborhood Share of | 0.039  | 0.046    | -0.007     | 0.003                                    | 0.042                       |
| Housing Units Vacant  | [0.052]  | [0.046]  | (0.001)*** | (0.002)+                                 | [0.042]                     |
| Median Household      | 63001  | 72431    | -9430      | 230                                      | 55445                       |
| Income                | [27623]  | [43627]  | ( 881)***  | ( 605)                                   | [28373]                     |
| Observations          | 2,797  | 7,979    | n/a        | n/a                                      | 1,819,021                   |

#### Table 1: Covariates of Assembly and Teardown Properties are Similar

Sources: See Data Appendix for complete information.

Notes: Columns (1), (2) and (5) display means. Column (3) displays the difference between columns (1) and (2) and the standard error for an equality of means test. Column (4) displays coefficient estimates from a regression of the variable in the row on an indicator variable for assembly and a set of census tract fixed-effects. [] denotes standard deviation and () denotes standard error. The standard errors in column (4) are clustered by city. Significance levels are denoted by \*\*\* for significant at the 0.1% level, \*\* for significant at the 1% level, \* for significant at the 5% level, and + for significant at the 10% level. Price per square foot in column (5) is not directly comparable to the other columns because it includes the value of capital and is therefore not comparable to the prices displayed in columns (1) and (2), which capture only the value of the land.

|  | Full Sample |         |         |         | Single Family |        |         |         |
|--|-------------|---------|---------|---------|---------------|--------|---------|---------|
|  | (1)         | (2)     | (3)     | (4)     | (5)           | (6)    | (7)     | (8)     |
| 1 if Parcel is in an assembly          | -0.01       | 0.42*** | 0.35*** | 0.39*** | 0.34***       | 0.37** | 0.39*** | 0.36*** |
|  | (0.18)      | (0.06)  | (0.05)  | (0.06)  | (0.10)        | (0.13) | (0.05)  | (0.04)  |
|  |             |         |         |         |               |        |         |         |
| Observations                           | 10,776      | 10,776  | 10,776  | 10,776  | 10,776        | 8,610  | 6,442   | 6,442   |
| Assemblies                             | 2,797       | 2,797   | 2,797   | 2,797   | 2,797         | 631    | 570     | 570     |
| Teardowns                              | 7,979       | 7,979   | 7,979   | 7,979   | 7,979         | 7,979  | 5,872   | 5,872   |
| Geographic Fixed Effects               |             |         |         |         |               |        |         |         |
| Tract                                  |             | Х       | Х       |         |               | Х      | Х       | Х       |
| Block Group                            |             |         |         | Х       |               |        |         |         |
| Tract * Year                           |             |         |         |         | Х             |        |         |         |
| Additional Covariates                  |             |         |         |         |               |        |         |         |
| Year-Quarter of Sale                   | Х           | Х       | Х       | Х       | Х             | Х      | Х       | Х       |
| Year-Quarter of Sale * Non-Residential | Х           | Х       | Х       | Х       | Х             | Х      |         |         |
| Use Classifications                    | Х           | Х       | Х       | Х       | Х             | Х      |         |         |
| Cubic in Lot Size                      | Х           | Х       | Х       | Х       | Х             | Х      | Х       | Х       |
| Neighborhood Demographics              |             |         | Х       |         |               | Х      |         | Х       |
| Distance to Key Amenities              |             |         | Х       |         |               | Х      |         | Х       |
| Assembly Teardowns                     |             |         |         |         |               | Х      |         |         |

#### Table 2: Excess Value of Land in Assembly

Sources: See Data Appendix for complete information.

Notes: The dependent variable is log(real sales price per square foot). We use the largest consistent sample of sales in the three years before teardown or assembly, where the teardown or assembly occurs between 1999 and 2011. Standard errors clustered at the city level are in parentheses. Significance levels are denoted by \*\*\* for significant at the 0.1% level, \*\* for significant at the 1% level, \* for significant at the 5% level, and + for significant at the 10% level. Neighborhood demographics include the following variables in both 2000 level form and as changes between 1990 and 2000 levels: poverty rate, neighborhood share black, neighborhood share Hispanic, share of housing units vacant and share of housing units owner-occupied. Use classifications include indicator variables for single family, non-condo multi-family, condo, vacant and other. Distance to key amenities include measures for the shortest distance from each parcel to three amenities: a highway entrance or exit, a metrolink stop (commuter rail), and a metrorail stop (subway or light rail).

|  | Full Sample |         |         | Single Family |         |
|--|-------------|---------|---------|---------------|---------|
|  | (1)         | (2)     | (3)     | (4)           | (5)     |
| 1*{Pre-Assembly Map Book Page}         | -0.005      | -0.01   | -0.004  | -0.033*       | -0.006  |
|  | (0.019)     | (0.013) | (0.010) | (0.016)       | (0.012) |
|  |             |         |         |               |         |
| Observations                           | 150,849     | 150,849 | 150,849 | 150,849       | 84,761  |
| Geographic Fixed Effects               |             |         |         |               |         |
| Tract                                  | x           | х       |         |               | х       |
| Block Group                            |             |         | х       |               |         |
| Tract-Year                             |             |         |         | х             |         |
| Additional Covariates                  |             |         |         |               |         |
| Year-Quarter of Sale                   | х           | х       | х       | х             | х       |
| Year-Quarter of Sale * Non-residential | x           | х       | х       | х             |         |
| Use Classification                     | x           | х       | х       | х             |         |
| Cubic in Lot Size                      | x           | х       | х       | х             | х       |
| Distance to Amenities                  | x           | х       |         |               | х       |
| Neighborhood Demographics              | x           | х       |         |               | х       |
| Capital Controls                       |             | х       | х       | х             | х       |
| Capital Controls * Non-residential     |             | x       | x       | x             | х       |

#### Table 3: Falsification Check - Do Future Assemblies Move the Current Price of Immediate Neighbors?

Sources: See Data Appendix for complete information.

Notes: The dependent variable in these regressions is log(real sales price per square foot). 1\*{Pre-Assembly Very Close Neighbor} equals one for parcels that share a mapbook page with a parcel that will assemble within three years. We use the largest consistent sample of sales in the three years before teardown or assembly, on mapbook pages where the teardown or assembly occurs between 1999 and 2011; however, we exclude all teardown and assembly sales. Standard errors clustered at the city level are given in parentheses. Significance levels are denoted by \*\*\* for significant at the 0.1% level, \*\* for significant at the 1% level, \* for significant at the 5% level, and + for significant at the 10% level. Use classifications, amenities and demographics are as given in the note on Table 2. The capital controls include the value of improvements to the land per square feet, the ratio of the value of improvements to the land (i.e., capital to land ratio) and a full set of interactions of indicators for decile of structure age with indicators for decile of structure square feet divided by lot square feet (a measure of the quantity of capital). For both deciles an 11<sup>th</sup> indicator variable is added to denote missing values.

|  | (1)    | (2)    | (3)    |
|--|--------|--------|--------|
| 1{Redevelopment}*1{Assembly}           | 0.13** | 0.15*  | 0.21** |
|  | (0.05) | (0.06) | (0.07) |
| Observations                           | 6,837  | 6,837  | 6,837  |
| No. Unique Assembled Parcels           | 853    | 853    | 853    |
| No. Unique Teardown Parcels            | 2,133  | 2,133  | 2,133  |
| Redevelopment and Capital Covariates   |        |        |        |
| 1{Redevelopment}                       | Х      | Х      | Х      |
| Capital Controls                       |        | Х      | Х      |
| 1{Redevelopment}*Capital Controls      |        | Х      | Х      |
| Geographic Fixed Effects               |        |        |        |
| Parcel Fixed Effect                    | Х      | Х      | Х      |
| Additional Covariates                  |        |        |        |
| Year-Quarter of Sale                   | Х      | Х      | Х      |
| Year-Quarter of Sale * Non-Residential | Х      | Х      | Х      |
| Use Classifications                    | Х      | Х      | Х      |
| City * Year Fixed Effects              |        |        | Х      |

#### **Table 4: Premium to Assembly in Repeat Sales Framework**

Sources: See Data Appendix for complete information.

Notes: The dependent variable in these regressions is log(real sales price per square foot). We use the largest consistent sample of repeat sales in the three years before teardown or assembly, where the teardown or assembly occurs between 1999 and 2011. Standard errors clustered at the city level are given in parentheses. Significance levels are denoted by \*\*\* for significant at the 0.1% level, \*\* for significant at the 1% level, \* for significant at the 5% level, and + for significant at the 10% level. 1{Redevelopment} is an indicator variable equal one when the sale is an assembly or teardown sale. Use classifications are as given in the note on Table 2. The capital controls (which are more parsimonious than the set used on Table 3) include capital per lot square foot (assessed improvements / lot size), the capital to land ratio (assessed improvements / assessed land value), quadratics in structure square feet per lot square foot, structure age (set to zero if missing), an indicator variable for structure age missing, the interaction of structure square feet per lot square foot with structure age, and the interaction of structure square feet per lot square foot with the indicator for structure age missing. Coefficients on the capital variables are allowed to vary by residential and non-residential pre-assembly use. The city\*year effects include a category for unincorporated portions of LA county.

|  | Non-Single<br>Family | Non-<br>Residential | Single<br>Family |
|--|----------------------|---------------------|------------------|
|  | (1)                  | (2)                 | (3)              |
| 1{Parcel in an Assembly}                   | 0.33*                | 0.32*               | 0.44***          |
|  | (0.13)               | (0.16)              | (0.09)           |
| 1{Parcel in an Assembly}*1{Owner Occupied} |                      |                     | -0.17            |
|  |                      |                     | (0.11)           |
| Observations                               | 4,334                | 3,573               | 4,225            |
| Assemblies                                 | 2,227                | 1,887               | 406              |
| Owner-Occupied                             |                      |                     | 241              |
| Teardowns                                  | 2,107                | 1,686               | 3,819            |
| Geographic Fixed Effects                   |                      |                     |                  |
| Tract                                      | Х                    | Х                   | Х                |
| Additional Covariates                      |                      |                     |                  |
| Year-Quarter of Sale                       | Х                    | Х                   | Х                |
| Year-Quarter of Sale * Non-Residential     | Х                    |                     |                  |
| Use Classifications                        | Х                    | Х                   |                  |
| Cubic in Lot Size                          | Х                    | Х                   | Х                |
| Neighborhood Demographics                  | Х                    | Х                   | х                |
| Distance to Key Amenities                  | Х                    | Х                   | Х                |

#### Table 5: Do Reservation Prices Explain the Premium to Assembly?

Sources: See Data Appendix for complete information.

Notes: The dependent variable in these regressions is log(real sales price per square foot). We use the sample of sales in the three years before teardown or assembly, where the teardown or assembly occurs between 1999 and 2011. Standard errors clustered at the city level are given in parentheses. Significance levels are denoted by \*\*\* for significant at the 0.1% level, \*\* for significant at the 1% level, \* for significant at the 5% level, and + for significant at the 10% level. Use classifications, amenities and demographics are as the note to Table 2.

|                                    | (1)       | (2)       | (3)       | (4)       |
|------------------------------------|-----------|-----------|-----------|-----------|
| Log(Lot Size)                      | 0.012***  | 0.011***  | 0.010**   | 0.010**   |
|                                    | (0.003)   | (0.003)   | (0.004)   | (0.003)   |
| Observations                       | 2,139,621 | 2,139,621 | 2,139,621 | 2,139,621 |
| Share Ever Assembled               | 0.014     | 0.014     | 0.014     | 0.014     |
| Geographic Fixed Effects           |           |           |           |           |
| Tract                              | х         |           | Х         |           |
| Block Group                        |           | х         |           | Х         |
| Additional Covariates              |           |           |           |           |
| Use Classifications                | х         | х         | Х         | Х         |
| Neighborhood Demographics          | х         |           | Х         |           |
| Distance to Key Amenities          | х         |           | х         |           |
| Capital Controls                   |           |           | Х         | Х         |
| Capital Controls * Non-Residential |           |           | x         | x         |

#### Table 6: Effect of Lot Size on Likelihood of Assembly

Sources: See Data Appendix for complete information.

Notes: The dependent variable is an indicator variable equal 1 if the parcel is ever engaged in assembly from 1999 to 2011; results are estimated via OLS (linear probability model). The sample is the cross-section of all parcels that exist in 1999, and excludes public land. Standard errors clustered at the city level are given in parentheses. Significance levels are denoted by \*\*\* for significant at the 0.1% level, \*\* for significant at the 1% level, \* for significant at the 5% level, and + for significant at the 10% level. Use classifications are as given in the note to Table 2. Census demographic variables and amenities are as listed in the notes to Table 2. The capital controls include the value of improvements to the land per square feet, the ratio of the value of improvements to the land to the value of the land (i.e., capital to land ratio) and a full set of interactions of indicators for decile of structure age with indicators for decile of structure square feet divided by lot square feet (for both deciles an 11<sup>th</sup> indicator variable is added to denote missing values).

|  | (1)       | (2)       |
|--|-----------|-----------|
| Log(Lot Size)                          | -0.825*** | -0.801*** |
|  | (0.191)   | (0.166)   |
|  |           |           |
| Observations                           | 2,916     | 2,916     |
| Number of Assembly Groups              | 1,358     | 1,358     |
| Geographic Fixed Effects               |           |           |
| Assembly Group                         | Х         | Х         |
| Additional Covariates                  |           |           |
| Year-Quarter of Sale                   | Х         | Х         |
| Year-Quarter of Sale * Non-Residential | Х         | Х         |
| Use Classifications                    | Х         | Х         |
| Distance to Key Amenities              |           | Х         |

### Table 7: Elasticity of Sales Prices with Respect to Lot Size

Sources: See Data Appendix for complete information.

Notes: The dependent variable in these regressions is log(real sales price per square foot). The sample is the same as in Table 2, but omits teardowns. Standard errors clustered at the city level are given in parentheses. Significance levels are denoted by \*\*\* for significant at the 0.1% level, \*\* for significant at the 1% level, \* for significant at the 5% level, and + for significant at the 10% level. Use classifications and distance to amenities are as given in the note to Table 2.

Appendix Figure 1: Map of Los Angeles County Census Tracts by Redevelopment Status



Notes: Each polygon in this map is a census tract, which is defined to have roughly between 3,000 and 4,000 people.

Appendix Figure 2: Los Angeles County Has Substantial Price Dispersion, and Higher Prices than Most of US



Prices in Los Angeles County

Los Angeles County Prices Relative to US



Note: The upper panel uses the Assessor sales data to illustrate dispersion across the county over time. The top line in the figure reports the 95th percentile of the log real price per square foot of land per year; subsequent lines report the 75th percentile, the mean, the 25th percentile and the 5th percentile. The bottom panel uses Zillow's estimates of mean price per square foot of structure by zip code and month to calculate annual averages by county and metropolitan area. The ten largest metros in the figure are the counties of New York, Chicago, Dallas-Fort Worth, Houston, Philadelphia, Washington, Miami-Fort Lauderdale, Atlanta and Boston.

|                                    | Maximal Sample |           | Sample with f | Sample with full covariates |  |  |
|------------------------------------|----------------|-----------|---------------|-----------------------------|--|--|
|                                    | Ever           | Never     | Ever          | Never                       |  |  |
|                                    | Assembled      | Assembled | Assembled     | Assembled                   |  |  |
|                                    | (1)            | (2)       | (3)           | (4)                         |  |  |
| Log(Lot Size)                      | 9.728          | 8.912     | 9.667         | 8.783                       |  |  |
|                                    | [1.569]        | [1.093]   | [1.393]       | [0.848]                     |  |  |
| Improvement/Land Assessed Value    | 36.878         | 3.937     | 59.933        | 4.285                       |  |  |
|                                    | [2156.9]       | [864.9]   | [2751]        | [902.7]                     |  |  |
| 1 if Single Family                 | 0.257          | 0.643     | 0.400         | 0.699                       |  |  |
|                                    | [0.437]        | [0.479]   | [0.49]        | [0.459]                     |  |  |
| Structure Square Feet in 100,000s  | 0.080          | 0.027     | 0.121         | 0.029                       |  |  |
|                                    | [0.421]        | [0.149]   | [0.514]       | [0.155]                     |  |  |
| Year Structure Was Built           | 1960.0         | 1955.2    | 1960.4        | 1955.2                      |  |  |
|                                    | [28.881]       | [21.613]  | [28.565]      | [21.612]                    |  |  |
| Census 2000 Block Group Covariates |                |           |               |                             |  |  |
| Poverty Rate                       | 0.128          | 0.135     | 0.133         | 0.134                       |  |  |
|                                    | [0.124]        | [0.113]   | [0.129]       | [0.113]                     |  |  |
| Share Black                        | 0.068          | 0.086     | 0.072         | 0.088                       |  |  |
|                                    | [0.119]        | [0.162]   | [0.122]       | [0.166]                     |  |  |
| Share Hispanic                     | 0.292          | 0.353     | 0.317         | 0.359                       |  |  |
|                                    | [0.271]        | [0.284]   | [0.28]        | [0.286]                     |  |  |
| Share Housing Units Vacant         | 0.050          | 0.043     | 0.046         | 0.038                       |  |  |
|                                    | [0.062]        | [0.053]   | [0.053]       | [0.037]                     |  |  |
| Share Housing Units                | 0.619          | 0.606     | 0.605         | 0.600                       |  |  |
| Owner-Occupied                     | [0.296]        | [0.262]   | [0.308]       | [0.264]                     |  |  |
| Distances to Amenities             |                |           |               |                             |  |  |
| Nearest Highway Entrance           | 1.741          | 1.637     | 1.342         | 1.272                       |  |  |
| or Exit                            | [2.907]        | [2.683]   | [1.771]       | [1.573]                     |  |  |
| Nearest Subway or Light Rail       | 10.362         | 8.172     | 8.743         | 7.040                       |  |  |
|                                    | [10.717]       | [9.316]   | [9.462]       | [7.506]                     |  |  |
| Nearest Commuter Rail              | 5.919          | 6.195     | 5.552         | 5.929                       |  |  |
|                                    | [4.891]        | [4.982]   | [4.414]       | [4.703]                     |  |  |
| Observations                       | 26,907         | 2,157,067 | 16,538        | 1,979,876                   |  |  |

# Appendix Table 1: Summary Statistics for 1999 Cross-Section

Sources: See Data Appendix.

## Data Appendix to "From Today's City to Tomorrow's City: An Empirical Investigation of Urban Land Assembly"

## NOT FOR PUBLICATION

Leah Brooks Research Division Federal Reserve Board of Governors

Byron Lutz Research Division Federal Reserve Board of Governors

August 2013

This appendix describes how we created the Los Angeles County parcel dataset. Section 1 describes the input datasets. Section 2 describes how we join these datasets, and reports statistics on the uncleaned data. Section 3 describes how we clean the joined dataset, and Section 4 reports statistics on the quality of the final dataset.

## 1 Data Sources

The basic unit of analysis is the parcel, which is an individual property as legally defined by the Los Angeles County Assessor and Recorder. In any year, there are roughly 2.3 million parcels in the County of Los Angeles. We rely on a number of different sources for information about parcels.

### 1.1 Parcel-Level Data

For detailed property information on parcels, we rely on data from three separate vendors: DataQuick, Applied Geodetics, and the Los Angeles County Assessor directly.

#### 1.1.1 DataQuick: 1999 - 2002

Dataquick is a property information vendor. It purchases property information from the Los Angeles County Assessor to sell to real estate professionals. We rely on Dataquick data for 1999-2002, and data are reported as of January of each year.

As far as we can ascertain, the Dataquick data are a slightly modified version of the file that the Assessor prepares (specifically, the "Secured Basic File" we discuss below). Dataquick modifies the data from the original Los Angeles County format, and we re-modify it to be consistent with the two following datasets. We discuss modifications at length in Section 3. This data vendor is abbreviated in tables as DQ.

#### 1.1.2 Applied Geodetics: 2004 - 2006

Applied Geodetics is a mapping firm in Los Angeles County. Applied Geodetics sold us data for 2004 (April), 2004 (February) and 2005 (May). To the best of our knowledge, these data are the unmodified Assessor's "Secured Basic File," which is the most complete record of property attributes available to the public from the Assessor. This data vendor is abbreviated in tables as AG.

#### 1.1.3 Los Angeles County Assessor, Local Roll: 2007 - 2009

From 2007 through 2009, we have purchased data directly from the Assessor. Due to financial constraints, we purchased the "Local Roll" database (roughly \$400) instead of the "Secured Basic File" (roughly \$13,000). The Local Roll has fewer parcel attributes than the Secured Basic File, and comes out annually in July. This data source is abbreviated in tables as LR.

#### 1.1.4 Los Angeles County Assessor, Secured Basic File: 2010 and 2011

In 2010 and 2011, we purchased the Secured Basic File, which is the County's most complete publicly available dataset about properties. These data are from July of each year, and this data source is abbreviated in tables as SB.

### 1.2 Sales Data

#### 1.2.1 Last Three Sales, 1980 to 2006

In 2006, we purchased a file from the County Assessor that contains information on the last three transactions for each property in the county. For each transaction, we observe transaction type, sale amount (if applicable), and date of transaction.

#### 1.2.2 Sales Within Two Years: 2008, 2009, 2010, and 2011 Files

In 2008, 2009, 2010, and 2011, we purchased additional lists of sales data from the County Assessor. These contain information on all transactions in the prior two years. For each transaction, we observe transaction type, sale amount (if applicable), and date of transaction.

These files leave a small gap from May through December 2006 which we have not been able to fill.

#### **1.3** Parcel Change Database

At our request, the Assessor made a special file that includes all parcel changes from July 1999 to January 2009. Specifically, for each change, this file includes the old parcel number(s), the new parcel number(s), and the effective date of the change. The County has electronic records for parcel changes starting in July 1999 and continuing to the end of our data.

This change database allows us to isolate land assembly and disassembly. The California Assessor's Handbook mentions only one reason for a parcel number to change: if the physical boundaries of a parcel are modified (California State Board of Equalization, 1997, page 26).

We purchased this change database again in July 2009, 2010 and 2011 (each covers all changes in the past two years) to allow us to link all later parcels with previous parcels.

#### 1.4 Digital Parcel Maps

For each year since 2006, we have an electronic map of all parcels that exist in that year. These maps have a boundary (a polygon) for each individual parcel. For each parcel, we use ArcGIS to calculate the x- and y-coordinates (latitude and longitude) of the polygon's geographic center (centroid).

### **1.5** Census Tract and Block Group Identification

The Census provides census tract and block group boundaries in shapefile format.<sup>1</sup> We use ArcGIS to intersect the 2000 census boundaries with the parcel boundaries from 2006 to 2011 to assign each parcel to a census block group.

The vast majority – 98 percent – of ever-existing parcels have block group identifiers.

#### **1.6 Block Group Data**

We use demographic block group level data from the 2000 Decennial Census (ICPSR file 13346, summary level 150), and from the 1990 Decennial Census (ICPSR 9782, summary level 150, but California file is damaged so we used a similar file downloaded from UCLA ATS).

We use ArcGIS and the Census's electronic maps to make a linkage between 1990 and 2000-based block groups, where relationships are based on land area overlap.

### 1.7 Non-Parcel Digital Maps

- Parks: Information from 2008 ESRI files of local and national parks for California
  - parks displayed on maps are only those more than 1.25 square miles
- Freeways: Data from State of California Cal-Atlas Geospatial Clearinghouse
  - website is http://www.atlas.ca.gov/download.html
  - transportation  $\rightarrow$  Census 2000  $\rightarrow$  state\_highways.\* and us\_highways.\*
- Freeway Entrances and Exits
  - Tele-Atlas US Data, contains federal interstate highway entrances and exits
- Coastline
  - layer of points every 1000 feet along LA County coastline

<sup>&</sup>lt;sup>1</sup>Files are at http://www.census.gov/geo/www/cob/bdy\_files.html.

- created by taking Census 2000 county map and deleting non-coastline portions
- used X-tools feature to points to convert coastline line to points
- Metrolink Stations
  - Commuter rail stations
  - File received from Javier Minjares, Southern California Association of Governments, 2010
- Metro Rail Stations
  - Intra-urban rail stations
  - File received from Javier Minjares, Southern California Association of Governments, 2010
- Major Roads
  - Tele-Atlas US Data, version 9.3
  - Major roads only
- Municipal Boundaries
  - Shapefile from 2000 Census

# 2 Initial Data Linking

Each parcel is identified by a 10-digit number: MMMM-PPP-XXX. The first four digits are the "map book" number – literally the number of the "book" in which the parcel appears. The second three digits are the map book page and identifies the page in the book on which the parcel is found. Each map book page contains a set of geographically contiguous parcels. The last three digits identify individual parcels on the map book page.

We began the data assembly by attempting to link all the annual cross-sections described above (1999-2009, without the missing 2003) using the parcel change database. Panel A of Table 1 presents the results of this original linkage. Slightly fewer than 2 million parcels never change their number throughout the sample (column 2). Column 3 reports the number of ever-changing parcels; this number varies by year as a changer could be 3 parcels in 1999, but 12 (or 1) parcels in 2009. The number of changes in any given year varies from 35,627 to 157,873. Column 4 reports the number of parcels that exist in this year of the sample, but not in all years of the sample (following a specific definition we describe later). Column 4 shows a striking number of parcels that exist after 2004, but not before. These missing, or "phantom," observations are due to DataQuick editing. We discuss how we deal with these discrepancies in the next section.

We also analyze total land area, measured in square feet, in each category as a check on our data assembly process. We report annual totals in Panel B of Table 1. Columns 2 through 4 report the total amount of parcel square footage in each of the parcel assembly categories by year. Unfortunately, for 2004 and 2005 we do not observe parcel size in the data. It is very clear that across all categories, but particularly in the "phantom" category, that the total land area is drastically smaller in the earlier data source (DataQuick). The California Department of Water Resources measures the total land area of Los Angeles County as 132,487,077,888 square feet, which is clearly much closer to the later data source (143 billion square feet) than the earlier one (22 billion square feet).<sup>2</sup> As we discuss later in great detail, we undertake two major modifications to make total County lot size consistent over time. First, we make sure that lot size for condominiums is coded consistently across all years. Second, we add in parcels that (mostly) do not appear in the DataQuick data that very likely existed at the time of these data (1999 to 2002).

Table 2 shows the total number of parcels (Panel A) and total land area (Panel B) by four categories of use types and by year. The Assessor assigns each property a "use code" that describes how the property is currently used, as distinct from how it is zoned. We use this code to make four major categories of use: single-family, multi-family, vacant, and other. The number of vacant parcels by year (column 4) shows a large break concomitant with the

<sup>&</sup>lt;sup>2</sup>Parcels cover all land and water area of the county with the exception of public roadways. We discuss later how the latter datasets "overcount" the county's land area.

break in data sources, and similar to the break for the "phantom" parcels of the previous table.<sup>3</sup> The bottom panel of the table shows that the break appears across types.

# 3 Cleaning and Consistency Changes

This section details the work we did to make the property data consistent across data sources and time.

### 3.1 Defining the Parcel Group

For ease of analysis, we define a "parcel group" to include all parcels linked by a change. For example, consider a change where parcels A and B combine to C, parcel C splits to D and E, and E splits to F and G. All parcels A through G are in one parcel group, following our definition. We create a unique identifier for each parcel group.

#### 3.2 No Observation of Land Area in 2004 and 2005

The data source we use for 2004 and 2005 does not contain information on parcel land area. We assume that all 2004 parcels existing in 2002 have the same land area that they did in 2002, and all 2005 parcels existing in 2006 have the same land area they will have in 2006. These are very reasonable assumptions, as lot size changes only when parcel number does. As we describe in greater detail below, we use a variety of methods to evaluate and clean lot size across time.

If a 2004 or 2005 parcel is also missing lot size information in 2002 or 2006, we take lot size from the closest year for which it is available.

 $<sup>^{3}</sup>$ Vacant parcels may be coded vacant residential, vacant commercial, or simply vacant. "Vacant" appears to refer to habitation or use, and not to the presence of structures.

#### **3.3** Vertically Stacked Parcels

The assessor draws the boundary for a single piece of land with multiple owners, such as condominiums, so that each owner's parcel has the land area of the entire lot. This type of separate ownership of the same piece of land is distinct from joint ownership of a single property. Separate ownership, or condominium-type ownership, parcels are "stacked" vertically on the Assessor's map, in the same way you would stack checkers. In this case, think of each checker as the condominium of an individual owner. A 100-unit condominium will consist of 100 parcels, one for each legally distinct unit of ownership. Stacked parcels are usually, but not always, condominiums.

In the DataQuick data, the land area of each condominium unit is the land area of the lot divided by the number of parcels covering that lot. In the remaining datasets, when units and parcels are co-terminus, each unit is assigned the land area of the entire lot. This causes problems with multifamily parcels in particular, and is one of the two causes of the discrepancy in land area between the 1999 to 2002 and 2003 to 2010 periods in Tables 1 and 2. We correct for this problem by modifying all parcels to have a land area consistent with the method in the DataQuick data (land area of a parcel = total land area of parcel / number of parcels occupying that lot). We prefer this method of measuring land area because the net land area of the county remains unchanged even if one parcel becomes three condos. The method used in the later datasets suggests that the total land area of the county increases when condominiums are created.

We identify vertically stacked parcels by calculating the centroid of every parcel on each map we have (annual maps, 2006-2010). In any given year, all parcels that share the same x-and y-coordinates are "stacked," and their land area requires correction. On average, a map-year has 251,518 unique parcels that are stacked, 15,299 stacks, and a mean of 16 parcels per stack. Unfortunately, we do not have maps for years before 2006, so this correction is incomplete for 2004 and 2005. Specifically, we can correct all stacked parcels that exist in

2004 and 2005 and which continue to exist in 2006. If parcels exist in 2004 and 2005 only – with no 2006 data – we are not able to re-calculate a lot size for those parcels. (Section 3.6 addresses how we handle observations which we are missing lot size information.)

## 3.4 "Phantom" Parcels

The third data challenge we faced was the presence of parcels that did not appear in all years, and the disappearance of which could not be explained by any observation in the parcel change database. We call these appearing and disappearing parcels "phantoms." Specifically, to be a "phantom," in our terms, a parcel must appear in at least five years of data.

Table 1 suggests that the DataQuick data do not include all parcels that exist in the county. We strongly suspect that, in an effort to satisfy customers primarily interested in residential (or at least tradeable) property, DataQuick deletes parcels that will not be transacted.

Table 3 describes the appearance of phantoms across the years of the final dataset. The horizontal axis has one column for each year of the dataset. The vertical axis has one row for each year of phantom non-appearance. The (1999,2000) cell says that 649 parcels that exist in the 1999 dataset do not exist in the 2000 dataset and are not accounted for by any parcel changes. The diagonal axis is by definition zero, since a parcel cannot be missing in the year in which it exists. The clear and striking pattern in this table is that the vast majority of phantom parcels are missing in the years 1999 to 2002, and come from datasets in years 2004 to 2010.

To resolve this problem, we say that any parcel that exists for at least five years in the dataset is a "true" parcel. Because these parcels are not in the change database, we assume that their size is constant over the period of interest and that they exist in all years. Roughly 200,000 of these parcels truly exist from 2004 to 2010 and do not appear from 1999 to 2002.

Only about 20,000 parcels appear from 1999 to 2002 and not from 2004 to 2010. Were we not to make this adjustment, we would have an severely unbalanced panel, and the early years of the panel would not account for a substantial amount of the land area of the County.

### 3.5 Relating Parcel Change to Parcel Cross-sectional Databases

We use the assessor's parcel change database to link parcels across years. The assessor reports on changes of four types: one to many, one to one, many to one, zero to one and one to zero. These last two changes (zero to one and one to zero) are the least frequent of all change types. They are used to move land in and out of the county. Land in use by roads is not in parcels, so the creation of a parcel from a road could cause a zero to one change. Similarly, a one to zero change could occur when a parcel becomes part of a road.

Table 4 reports on the quality of the match between the datasets. The table reports statistics for all changing parcels, and by change type. Columns 2 and 3 show that more parcels enter the county than leave over the period. Columns 4 and 5 report the number of parcels from the change database that find a match in the cross-sectional database, and shows whether these parcels are entering or leaving the county.

Column 8 reports the share of parcels from the change database that we consider "successfully matched." Across all years, we are able to match 97 percent of parcels in the change database with parcels in the cross-sectional databases. Matches can be one of two types. The first type of match is when a parcel in the change database directly matches a parcel in the cross-section. These are a majority of the matches, as shown in Column 6. However, many-to-many changes (and some others) require the use of placeholder parcels; this is the second type of match. The Assessor's file structure does not allow for a direct many-to-many match. Instead, many-to-many matches are coded as many-to-one and then one-to-many. In this case, the "one" is the placeholder parcel. We report the share of matches to placeholder parcels in Column 7. Across all types of changes, roughly 17 percent of changing parcels match to a temporary placeholder parcel. As shown in Column 9, only three percent of parcels have no match at all. The final column in the table reports the time from appearance in the parcel change database to appearance in the cross-sectional database. Parcels usually take about one year to appear in the cross-section database after appearing in the change database.

To date to a parcel change, then, we must make a choice between two possible dates – one from the cross-sectional data, and one from the parcel change database. In cases where the parcel change database and the cross-sectional databases do not agree, we defer to the change database. The change database records an administrative, or policy change, while the cross-sectional data more likely reflects what is actually physically changing. Since we are most interested in policy choices regarding land use, the date of the decision to change a parcel seemed more relevant than the physical change itself.

Finally, we drop a very small number of parcels we cannot link to other parcels in the cross section to form a change group. Specifically, we drop 4,812 observations (470 unique parcel numbers) that do not fit into a change group.

#### **3.6** Reported Changes in Land Area for Non-Changing Parcels

We check the internal consistency of the data by examining whether there are changes in land area for parcels that do not change parcel number across the years of the sample. By construction, such parcels should have the same land area for all years of the sample. For each parcel, we calculate the percentage difference between the largest land area and the smallest land area reported (relative to the mean lot size of the entire period). Though the median change is a very small 0.7 percent, changes at the ninetieth percentile and above can be quite large.

Given this, we make two adjustments to clean the land area variable. First, if a parcel's land area changes by less than 20 percent from 1999 to 2011, we give the parcel the 2011

land area value for each year. We rely on data from the later years of the sample because by all measures we report here it is more reliable than the DataQuick data. This change modifies the values for virtually all of the parcels in the dataset. Of the modified parcels, about 19 percent of them have very minor differences of less than 0.1 percent, which seem to be due to rounding differences between the different datasets. Another 60 percent have differences between 0.1 and 5 percent of land area over time, and 6 percent have differences of between 5 and 20 percent.

The second adjustment to clean the land area variable modifies parcels that are vertically stacked. After correcting for the different lot size definitions, if the percentage change in lot size for the same parcel is less than 20 percent, we replace the area for all years with the average across all years in the sample.

Finally, we delete 74,787 unique parcels, between 2 and 4 percent of the observations in each cross-section, that have land area changes of more than 20 percent, or which have always-missing values for land area.

#### 3.7 Land Area Adjustments for Parcel Changes

When parcels change, the total initial lot size of changing parcels should equal the final total lot size of the changing parcels. Unfortunately, this is not always the case.

For example, if a single parcel changes number and becomes only one new parcel, the parcel's geographic area does not change by construction. However, due to dataset breaks and other inconsistencies, differences in lot size sometimes appear. These create inconsistencies in geographical areas over time that do not reflect actual changes to parcels.

We resolve this problem by inflating or deflating each parcel's lot size so that the total lot size for parcels that change within a parcel group is the same for each year of data. Let g be the parcel group, which is all parcels ever associated with a given change. The adjustment is

lot size<sub>*i,g,t*</sub> = reported lot size<sub>*i,t*</sub> \* 
$$\frac{\text{average total lot size}_g}{\text{total lot size}_{g,t}}$$
.

The final value for parcel *i* in year *t* (lot size<sub>*i*,*t*</sub>) is equal to parcel *i*'s reported square footage in year *t* (reported lot size<sub>*i*,*t*</sub>) multiplied by the ratio of the average of the parcel group's total square feet (average total lot size<sub>*g*</sub>) divided by the current year total square feet (total lot size<sub>*g*,*t*</sub>). By adjusting at the parcel group level, this method allows us to account for both simple and complex changes.

The result of this adjustment is that each parcel group has a constant lot square footage across time. This better reflects the reality of a parcel change. Square footage at the beginning should equal square footage at the end, regardless of how the land is divided or combined.

#### **3.8** Missing Values for Parcel Attributes

In addition to the work with lot square feet that we do above, we fill in missing values for parcel attributes to a limited extent. For non-geographic information, we edit parcels that do not change number, and for which a later observation of the parcel exists. In these cases, we attribute the later parcel characteristics to the earlier parcel if the structure on the parcel was built before the year of the data. For example, imagine parcel A exists from 1999 to 2009, but we observe attributes starting only in 2005. If the structure on parcel A was built in 1995, we fill in the missing values for parcel A from 1999 through 2004 with the 2005 values. If the structure on parcel A was built in 2005, we do not fill in missing values (save for lot square footage as described above).

Geographic identifying information – tract and block group information, and latitude and longitude information – come from merging the panel data with information generated from electronic maps. We have these electronic maps starting only in 2006. For parcels that do not change, we know that the physical border has not changed and therefore use the 2006 geographic information.

For parcels that existed before 2006 and not afterward we impute geographic information. Specifically, we assign old parcels to the same block group or tract as the later-existing parcels (no parcel change groups cross block groups). For the latitude and longitude, we give the old parcels the average of the latitude and longitude of the later-existing parcels. This should be a very close approximation to the actual location.

For parcels that are not assigned a block group through these methods (changing or not), we interpolate to find the block group. Specifically, if a parcel is missing a block group, we assign the parcel to have the block group of all other parcels on that map book page. A map book page is identifiable by the first seven digits of the parcel number. Map book pages have very few parcels and very rarely cross block group boundaries.

#### 3.9 Determining Owner Occupancy Status and Housing Tenure

In order to determine whether a parcel is owner-occupied, we compare the parcel's physical address to the parcel's mailing address used for tax purposes (i.e., where the property tax bill is sent). If the two addresses match, we infer that the property is owner-occupied. In practice this procedure is complicated by the fact that the address data is often missing or incomplete. We therefore use an algorithm to classify each parcel into one of three categories: match, non-match, and insufficient data.

The address fields in our data are state, zip code, city name, street name and street number. We categorize a parcel as a match, and hence as owner-occupied, if the street number is not missing *and* either of the following two conditions pertaining to the mailing and physical addressed are met:

1. The street number in the first N places is the same in the physical and mailing address, where N is the length of the shorter of the two street numbers (if N < 3 then the length

of the two street numbers must also be equal), *and* the first three letters of the street name are the same in both the physical and mailing address.

- 2. The same street number in the first N places is the same in the physical and mailing address (if N < 3 then the length of the two street numbers must also be equal) and the physical and mailing addresses are equal for any two of the following :
  - First letter of the street name
  - First letter of city name
  - First letter of state name
  - First digit of the zip code

We categorize a parcel as a non-match, and hence as not owner-occupied, if it is not categorized as a match *and* any one of the following conditions for the mailing and physical addresses are met:

- Both state names are valid and differ *and* one of the following:
  - First three letters of city name are not identical (and neither is missing)
  - First three letters of street name are not identical (and neither is missing)
  - First two digits of zip code are not identical (and neither is missing)
  - First two digits of street number are not identical (and neither is missing)
- Mailing address is a PO Box
- Street numbers differ in the first N places and are not missing
- Both zip codes are valid, differ *and* one of the following:
  - First three letters of city name are not identical (and neither is missing)
  - First three letters of street name are not identical (and neither is missing)
  - First two digits of street number are not identical (and neither is missing)
- The first four characters of the street name do not match *and* the first four characters of the city name do not match
- Street name matches exactly and the street number does not

We categorize a parcel as having insufficient data if it is neither a match nor a non-match.
Among the subset of parcels identified as being owner-occupied, we infer housing tenure from the number of years since the house last sold. For instance, an owner-occupied single family house which last sold in 2000 is categorized as having housing tenure equal to 5 in 2005.

#### 3.10 Determining Which Assembly Sales are Teardowns

For a given assembly, pre-assembly, we observe attributes for one structure per parcel. However, a parcel may, in fact, contain more than one structure. For each parcel change group before assembly, we find the minimum and maximum of the age of the structure on each parcel:  $\max(age_{g,before})$  and  $\min(age_{g,before})$  (g indicates a parcel change group). We observe similar ages post-assembly:  $\max(age_{g,after})$  and  $\min(age_{g,after})$ . We define parcels as being assembly teardowns only if  $\max(age_{g,before}) < \min(age_{g,after})$ . This definition therefore excludes change groups where some, but not all, structures were demolished.

### 3.11 Preparing Sales Data

#### 3.11.1 Basic Cleaning

To prepare the sales data, we restrict our analysis to arm's length transactions. We drop sales with dates not on any calendar (e.g., day 45; this is a problem only for the 2006 dataset), sales with values of 10 or less (which the assessor uses as a special code for adjoining pools or structures), and sale amounts equal to 999,999,999. We drop sales before 1985. We keep only one observation if the data report two sales on the same day of the same amount. A small share of observations have two sales on the same day of differing amounts. In this case, we drop both observations. Finally, we drop all sales with a value of less than \$0.15 per square foot, as we believe that this value is too small to be credible. This restriction drops less than 1 percent of the remaining sales.

#### 3.11.2 Multi-sales

In some cases, parcels do not sell individually. For example, suppose that three parcels sell as a group at one time. In such a case, the assessor observes only the total purchase price. In these cases, the Assessor codes the sale as a "multi-sale." Multi-sales are 6 percent of all sales.

The Assessor codes multi-sales as 2 through 9 or M (more than 9) to designate the number of parcels in the multi-sale. In the raw data, the assessor puts the total price as the price for each parcel, therefore overstating total sale amounts.

We correct for this overstatement by re-scaling prices. To begin, we double-check the assessor's coding of multi-sales. We count the number of sales that occur on the same day, on the same map book page, for the same sale amount. Any sales that are duplicates by these terms we designate as multi-sales. If we count the same number of parcels in the multi-sale as the Assessor reports, we call this a verified multi-sale. Our analysis uses only verified multi-sales and drops non-verified multi-sales.

We correct sale values for multi-sale properties by giving each property a share of the total sale price equal to its share of the lot square footage of all the properties in the group.

### 4 Cleaning and Consistency Results

Tables 5 and 6 present the results of the data cleaning. In the first of these tables, Panel A shows that the overall pattern of the number of parcels over time no longer exhibits any breaks when we change data sources. The number of phantom parcels and never-changing parcels is now consistent across all years. The number of changing parcels increases over time: there are roughly 60,000 of these parcels in 1999 and 160,000 of them in 2011. This means that, on net, more parcels have split than merged. Panel B shows that this increase is not driven by substantial changes in land area. This table no longer shows discontinuous

patterns across dataset breaks.

Panel A of Table 6 shows that, from 1999 to 2011, the number of single- and multifamily and vacant parcels increased, while the number of "other" parcels decreased. Land area trends, shown in Panel B, also no longer show sharp discontinuities related to dataset breaks.

# References

California State Board of Equalization, 1997. "Assessor's Handbook 215: Assessment Map Standards." Tech. rep., California State Board of Equalization.

|                  |        | (1)       | (3)            | (4)      |          |  |  |  |  |  |  |
|------------------|--------|-----------|----------------|----------|----------|--|--|--|--|--|--|
| Panel A: Parcels |        |           |                |          |          |  |  |  |  |  |  |
| Year             | Source | Total     | Never-Changers | Changers | Phantoms |  |  |  |  |  |  |
| 1999             | DQ     | 2,054,630 | 1,948,502      | 35,627   | 6,187    |  |  |  |  |  |  |
| 2000             | DQ     | 2,062,666 | 1,948,502      | 38,239   | 11,294   |  |  |  |  |  |  |
| 2001             | DQ     | 2,087,855 | 1,948,502      | 46,323   | 26,114   |  |  |  |  |  |  |
| 2002             | DQ     | 2,087,417 | 1,948,502      | 54,704   | 17,449   |  |  |  |  |  |  |
| 2004             | AG     | 2,311,387 | 1,948,502      | 86,692   | 206,481  |  |  |  |  |  |  |
| 2005             | AG     | 2,319,526 | 1,948,502      | 94,412   | 207,219  |  |  |  |  |  |  |
| 2006             | AG     | 2,337,233 | 1,948,502      | 111,539  | 207,947  |  |  |  |  |  |  |
| 2007             | LR     | 2,354,414 | 1,948,502      | 128,751  | 208,400  |  |  |  |  |  |  |
| 2008             | LR     | 2,366,832 | 1,948,502      | 140,705  | 208,935  |  |  |  |  |  |  |
| 2009             | LR     | 2,376,360 | 1,948,502      | 150,182  | 209,330  |  |  |  |  |  |  |
| 2010             | SB     | 2,380,386 | 1,948,502      | 154,261  | 209,647  |  |  |  |  |  |  |
| 2011             | SB     | 2,383,988 | 1,948,502      | 157,873  | 209,677  |  |  |  |  |  |  |

### Table 1: Original Dataset: By Type of Appearance in Dataset

#### Panel B: Land Area, Square Feet

| Year | Source | Total           | Never-Changers | Changers       | Phantoms       |
|------|--------|-----------------|----------------|----------------|----------------|
| 1999 | DQ     | 27,764,230,819  | 19,504,676,079 | 1,124,038,343  | 418,477,358    |
| 2000 | DQ     | 22,403,685,543  | 19,505,037,945 | 1,011,368,413  | 546,960,947    |
| 2001 | DQ     | 23,929,788,784  | 19,505,038,780 | 1,224,775,567  | 1,668,968,971  |
| 2002 | DQ     | 22,411,729,309  | 19,505,073,265 | 1,111,671,740  | 403,388,448    |
| 2004 | AG     |                 |                |                |                |
| 2005 | AG     |                 |                |                |                |
| 2006 | AG     | 138,025,141,417 | 45,151,780,276 | 9,510,647,933  | 72,011,080,923 |
| 2007 | LR     | 141,249,361,785 | 45,154,323,925 | 10,296,129,745 | 72,162,588,150 |
| 2008 | LR     | 142,126,764,819 | 45,153,989,843 | 11,197,350,796 | 72,244,152,453 |
| 2009 | LR     | 142,791,412,096 | 45,153,547,001 | 11,765,966,573 | 72,389,504,481 |
| 2010 | SB     | 143,083,855,131 | 45,153,039,051 | 12,139,840,358 | 72,466,658,923 |
| 2011 | SB     | 143,215,366,799 | 45,131,956,782 | 12,500,698,840 | 72,345,117,341 |

Note: Columns 2, 3 and 4 do not add up to the total in Column 1. Column 1 reports the total number of parcels in the original data. As described in this appendix, we drop some parcels. These dropped parcels are not classified as being one of the three types in Columns 2, 3 and 4.

|                  |        | (1)       | (2)           | (3)          | (4)     | (5)     |  |  |  |  |
|------------------|--------|-----------|---------------|--------------|---------|---------|--|--|--|--|
| Panel A: Parcels |        |           |               |              |         |         |  |  |  |  |
| Year             | Source | Total     | Single-Family | Multi-Family | Vacant  | Other   |  |  |  |  |
| 1999             | DQ     | 2,054,630 | 1,399,650     | 436,693      | 16,332  | 127,788 |  |  |  |  |
| 2000             | DQ     | 2,062,666 | 1,400,390     | 436,686      | 16,448  | 127,915 |  |  |  |  |
| 2001             | DQ     | 2,087,855 | 1,400,729     | 436,658      | 28,288  | 127,898 |  |  |  |  |
| 2002             | DQ     | 2,087,417 | 1,401,260     | 436,982      | 15,780  | 127,509 |  |  |  |  |
| 2004             | AG     | 2,311,387 | 1,435,886     | 455,012      | 184,643 | 163,087 |  |  |  |  |
| 2005             | AG     | 2,319,526 | 1,438,901     | 459,453      | 184,088 | 163,251 |  |  |  |  |
| 2006             | AG     | 2,337,233 | 1,448,345     | 469,750      | 182,511 | 163,790 |  |  |  |  |
| 2007             | LR     | 2,354,414 | 1,455,306     | 480,489      | 181,534 | 163,177 |  |  |  |  |
| 2008             | LR     | 2,366,832 | 1,458,101     | 492,126      | 179,772 | 164,284 |  |  |  |  |
| 2009             | LR     | 2,376,360 | 1,459,853     | 501,540      | 178,657 | 164,334 |  |  |  |  |
| 2010             | SB     | 2,380,386 | 1,460,179     | 507,560      | 177,099 | 164,267 |  |  |  |  |
| 2011             | SB     | 2,383,988 | 1,461,338     | 511,307      | 176,493 | 163,971 |  |  |  |  |

### Table 2: Original Dataset: Use Type By Year

#### Panel B: Land Area, Square Feet

| Year | Source | Total           | Single-Family  | Multi-Family   | Vacant         | Other          |
|------|--------|-----------------|----------------|----------------|----------------|----------------|
| 1999 | DQ     | 27,764,230,819  | 13,720,572,239 | 2,692,490,776  | 663,269,634    | 3,970,844,134  |
| 2000 | DQ     | 22,403,685,543  | 13,730,266,220 | 2,690,080,642  | 662,540,286    | 3,940,631,065  |
| 2001 | DQ     | 23,929,788,784  | 13,759,487,498 | 2,688,552,227  | 1,849,059,181  | 3,949,719,842  |
| 2002 | DQ     | 22,411,729,309  | 13,775,928,500 | 2,689,535,220  | 398,909,792    | 3,902,599,372  |
| 2004 | AG     |                 |                |                |                |                |
| 2005 | AG     |                 |                |                |                |                |
| 2006 | AG     | 138,025,141,417 | 18,963,069,965 | 27,976,927,390 | 40,656,243,630 | 38,946,710,876 |
| 2007 | LR     | 141,249,361,785 | 19,175,924,780 | 28,711,136,882 | 40,648,767,481 | 39,015,143,144 |
| 2008 | LR     | 142,126,764,819 | 19,290,606,188 | 29,473,535,434 | 40,622,373,977 | 39,161,389,659 |
| 2009 | LR     | 142,791,412,096 | 19,375,483,660 | 30,192,800,865 | 40,494,279,725 | 39,239,717,086 |
| 2010 | SB     | 143,083,855,131 | 19,397,286,430 | 30,464,889,679 | 40,624,729,739 | 39,265,223,833 |
| 2011 | SB     | 143,215,366,799 | 19,447,253,562 | 30,642,750,267 | 40,691,087,395 | 39,196,681,739 |

Note: Columns 2 through 5 do not add up to the total in column 1, as a very few parcels have no intelligible use codes.

| Total # of parcels that do not appear in vertical axis year, but do appear in this year & for at least five years |       |       |        |        |         |         |         |         |         |         |         |         |
|---|-------|-------|--------|--------|---------|---------|---------|---------|---------|---------|---------|---------|
|   | 1999  | 2000  | 2001   | 2002   | 2004    | 2005    | 2006    | 2007    | 2008    | 2009    | 2010    | 2011    |
|   | (1)   | (2)   | (3)    | (4)    | (5)     | (6)     | (7)     | (8)     | (9)     | (10)    | (11)    | (12)    |
| 1999  | 0     | 7,370 | 25,193 | 20,302 | 223,624 | 222,504 | 219,819 | 218,016 | 215,840 | 214,836 | 213,393 | 212,424 |
| 2000  | 649   | 0     | 20,766 | 17,617 | 222,718 | 221,518 | 218,834 | 217,033 | 214,830 | 213,830 | 212,391 | 211,405 |
| 2001  | 1,034 | 425   | 0      | 11,623 | 209,179 | 208,135 | 205,805 | 204,202 | 202,147 | 201,239 | 199,859 | 198,945 |
| 2002  | 5,406 | 4,962 | 16,085 | 0      | 220,580 | 219,347 | 216,788 | 215,007 | 212,863 | 211,873 | 210,423 | 209,403 |
| 2004  | 2,048 | 1,465 | 1,200  | 1,070  | 0       | 6,924   | 8,326   | 8,934   | 9,601   | 10,219  | 10,561  | 10,793  |
| 2005  | 2,054 | 1,473 | 1,208  | 1,078  | 37      | 0       | 12,621  | 13,231  | 13,872  | 14,510  | 14,822  | 15,022  |
| 2006  | 2,058 | 1,477 | 1,213  | 1,083  | 119     | 94      | 0       | 4,475   | 5,269   | 5,946   | 6,329   | 6,580   |
| 2007  | 2,063 | 1,482 | 1,218  | 1,089  | 159     | 136     | 58      | 0       | 2,225   | 2,967   | 3,400   | 3,703   |
| 2008  | 2,063 | 1,482 | 1,218  | 1,089  | 188     | 168     | 96      | 51      | 0       | 3,337   | 3,825   | 4,129   |
| 2009  | 2,064 | 1,483 | 1,219  | 1,090  | 210     | 190     | 119     | 76      | 36      | 0       | 1,455   | 1,916   |
| 2010  | 2,065 | 1,484 | 1,220  | 1,091  | 225     | 207     | 140     | 103     | 69      | 44      | 0       | 1,372   |
| 2011  | 4,171 | 3,600 | 3,422  | 3,253  | 3,551   | 3,537   | 3,475   | 3,458   | 3,431   | 3,410   | 3,400   | 0       |

**Table 3: Phantom Parcels Across Years** 

Notes: This table reports a count of only phantom parcels. Each cell reports the number of parcels in the year of the horizontal axis that are "phantoms" that we have added in the year of the vertical axis. For example, cell (1999,2004) has a value of 2,048 which means that 2,048 parcels exist in 1999 and at least four other years and do not appear in the 2004 cross-sectional data. This table accounts for changes from the change database. By construction, the diagonal is zero – a parcel that we add to the database because it does not appear (a phantom) cannot exist in the original data in that year.

|             | Parcels ir<br>Data | n Change<br>base  | Parcels in the Cross-<br>Section with match<br>ange in parcel change<br>e database |                   | Sha   | Share of all parcels in Change Database                           |                                    |  |  |  |
|-------------|--------------------|-------------------|--|-------------------|---|---|------------------------------------|--|--|--|
| Change type | Entering<br>County | Leaving<br>County | Entering<br>County   | Leaving<br>County | Match in<br>Cross-<br>sectional<br>database | Temporary<br>parcel is not<br>matched, but<br>change parcel<br>is | Successfully<br>Matched<br>Parcels | No match in<br>Cross-<br>sectional<br>database | Matches:<br>Average years<br>for change to<br>appear in<br>cross-section * |  |
| (1)         | (2)                | (3)               | (4)  | (5)               | (6)   | (7)   | (8)                                | (9)  | (10)   |  |
| All types   | 204,832            | 99,748            | 178,963  | 64,137            | 0.80  | 0.17  | 0.97                               | 0.03   | 1.11   |  |
| Zero to One | 2,400              | 1                 | 661  |                   | 0.28  | 0.62  | 0.90                               | 0.10   | 1.49   |  |
| One to Zero | 1                  | 918               |  | 458               | 0.50  | 0.00  | 0.50                               | 0.50   |  |  |
| One to One  | 32,295             | 32,338            | 26,535   | 23,770            | 0.78  | 0.20  | 0.97                               | 0.03   | 1.18   |  |
| One to Many | 152,914            | 18,025            | 139,377  | 9,456             | 0.87  | 0.11  | 0.98                               | 0.02   | 1.08   |  |
| Many to One | 17,222             | 48,466            | 12,390   | 30,453            | 0.65  | 0.30  | 0.95                               | 0.05   | 1.15   |  |

\* This figure is calculated using only cases where the time for the parcel to appear in the cross-section is positive. In a very small set of cases, a parcel may appear in the cross-sectional data before it appears in the change database, in which case the time for the parcel to appear is negative. Approximately 6 percent of matched parcels have a negative time to appear; 83 percent of these negative values appear in the first year of our sample, 1999.

|                  |        | (1)       | (3)            | (4)      |          |  |  |  |  |  |  |
|------------------|--------|-----------|----------------|----------|----------|--|--|--|--|--|--|
| Panel A: Parcels |        |           |                |          |          |  |  |  |  |  |  |
| Year             | Source | Total     | Never-Changers | Changers | Phantoms |  |  |  |  |  |  |
| 1999             | DQ     | 2,220,095 | 1,948,502      | 59,214   | 212,379  |  |  |  |  |  |  |
| 2000             | DQ     | 2,226,745 | 1,948,502      | 65,864   | 212,379  |  |  |  |  |  |  |
| 2001             | DQ     | 2,232,094 | 1,948,502      | 71,213   | 212,379  |  |  |  |  |  |  |
| 2002             | DQ     | 2,239,496 | 1,948,502      | 78,615   | 212,379  |  |  |  |  |  |  |
| 2004             | AG     | 2,252,467 | 1,948,502      | 91,586   | 212,379  |  |  |  |  |  |  |
| 2005             | AG     | 2,263,834 | 1,948,502      | 102,953  | 212,379  |  |  |  |  |  |  |
| 2006             | AG     | 2,275,592 | 1,948,502      | 114,711  | 212,379  |  |  |  |  |  |  |
| 2007             | LR     | 2,291,112 | 1,948,502      | 130,231  | 212,379  |  |  |  |  |  |  |
| 2008             | LR     | 2,304,070 | 1,948,502      | 143,189  | 212,379  |  |  |  |  |  |  |
| 2009             | LR     | 2,312,035 | 1,948,502      | 151,154  | 212,379  |  |  |  |  |  |  |
| 2010             | SB     | 2,316,179 | 1,948,502      | 155,298  | 212,379  |  |  |  |  |  |  |
| 2011             | SB     | 2,321,904 | 1,948,502      | 161,023  | 212,379  |  |  |  |  |  |  |

## Table 5: Cleaned Dataset: By Type of Appearance in Dataset

### Panel B: Land Area, Square Feet

| Year | Source | Total           | Never-Changers | Changers      | Phantoms       |
|------|--------|-----------------|----------------|---------------|----------------|
| 1999 | DQ     | 101,045,575,947 | 19,953,151,787 | 8,676,099,862 | 72,416,324,298 |
| 2000 | DQ     | 101,043,896,594 | 19,953,151,787 | 8,674,420,509 | 72,416,324,298 |
| 2001 | DQ     | 101,027,910,070 | 19,953,151,787 | 8,658,433,985 | 72,416,324,298 |
| 2002 | DQ     | 101,035,449,637 | 19,953,151,787 | 8,665,973,552 | 72,416,324,298 |
| 2004 | AG     | 101,055,095,260 | 19,953,151,787 | 8,685,619,175 | 72,416,324,298 |
| 2005 | AG     | 101,042,116,217 | 19,953,151,787 | 8,672,640,132 | 72,416,324,298 |
| 2006 | AG     | 101,039,254,204 | 19,953,151,787 | 8,669,778,119 | 72,416,324,298 |
| 2007 | LR     | 101,041,168,637 | 19,953,151,787 | 8,671,692,552 | 72,416,324,298 |
| 2008 | LR     | 101,051,694,982 | 19,953,151,787 | 8,682,218,897 | 72,416,324,298 |
| 2009 | LR     | 101,055,022,345 | 19,953,151,787 | 8,685,546,260 | 72,416,324,298 |
| 2010 | SB     | 101,080,139,060 | 19,953,151,787 | 8,710,662,975 | 72,416,324,298 |
| 2011 | SB     | 101,319,598,704 | 19,953,151,787 | 8,950,122,619 | 72,416,324,298 |

|                  |        | (1)       | (2)           | (3)          | (4)     | (5)     |  |  |  |  |
|------------------|--------|-----------|---------------|--------------|---------|---------|--|--|--|--|
| Panel A: Parcels |        |           |               |              |         |         |  |  |  |  |
| Year             | Source | Total     | Single-Family | Multi-Family | Vacant  | Other   |  |  |  |  |
| 1999             | DQ     | 2,220,095 | 1,418,218     | 445,884      | 175,673 | 179,318 |  |  |  |  |
| 2000             | DQ     | 2,226,745 | 1,422,309     | 448,362      | 176,446 | 178,424 |  |  |  |  |
| 2001             | DQ     | 2,232,094 | 1,426,138     | 449,673      | 178,323 | 176,744 |  |  |  |  |
| 2002             | DQ     | 2,239,496 | 1,431,767     | 451,753      | 177,975 | 176,840 |  |  |  |  |
| 2004             | AG     | 2,252,467 | 1,438,890     | 458,633      | 188,018 | 165,919 |  |  |  |  |
| 2005             | AG     | 2,263,834 | 1,443,309     | 465,230      | 188,135 | 166,199 |  |  |  |  |
| 2006             | AG     | 2,275,592 | 1,450,579     | 473,405      | 184,990 | 165,757 |  |  |  |  |
| 2007             | LR     | 2,291,112 | 1,457,071     | 483,336      | 184,411 | 165,456 |  |  |  |  |
| 2008             | LR     | 2,304,070 | 1,459,917     | 495,133      | 182,062 | 166,137 |  |  |  |  |
| 2009             | LR     | 2,312,035 | 1,461,376     | 503,333      | 180,483 | 166,048 |  |  |  |  |
| 2010             | SB     | 2,316,179 | 1,461,672     | 509,142      | 178,578 | 165,992 |  |  |  |  |
| 2011             | SB     | 2,321,904 | 1,463,549     | 512,388      | 178,865 | 166,307 |  |  |  |  |

### Table 6: Cleaned Dataset: Use Type By Year

#### Panel B: Land Area, Square Feet

| Year | Source | Total           | Single-Family  | Multi-Family  | Vacant         | Other          |
|------|--------|-----------------|----------------|---------------|----------------|----------------|
| 1999 | DQ     | 101,045,575,947 | 15,181,544,607 | 3,262,746,685 | 42,464,887,994 | 39,940,194,500 |
| 2000 | DQ     | 101,043,896,594 | 15,604,921,648 | 3,290,852,310 | 42,161,273,648 | 39,627,551,166 |
| 2001 | DQ     | 101,027,910,070 | 16,672,977,896 | 3,367,821,918 | 41,039,886,406 | 39,609,467,767 |
| 2002 | DQ     | 101,035,449,637 | 16,521,335,291 | 3,673,878,015 | 40,945,640,746 | 39,698,352,668 |
| 2004 | AG     | 101,055,095,260 | 16,776,334,754 | 3,702,529,345 | 41,432,839,634 | 39,101,202,122 |
| 2005 | AG     | 101,042,116,217 | 16,852,988,610 | 3,699,048,520 | 41,330,749,509 | 39,128,076,071 |
| 2006 | AG     | 101,039,254,204 | 16,855,731,407 | 3,768,149,464 | 41,267,359,016 | 39,129,931,492 |
| 2007 | LR     | 101,041,168,637 | 16,975,273,672 | 3,892,363,301 | 41,006,052,815 | 39,163,785,952 |
| 2008 | LR     | 101,051,694,982 | 17,045,615,339 | 3,919,442,451 | 40,889,968,790 | 39,193,243,887 |
| 2009 | LR     | 101,055,022,345 | 17,067,792,621 | 3,953,296,032 | 40,775,960,747 | 39,257,972,945 |
| 2010 | SB     | 101,080,139,060 | 17,079,013,486 | 3,972,852,109 | 40,772,913,715 | 39,255,359,749 |
| 2011 | SB     | 101,319,598,704 | 17,112,122,808 | 3,985,296,710 | 40,903,686,375 | 39,318,492,811 |

Note: Columns 2 through 5 do not add up to the total in column 1, as a very few parcels do not have an intelligible use code.